NATIONAL QUALITY FORUM

Resource Use Measure Evaluation 1.0
January 2011

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

Resource Use Definition:

- Resource use measures are broadly applicable and comparable measures of input counts—(in terms of units or dollars)-- applied to a population or population sample
- Resource use measures count the frequency of specific resources; these resource units may be monetized, as appropriate.
- The approach to monetizing resource use varies and often depends on the perspective of the measurer and those being measured. Monetizing resource use allows for the aggregation across resources.

NQF Staff: NQF staff will complete a preliminary review of the measure to ensure conditions are met and the form has been completed according to the developer's intent. Staff comments have been highlighted in green.

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

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

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

Evaluation ratings of the extent to which the subcriteria are met (TAP or Steering Committee)
High (H) – based on the information submitted, there is high confidence (or certainty) that the criterion is met
Moderate (M) – based on the information submitted, there is moderate confidence (or certainty) that the criterion is met
Low (L) - based on the information submitted, there is low confidence (or certainty) that the criterion is met
Insufficient (I) – there is insufficient information submitted to evaluate whether the criterion is met, e.g., blank, incomplete, or information is not relevant, responsive, or specific to the particular question (unacceptable)
Not Applicable (NA) - Not applicable (only an option for a few subcriteria as indicated)

Evaluation ratings of whether the measure met the overall criterion (Steering Committee)
Yes (Y)- The overall criteria has been met
No (N)-The overall criterion has NOT been met
High (H) – There is high confidence (or certainty) that the criterion is met
Moderate (M) – There is moderate confidence (or certainty) that the criterion is met
Low (L) - There is low confidence (or certainty) that the criterion is met

Recommendations for endorsement (Steering Committee)
Yes (Y) – The measure should be recommended for endorsement
No (N)-The measure should NOT be recommended for endorsement
Abstain (A)- Abstain from voting to recommend the measure

TAP/Workgroup Reviewer Name:
Steering Committee Reviewer Name:
Staff Reviewer Name(s):
NQF Review #: 1585 **NQF Project:** Endorsing Resource Use Standards- Phase II

BRIEF MEASURE INFORMATION
Measure Title: Episode of care for simple, non-specific lower back pain (acute and subacute)
Measure Steward (IP Owner): American Board of Medical Specialties Research and Education Foundation, 222 N. LaSalle St, Suite 1500, Chicago, Illinois, 60601
Brief description of measure: Resource use and costs associated with management of an episode-of-care for simple non-specific lower back pain. The episode is triggered by an initial ambulatory care visit for non-specific lower back pain (LBP). The episode lasts three months (90 days) from the time of the trigger ambulatory visit. An episode only begins if there are no LBP ambulatory care visits within 90 days prior to the initial LBP visit. Also, all individuals with a radiculopathy diagnosis during the measurement or prior period are excluded. LBP-related resource use and costs are measured during the episode, including 14 days prior to the initial visit that triggers the episode.
Resource use service categories: Inpatient services: Inpatient facility services Inpatient services: Evaluation and management Inpatient services: Procedures and surgeries Inpatient services: Imaging and diagnostic Inpatient services: Lab services Inpatient services: Admissions/discharges Ambulatory services: Outpatient facility services Ambulatory services: Emergency Department Ambulatory services: Pharmacy Ambulatory services: Evaluation and management Ambulatory services: Procedures and surgeries Ambulatory services: Imaging and diagnostic Ambulatory services: Lab services Durable Medical Equipment (DME)
Brief description of measure clinical logic: Resource use and costs associated with management of an episode-of-care for simple non-specific lower back pain. The episode is triggered by an initial ambulatory care visit for non-specific lower back pain (LBP). The episode lasts three months (90 days) from the time of the trigger ambulatory visit. An episode only begins if there are no LBP ambulatory care visits within 90 days prior to the initial LBP visit. Also, all individuals with a radiculopathy diagnosis during the measurement or prior period are excluded. LBP-related resource use and costs are measured during the episode, including 14 days prior to the initial visit that triggers the episode.
If included in a composite or paired with another measure, please identify composite or paired measure:
Subject/ Topic Areas: Musculoskeletal
Type of resource use measure: Cost/Resource Use
Data Type: Administrative claims Other

CONDITIONS FOR CONSIDERATION BY NQF	
Four conditions must be met before proposed measures may be considered and evaluated for suitability as voluntary consensus standards:	**NQF Staff**
A. Measure Steward Agreement. *The measure is in the public domain or an intellectual property (measure steward agreement) is signed. Public domain only applies to governmental organizations. All non-government organizations must sign a measure steward agreement even if measures are made publicly and freely available.*	**A** **Y☐** **N☐**

A.1.Do you attest that the measure steward holds intellectual property rights to the measure? (If no, do not submit) Yes **A.2. Please check if either of the following apply:** **A.3. Measure Steward Agreement.** Agreement signed and submitted **A.4. Measure Steward Agreement attached:** Signed_NQFMeasureSteward Agreement_020309-634386996326671655.pdf	
B. Maintenance. *The measure owner/steward verifies there is an identified responsible entity and process to maintain and update the measure on a schedule that is commensurate with the rate of clinical innovation, but at least every 3 years. (If no, do not submit)* Yes, information provided in contact section	**B** **Y☐** **N☐**
C. Purpose/ Use *(All the purposes and/or uses for which the measure is specified and tested:* Quality Improvement (Internal to the specific organization)	**C** **Y☐** **N☐**
D. Testing. *The measure is fully specified and tested for reliability and validity (See guidance on measure testing).* Yes, reliability and validity testing completed	**D** **Y☐** **N☐**
E. Harmonization and Competing Measures. *Have NQF-endorsed measures been reviewed to identify if there are related or competing measures? (List the NQF # and title in the section on related and competing measures)* Yes *E.1.***Do you attest that measure harmonization issues with related measure (either the same measure focus or the same target population) have been considered and addresses as appropriate? (List the NQF # and title in the section on related and competing measures)** No related measures *E.2.Do you attest that competing measures (both the same measure focus and the same target population) have been considered and addressed where appropriate?* No competing measures	**E** **Y☐** **N☐**
F. Submission Complete. *The requested measure submission information is complete and responsive to the questions so that all the information needed to evaluate all criteria is provided.*	**F** **Y☐** **N☐**
Have all conditions for consideration been met? **Staff Notes to Steward** (*if submission returned*)**:**	**Y ☐** **N ☐**

Staff Notes to Reviewers (*issues or questions regarding any criteria*)**:**

File Attachments Related to Measure/Criteria:
Attachment:
Attachment: S5_Data Dictionary-634349417472798395.pdf
Attachment:
Attachment:
Attachment:
Attachment:
Attachment:
Attachment:
Attachment: 10.1_Risk adjustment method-634349436185183385.pdf
S12_sample score report LBP.pdf
Attachment: SA_Reliability_Validity Testing LBP Simple.pdf

IMPORTANCE TO MEASURE AND REPORT	
Extent to which the specific measure focus is important to making significant gains in health care quality (safety, timeliness, effectiveness, efficiency, equity, patient-centeredness) and improving health outcomes for a specific high impact aspect of healthcare where there is variation in performance. **Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All subcriteria must be met to pass this criterion.**	**Eval Rating**
High Impact **IM1. Demonstrated high impact aspect of healthcare:** Affects large numbers High resource use **IM1.1. Summary of evidence of high impact:** Low back pain contributes a major burden for individuals and society (1). According to data from the second National Health and Nutrition Examination Survey (NHANES II), it is the second most common symptom-related reason for visits in the United States (2). According to this survey, most people with LBP sought care from general practitioners, with orthopaedists and chiropractors being the next most common sources of care. In the 2002 US National Health Interview Survey (NHIS), the prevalence of low back pain was 25%, defined as lasting at least a whole day in the prior three months (3). The total costs of low-back pain in the United States exceed $100 billion a year (4). Two-thirds of this cost is indirect, due to lost wages and reduced productivity. Likewise, back pain has a huge impact on lifestyle and quality of life. The Institute of Medicine and AQA have identified LBP as one of 20 conditions that should be considered priority areas in need of quality improvement based on its relevance to a significant volume of patients, its impact on those patients, and the perception of opportunity to significantly improve the quality of related care. LBP had also been previously identified as a priority area in other national initiatives including the U.S. Department of Veterans Affairs' Quality Enhancement Research Initiative (5). In addition, LBP episodes are increasingly high-resource episodes in large part because of increasing utilization (and over-utilization) of imaging services during the diagnostic process, as has been widely publicized in recent years. For example, a 2008 study of approximately 45,000 patients with back problems demonstrated that the average costs of their treatment increased 65 percent between 1997 and 2005, considerably faster than the costs of medical services in general. Most of this increase could be attributed to growth in imaging (6). Risk factors for low back pain include: smoking, obesity, older age, female gender, sedentary work, physically and psychologically strenuous work, workers' compensation insurance, low educational attainment, job dissatisfaction, anxiety and depression (7,8,9). A prospective study on two general medicine practices in south Manchester by Croft et al showed that 90% of patients with low back pain will have stopped seeking care after three months (10). Low back pain has a benign course in 90% of patients (11). LBP is a condition that, depending on the presence or absence of radiculopathy, can be treated in very different ways. The prevalence of lumbosacral radiculopathy, or sciatica is approximately 3-5% distributed equally between men and women (12). Approximately 13% of the employed population suffer from lumbar complaints--out of these, 11% suffer from lumbar radiculopathy (13). Costs for treatment range from near $4,700 for conservative therapy to over $42,000 per	**1a** **H☐** **M☐** **L☐** **I☐**

patient for operative treatment (14). Lumbar discectomy for radiculopathy is the most commonly performed spinal procedure in the US with approximately 300,000 cases each year (14). A single surgery isn't always effective. Recurrent lumbar disc herniation occurs in a subset of patients in over 10% of patients and is associated with substantial additional health care costs (15).
The ABMS REF has developed two measures—one designed to observe variation in resource use for patients presenting with LBP without radiculopathy and separate measure for patients experiencing radiculopathy with or without LBP.

IM1.2. Citations for evidence of high impact cited in IM1.1.:

1.Williams CM, Maher CG, Hancock MJ, et al. Low back pain and best practice care: A survey of general practice physicians. Arch Intern Med. Feb 8 2010;170(3):271-277.
2.Deyo RA, Tsui-Wu YJ. Descriptive epidemiology of low-back pain and its related medical care in the United States. Spine (Phila Pa 1976). Apr 1987;12(3):264-268.
3.Deyo RA, Mirza SK, Martin BI. Back pain prevalence and visit rates: estimates from U.S. national surveys, 2002. Spine (Phila Pa 1976). Nov 1 2006;31(23):2724-2727.
4.Katz JN. Lumbar disc disorders and low-back pain: socioeconomic factors and consequences. J Bone Joint Surg Am. Apr 2006;88 Suppl 2:21-24.
5. Priority Areas for National Action: Transforming Health Care Quality. Institute of Medicine. Karen Adams and Janet Corrigan Editors. March 10, 2003.
6. Martin, Brook, Deyo, Richard, Mirza, Sohail. Expenditures and Health Status Among Adults with Back and Neck Problems. JAMA 2008: 299(6): 656-664
7.Cassidy JD, Carroll LJ, Cote P. The Saskatchewan health and back pain survey. The prevalence of low back pain and related disability in Saskatchewan adults. Spine (Phila Pa 1976). Sep 1 1998;23(17):1860-1866; discussion 1867.
8.Lean ME, Han TS, Seidell JC. Impairment of health and quality of life using new US federal guidelines for the identification of obesity. Arch Intern Med. Apr 26 1999;159(8):837-843.
9.Frank J, Sinclair S, Hogg-Johnson S, et al. Preventing disability from work-related low-back pain. New evidence gives new hope--if we can just get all the players onside. CMAJ. Jun 16 1998;158(12):1625-1631.
10.Croft PR, Macfarlane GJ, Papageorgiou AC, Thomas E, Silman AJ. Outcome of low back pain in general practice: a prospective study. BMJ. May 2 1998;316(7141):1356-1359.
11.Bach SM, Holten KB. Guideline update: what´s the best approach to acute low back pain? J Fam Pract. Dec 2009;58(12):E1.
12. Tarulli AW, Raynor EM. Lumbosacral radiculopathy. Neurologic Clinics 2007;25:387-405.
13. Jordon J, Konstantinou TS, Weinstein J. Herniated lumbar disc, Clin Evid 2007.
14. Parker L, Risheng X, McGirt MJ, et al. Long-term back pain after single level discectomy for raiculopathy: incidence and health care cost analysis. J Neurosurg Spine 2010:12;178-182.
15. Ambrossi GL, McGirt MJ, SciubbaDM et al. Recurrent lumbar disc herniation after single-level lumbar discectomy: incidence and health care cost analysis. Neurosurgery 2009;65:574-8.

IM2. Opportunity for Improvement

IM2.1. Briefly explain the benefits envisioned by use of this measure:

To identify actionable information on the underlying causes of differences in patterns of care for episodes of simple non-specific acute/subacute LBP, it is useful to examine resource use and costs during an episode of care. If results from these analyses can provide clear and actionable information on which components of care can (or should) be reduced and which components of care can (or should) be increased, this information can help reduce spending while maintaining or even improving clinical quality and outcomes. This measure can be used to identify and, if necessary, address unwarranted variability in the resources used to treat LBP patients on an outpatient basis. In addition where gaps in utilization occur leading to suboptimal quality, education and care coordination can implemented.

1b

IM2.2. Summary of data demonstrating variation across providers or entities:

--Chou, et al. analyzed randomized controlled trials examining the effects of routine, immediate lumbar imaging versus usual clinical care without immediate imaging on clinical outcomes in patients with low-back pain and no indication of serious underlying conditions. These trials reported pain or function, quality of life, mental health, overall patient-reported improvement, and patient satisfaction. The investigators did not find any evidence of significant differences between immediate lumbar imaging and usual care without immediate imaging for primary outcomes at either short-term or long-term follow-up leading the authors to conclude that lumbar imaging for low-back pain without indications of serious underlying conditions does not improve clinical outcomes and recommend clinicians refrain from routine,

H☐
M☐
L☐
I☐

immediate lumbar imaging in patients with acute or subacute low-back pain and without features suggesting a serious underlying condition (1).

--Wilson et al. examined patients presenting for outpatient treatment for low back pain and found a significant association between patients' perceived need for radiological studies and use of services. The investigators concluded that differences in physicians' adherence to guidelines regarding radiology utilization may in part reflect variations in patient's perceived need for those services. (2)
--In a recent study by Gelhorn et al, the investigators note that wide practice variations exist in the treatment of acute low back pain and use of physical therapy (PT) as part of early treatment. The investigators studied the relationship between early PT for acute low back pain and subsequent use of lumbosacral injections, lumbar surgery, and frequent physician office visits for low back pain. The investigators found a lower risk of subsequent medical service usage among patients who received PT early after an episode of acute low back pain relative to those who received PT at later times. The investigators also noted medical specialty variations exist regarding early use of PT, with potential underutilization among generalist specialties (3).

--In 2007, Fritz et al examined the association between adherence to the guideline recommendation to use active versus passive treatments with clinical outcomes and costs for patients with acute low back pain receiving physical therapy and found adherence to the guideline recommendation for active care was associated with better clinical outcomes and reduced cost (4).

--In 1994, Volinn et al used data from the National Hospital Discharge Survey to examine patterns in both surgical and nonsurgical low back pain (LBP) hospitalizations among geographic regions in the US (1979-1987). The investigators found rates of both surgical and nonsurgical LBP hospitalization varied twofold among regions of the U.S., and average lengths of stay for these types of hospitalization varied considerably as well. The U.S. rate of LBP surgery increased sharply during the period covered by the study. Over the same time, the U.S. rate of nonsurgical LBP hospitalization declined. To explore the cost implications of regional variations, the investigators estimated how much would be saved if all regions of the U.S. had the same rates of surgical and nonsurgical LBP hospitalization as the region with the lowest rates and shortest average length of stay concluding over $500 million in health care costs would be saved (5).

--Watters and McGirt conducted an evidence-based review of the clinical literature supporting the performance of conservative versus aggressive technique for discectomy for the treatment of lumbar disc herniation with radiculopathy and found fair quality evidence that conservative discectomy may result in shorter operative times and a quicker return to work despite similar lengths of hospital stay, similar pain levels and similar 2 year incidence of persistent back and leg pain. However, there is also fair quality evidence that conservative discectomy will result in higher incidence of recurrent disc herniation (5).
-- Using Medicare part B claims data, Carrino et al, evaluated spinal injection procedures for trends in volume, reimbursement, and physician specialty participation finding that spinal injection volume and reimbursement have increased substantially in the Medicare population from 1993 to 1999. During this interval, radiologist participation has decreased with non-radiologists performing most spinal injection procedures (6).

--Injection therapy is well established in the treatment of lumbar radiculopathy and for years has been performed without image guidance. More recently, minimally invasive imaging guided techniques are being used (CT, fluoroscopy and ultrasound) to increase the precision of spinal injection and improve the success rates of pain management (6, 7, 8).

IM2.3. Citations for data on variation:

1.Chou R, Fu R, Carrino JA, Deyo RA. Imaging strategies for low back pain: systematic review and meta-analysis. Lancet 2009;373:463-72.
2.Wilson IB, Dukes K, Greenfield S, et al. Patients' role in the use of radiology testing for common office practice complaints. Arch Intern Med 2001:161:256-63.
3.Gellhorn AC, Chan L, Martin B, Friedly J. Management Patterns in Acute Low Back Pain: The Role of Physical Therapy. Spine;2010 Nov 19. [Epub ahead of print].
4.Fritz JM, Cleland JA, Brennan GP. Does adherence to the guideline recommendation for active treatments improve the quality of care for patients with acute low back pain delivered by physical therapists? Med Care 2007;45:973-80.
5.Volinn E, Turczyn KM, Loeser JD. Patterns in low back pain hospitalizations: implications for the treatment of low back pain in an era of health care reform. Clin J Pain 1994;10:64-70.
6.Watters WC, McGirt MJ. An evidence-based review of the literature on the consequences of conservative versus aggressive discectomy for the treatment of primary disc herniation with radiculopathy. The Spine Journal 2008;9:240-

257.
7.Carrino JA, Morrison WB, Parker L, et al. Spinal injection procedures: volume, provider distribution, and reimbursement in the U.S. medicare population from 1993 to 1999. Radiology. 2002;225:723-9.
8.Kim PS. Role of injection therapy: review of indications for trigger point injections, regional blocks, facet joint injections, and intra-articular injections. Curr Opin Rheumatol 2002;14:52-57.
9.Loizides A, Siegfried P, Plaikner M, et al. Ultrasound-guided injections in the lumbar spine. Medical Ultrasonography 2011;13:54-58.

IM2.4. Summary of data on disparities by population group:

Racial, ethnic and gender disparities in the treatment of pain is a complex issue with little research to date focused specifically on disparities and low back pain. Research to date offers mixed findings.

--In a study conducted by Carey et al, four strata of randomly selected health care providers in North Carolina (primary care MDs, Doctors of Chiropractic, orthopedic surgeons, and group model HMO primary care providers) enrolled 1633 consecutive patients with low back pain into a cohort study to determine whether race had an independent effect on rate of recovery from low back pain, and whether there was any racial disparity in the treatments provided to patients with low back pain. The findings showed that the blacks had somewhat higher pain and disability than the whites at the time of their presentation to health care providers, and that this somewhat greater level of disability continued throughout the recovery phase of their low back pain episode. This contrasted with the providers' assessment of these patients' condition. The providers assessed them as having less pain and as being less likely to have a serious underlying cause of back pain. Similar to studies in other conditions, black patients were treated less intensively in terms of both diagnostic and therapeutic interventions (1).

--Selim et al. analyzed data from the Veterans Health Study to determine whether clinical differences and/or race account for disparities between white and nonwhite patients in the use of lumbar spine radiographs and found there was greater use of lumbar spine radiographs by nonwhite patients compared with white patients. This remained true when patients were subcategorized by severity of low back pain. However, race had no influence when multiple clinical characteristics of the patients were controlled for simultaneously. The investigators conclude the study demonstrates the importance of careful and comprehensive case-mix adjustment when assessing apparent differences in the use of medical services (2).

--In a retrospective cohort study of 20,125 adults who presented to 2 urban EDs with back or abdominal pain over a 4-year period, Mills et al, assessed whether patient race affects analgesia administration for patients presenting with back or abdominal pain and demonstrated nonwhite patients were significantly less likely to receive analgesia and specifically opiate analgesia and waited longer for their opiate medication than white patients when presenting to the ED for these complaints despite higher pain scores (3).

--In a retrospective review comparing physician workup of degenerative lumbosacral pathologies between different genders and ethnic groups, researchers identified 5690 patients with degenerative lumbosacral pathologies and found that although females were more likely than males to have imaging tests ordered, male patients were significantly more likely to have surgery recommended than females; nonwhite females were 52% less likely to have surgery offered at initial visit, as compared to white males; more imaging tests were ordered or reviewed among whites than among any other ethnic group; and white and asian patients were significantly more likely to have surgery recommended or prescribed than black and hispanic patients. The investigators conclude the study findings suggest that ethnicity and gender affect the workup and surgical management of degenerative spinal disorders. However, they also note that there are a number of confounding factors that were not identified in their database, including managed care and insurance status and cultural differences, which may affect both test ordering and treatment recommendations and thus recommend further study of bias in clinical decision-making (4).

IM2.5. Citations for data on disparities cited in IM2.4:

1.Carey TS, Garrett JM. The relation of race to outcomes and the use of health care services for acute low back pain. Spine 2003;28:390-4.
2.Selim AJ, Fincke G, Ren XS, Deyo RA, Lee A, Skinner K, Kazis L. Racial differences in the use of lumbar spine radiographs: results from the Veterans Health Study. Spine (Phila Pa 1976). 2001 Jun 15;26(12):1364-9.
3.Mills AM, Shofer FS, Boulis AK, Holena DN, Abbuhl SB. Racial disparity in analgesic treatment for ED patients with abdominal or back pain. Am J Emerg Med. 2010 Apr 30. [Epub ahead of print].
4.Taylor BA, Casas-Ganem J, Vaccaro AR, Hilibrand AS, Hanscom BS, Albert TJ. Differences in the work-up and

treatment of conditions associated with low back pain by patient gender and ethnic background. Spine 2005;30:359-64.	
IM3. Measure Intent **IM3.1. Describe intent of the measure and its components/ Rationale (including any citations) for analyzing variation in resource use in this way** While documentation of regional variability in the overall costs of care reveals that inefficiencies exist in the healthcare system, it does not provide actionable information on the underlying causes of these differences or how they can be reduced. One potential solution is to focus on episode-based resource use and costs so that variations within a particular clinical area can be examined and areas of variability can be optimized. Moreover, episode-based resource measures can be combined with surrogate measures of quality care to identify highly efficient care where quality is high and costs are low. With this information, all parties involved (consumers, purchasers, and providers) can optimize treatment decisions that affect the balance of costs and quality of care.	**1c** **H☐** **M☐** **L☐** **I☐**
IM4. Resource use service categories are consistent with measure construct *Refer to IM3.1. & all S9 items to evaluate this criteria.*	**1d** **H☐** **M☐** **L☐** **I☐**
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Importance to Measure and Report?*	
Steering Committee: Was the threshold criterion, *Importance to Measure and Report*, **met?** **Rationale:**	**Y☐** **N☐**

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.	
MEASURE SPECIFICATIONS	
S1. Measure Web Page: *Do you have a web page where current detailed measure specifications can be obtained?* Yes http://www.healthqualityalliance.org/hvhc-project/cost-care-measurement-development	**Eval Rating 2a1/2b1**
S2. General Approach *If applicable, summarize the general approach or methodology to the measure specification. This is most relevant to measures that are part of or rely on the execution of a measure system or applies to multiple measures.*	

The ABMS REF episode-based resource use measures were created in an open and transparent manner with input from a wide range of clinical experts, methodologists, health care economists and other stakeholders. The measure development process involved a series of deliberate steps where participating clinicians took into account the natural progression of a condition and existing best practices before carefully considering how to best use administrative claims data to construct the episode. They aimed to identify clinically homogenous populations so that the measures would be sensitive to provider decisions and existing practice protocols for like patients. Workgroup members were then asked to conceptualize the measure specifications based on their combined knowledge of guidelines, evidence, and clinical experience. The workgroups helped to define the denominator, duration, clinically relevant services and attribution of each episode as related to the clinical progression and treatment of the condition. Project staff then worked to translate the concepts into detailed written measure specifications and test the measures on a commercial database. The workgroups subsequently re-convened via a series of conference calls to review data analyses, share expert opinions, consider additional evidence-based literature, revise and finalize the measure specifications. Each measure was developed independently and, as such, they are not summative.

Attachment:

S3. Type of resource use measure:

Per episode

S4. Target Population:

Adult/Elderly Care

S4.1. Subject/Topic Areas:

Musculoskeletal

S4.2. Cross Cutting Areas *(HHS or NPP National health goal/priority)*

Care Coordination

S5. Data dictionary or code table
Please provide a web page URL or attachment if exceeds 2 pages. NQF strongly prefers URLs. Attach documents only if they are not available on a web page and keep attached file to 5MB or less.

Data Dictionary:

URL:
Please supply the username and password:
Attachment: S5_Data Dictionary-634349417472798395.pdf

Code Table:

URL:
Please supply the username and password:
Attachment:

S6.Data Protocol (Resource Use Measure Module 1)
The measure developer must determine which of the following data protocol steps: data preparation, data inclusion criteria, data exclusion criteria, and missing data, are submitted as measure specifications or as guidelines. Specifications limit user options and flexibility and must be strictly adhered to; whereas guidelines are well thought out guidance to users while allowing for user flexibility. If the measure developer determines that the requested specification approach is better suited as guidelines, please select and submit guidelines, otherwise specifications must be provided.

Data Protocol Supplemental Attachment or URL:
If needed, attach document that supplements information provided for data protocol for analysis, data inclusion criteria, data exclusion criteria, and missing data (Save file as: S6_Data Protocol).

All fields of the submission form that are supplemented within the attachment must include a summary of important information included in the attachment and its intended purpose, including any references to page numbers, tables, text, etc.

URL: http://www.healthqualityalliance.org/hvhc-project/cost-care-measurement-development
Please supply the username and password:
Attachment:

S6.1. Data preparation for analysis
Detail (specify) the data preparation steps and provide rationale for this methodology.

Guidelines : Approach to Data Cleaning:
If a standardized cleaning methodology or logic for the claims data exists, users are encouraged to apply the existing methodology, or conversely, encouraged not to remove data cleaning steps already implemented. If however, organizations impute missing data, we recommend using only non-imputed data.

Rationale: Each organization will be more familiar with the nature of their data therefore any standard cleaning procedures are likely to be appropriate. Imputation can produce unpredictable biases in the results.

S6.2.Data inclusion criteria
Detail initial data inclusion criteria and rationale(related to claim-line or other data quality, data validation, e.g. truncation or removal of low or high dollar claim)

Guidelines : Paid claims with non-missing enrollee identification numbers, primary procedure and diagnosis codes should be included in the measure.
Note: The ABMS REF resource use measures are constructed based on date of service, not date of payment. Therefore, we recommend applying the measures to finalized or "closed" datasets so that complete claims histories during the measurement period are captured in the data.
Including enrollees with at least 24 months of continuous medical and pharmacy benefit enrollment during the identification year and the measurement year is recommended. However, the measure has been tested on enrollees with at least 320 total days of coverage during each year. If precise information regarding persons' total days of coverage is not available, it is recommended that measure implementers estimate this information to the best of their ability using available data elements (e.g., monthly enrollment indicators). This approach is based on the similar eligibility requirements used by NCQA for HEDIS measure denominators.

S6.3. Data exclusion criteria
Detail initial data exclusion criteria and rationale (related to claim-line or other data quality, data validation, e.g. truncation or removal of low or high dollar claim)

Guidelines : Beyond the standard data cleaning steps, we recommend that claim lines with missing or zero quantity values be set to a quantity of one and claim lines missing enrollee identification variables, primary diagnosis and procedure codes, and service date be eliminated. We also recommend eliminating all rejected or unpaid claims. Because a single provider id could have multiple specialties, we also recommend generating a uniform specialty for all providers by assigning each provider the specialty which is most frequently observed from all their Evaluation and Management visits.

Rationale: Converting missing or zero quantities to a minimum value of 1 allows for the pricing of these services.
Claim lines missing enrollee identifiers, or primary procedure and diagnosis codes cannot be attributed to an individual, and without procedure and diagnosis codes, services cannot be properly identified and categorized. The resource use measures are intended to track costs to the payer, not general or societal costs, so rejected or unpaid claims should be eliminated.
Standardizing the specialty of all providers eliminates the possibility that providers are classified as one specialty for one enrollee and another specialty for others.

S6.4. Missing Data
Detail steps associated with missing data and rationale(e.g., any statistical techniques used)

Guidelines : Users are encouraged to eliminate claim lines missing enrollee identification variables or primary

procedure and diagnosis codes. We do not recommend using any imputation methods to replace missing data.

Rationale: Claim lines missing enrollee identifiers cannot be attributed to an individual, and without procedure and diagnosis codes, services cannot be properly identified and categorized. Imputation of missing information could introduce bias into the measure, so we do not recommend the use of imputed data.

S7. Data Type: Administrative claims
Other

S7.1. Data Source or Collection Instrument
Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc.)

Sources for administrative claims: commercial databases
Standardized price tables: Users can download tables from the NCQA website (see url below) or use the guidelines in the technical appendix of the written measure specification to create their own standardized prices.

S7.2. Data Source or Collection Instrument Reference
(Please provide a web page URL or attachment). NQF strongly prefers URLs. Attach documents only if they are not available on a web page and keep attached file to 5MB or less)

URL: http://www.ncqa.org/tabid/1092/Default.aspx
Please supply the username and password:
Attachment:

S8.Measure Clinical Logic (Resource Use Measure Module 2)
The measure's clinical logic includes the steps that identify the condition or event of interest and any clustering of diagnoses or procedures. For example, the diagnoses and procedures that qualifies for a cardiac heart failure episode, including any disease interaction, comorbid conditions, or hierarchical structure to the clinical logic of the model. (Some of the steps listed separately below may be embedded in the risk adjustment description, if so, please indicate NA and in the rationale space list 'see risk adjustment details.')

Clinical Logic Supplemental Attachment or URL:
If needed, provide a URL or document that supplements information provided for the clinical framework, co-morbid interactions, clinical hierarchies, clinical severity levels, and concurrency of clinical events

URL: http://www.healthqualityalliance.org/hvhc-project/cost-care-measurement-development
Please supply the username and password:
Attachment:

S8.1. Brief Description of Clinical Framework
Briefly describe your clinical logic approach including clinical topic area, whether or not you account for comorbid and interactions, clinical hierarchies, clinical severity levels and concurrency of clinical events.

Resource use and costs associated with management of an episode-of-care for simple non-specific lower back pain. The episode is triggered by an initial ambulatory care visit for non-specific lower back pain (LBP). The episode lasts three months (90 days) from the time of the trigger ambulatory visit. An episode only begins if there are no LBP ambulatory care visits within 90 days prior to the initial LBP visit. Also, all individuals with a radiculopathy diagnosis during the measurement or prior period are excluded. LBP-related resource use and costs are measured during the episode, including 14 days prior to the initial visit that triggers the episode.

S8.2. Clinical framework
Detail any clustering and the assignment of codes, including the grouping methodology, the assignment algorithm, and relevant codes and rationale for these methodologies.

The following steps are used to create the clinical framework for the measure.

Identify the measure population

Step 1: Identify patients that have a diagnostic code for an ambulatory care visit (including E&M visits and specified chiropractic and physical therapy visits) during the event identification year (see also Table LBP-A and LBP-A1). These ICD-9 codes may be present in any diagnostic field, regardless of corresponding CPT and UB revenue codes. Diagnosis Codes: Lumbago: ICD9: 724.2; Backache, NOS: ICD9: 724.5; Other back symptoms: ICD9: 724.8; Back disorders, NOS: ICD9: 724.9; Somatic disfunction, lumbar region:ICD9: 739.3; Somatic disfunction, sacral region: ICD9: 739.4; Sprains and strains of lumbar: ICD9: 847.2; Office or Other Outpatient Services: CPT: 99201–99215; Chiropractic-specific codes: CPT: 98940-98942; Physical therapy-specific codes: CPT: 97110, 97112, 97113, 97124, 97140; Hospital Observation Services: CPT: 99217–99220; Hospital Inpatient Services: CPT: 99221–99239; Consultations: CPT: 99241–99275; Critical Care and Intensive Care Services: CPT: 99289–99298; Nursing Facility, Domiciliary and Home Services: CPT: 99301–99350; Case Management Services and Care Plan Oversight Services: CPT: 99361–99380; Preventive Medicine Services: CPT: 99381–99429; Other E&M Services: CPT: 99450–99456, 99354–99357

Step 2: Identify patients that meet age, eligibility and continuous enrollment criteria

1. Age: Identify patients aged 18 to 64.

2. Eligibility
a. Identify benefits during both the measurement period and prior period.
b. To be included persons must have both of the following benefits in both periods
i. Medical benefit
ii. Pharmacy benefit

3. Continuous enrollment
c. Determine enrollment during both the measurement and prior periods.
d. Identify (or estimate) total days of coverage in prior year
e. To be eligible, persons must have at least 320 total days of coverage during prior year and persons must be fully covered during measurement period.

Step 3: Identify patients with exclusion criteria

Exclusion criteria:

- Patient had LBP diagnosis (see Table LBP-A) for ambulatory visit (see Table LBP-A1 for included CPT codes) within 90 days prior to potential trigger LBP visit: Diagnosis Codes: Lumbago: 724.2; Backache, NOS: 724.5; Other back symptoms: 724.8; Back disorders, NOS: 724.9; Somatic disfunction, lumbar region: 739.3; Somatic disfunction, sacral region: 739.4; Sprains and strains of lumbar: 847.2
  CPT Codes: Office or Other Outpatient Services: 99201–99215; Chiropractic-specific codes: 98940-98942; Physical therapy-specific codes: 97110, 97112, 97113, 97124, 97140; Hospital Observation Services: 99217–99220; Hospital Inpatient Services: 99221–99239; Consultations: 99241–99275; Critical Care and Intensive Care Services: 99289–99298; Nursing Facility, Domiciliary and Home Services: 99301–99350; Case Management Services and Care Plan Oversight Services: 99361–99380; Preventive Medicine Services: 99381–99429; Other E&M Services: 99450–99456, 99354–99357

- Patient had fusion surgery, other back surgery, or fracture (see Table LBP-E) within 6 months prior to potential trigger LBP visit: Fusion Surgery: CPT: 22840,22851,22630,22612,22614; Other Back Surgery: CPT: 63001- 63051; Fracture (recent trauma codes): ICD9: 800, 805, 806, 839, 850-854, 860-869, 905-909, 926.11, 926.12, 929, 952, 958-959

- Patient has radiculopathy diagnosis (see Table LBP-F1) during measurement or prior period: Lumbosacral spondylosis without myelopathy: ICD9: 721.3; Spondylosis of unspecified site: ICD9: 721.9; Lumbar disc displacement w/o myelopathy: ICD9: 722.1; Degeneration of thoracic or lumbar intervertebral disc: ICD9: 722.5; Sciatica: ICD9: 724.3; Back pain with radiation, unspecified: ICD9: 724.4

- Patient has coincident UTI or sacroiliatis diagnosis on trigger claim (see Table LBP-F3) – a claim with these

Eval Rating 2a1

H☐ M☐ L☐ I☐

Eval Rating 2b1

H☐

diagnostic codes cannot be a trigger event: UTI: ICD9: 599.0; Sacroiliitis: ICD9: 720.7

M☐
L☐
I☐

- Patients meets one or more of the following exclusion criteria during the identification OR the measurement year (see Tables LBP-F2, F5-F7): active cancer (excluding melanoma, skin, prostate, and chronic lymphocytic leukemia): ICD9: 140-171, 174-184, 187-203, 204.0, 204.2, 204.8, 205-208, 230-239 WITH CPT: 38230, 38240-38242, 77261-77799, 79000-79999, 96400-96549; ICD-9-CM Procedure: 41.0, 41.91, 92.2; UB Revenue 028x, 033x, 0342, 0344, 0973; end stage renal disease (ESRD) including renal dialysis: CPT36145, 36800-36821, 36831-36833, 90919-90921, 90923-90925, 90935, 90937, 90939, 90940, 90945, 90947, 90989, 90993, 90997, 90999, 99512; HCPCS: G0257, G0311-G0319, G0321-G0323, G0325-G0327, G0392, G0393, S9339;ICD-9-CM Diagnosis:585.5, 585.6, V42.0, V45.1, V56; ICD-9-CM Procedure: 38.95, 39.27, 39.42, 39.43, 39.53, 39.93, 39.94, 39.95, 54.98; UB Revenue: 080x, 082x-085x, 088x ; UB Type of Bill: 72x; POS: 65; organ transplant: CPT: 32850-32856, 33930-33945, 44132-44137, 44715-44721, 47133-47147, 48160, 48550-48556, 50300-50380; HCPCS: S2152, S2053-S2055, S2060, S2061, S2065; ICD-9-CM Procedure: 33.5, 33.6, 37.5, 41.94, 46.97, 50.5, 52.8, 55.6; UB Revenue: 0362, 0367, 0810-0813, 0819; HIV/AIDS: ICD-9: 042; Neurological impairment: ICD9: 344.60, 729.2; Intraspinal abscess : ICD9: 324.1 324.9; Thoracic or Lumbar Sondylosis with Myelopathy (progressive symptoms) : ICD9: 721.4; Intervertebral Disk Disorder with Myelopathy (progressive symptoms): ICD9: 722.7; IV Drug Abuse: ICD9: 304.0,304.1x,304.2x,304,4x,305.4x,305.5x,305,6x,305.7x

Step 4: Combine prior steps to identify measure population

1. Identify LBP eligible population
2. Exclude those patients not meeting general inclusion criteria (e.g. age, continuous eligibility)
3. Exclude those patients meeting one or more measure exclusion criteria
4. The resulting collection of patients is the measure population

Eligible Event Identification:

For each individual in the measure population, identify the low back pain-related claims for services rendered during the measurement year. Claims / encounters will be identified based on the presence of LBP-related diagnosis codes or procedure codes. These events will be used to determine the LBP-related resource use.

Inpatient hospitalization events

Identify all inpatient claims / encounters with a LBP-related diagnostic code appearing in the primary diagnosis field only (see Table LBP-B1). Also identify any inpatient claims with surgery CPT codes listed in Table LBP-B2 regardless of diagnostic code: Lumbosacral spondylosis without myelopathy: ICD9: 721.xx; Spondylosis of unspecified site: ICD9: 722.xx; Lumbar disc displacement w/o myelopathy: ICD9: 724.xx; Somatic dysfunction lumbar region: ICD9: 739.3; Somatic dysfunction sacral region: ICD9: 739.4; Degeneration of thoracic or lumbar intervertebral disc: ICD9: 847.2

Outpatient events

Identify all outpatient claims / encounters with a LBP-related diagnostic code appearing in any position (see Table LBP-B1): Lumbosacral spondylosis without myelopathy: ICD9: 721.xx; Spondylosis of unspecified site: ICD9: 722.xx; Lumbar disc displacement w/o myelopathy: ICD9: 724.xx; Somatic dysfunction lumbar region: ICD9: 739.3; Somatic dysfunction sacral region: ICD9: 739.4; Degeneration of thoracic or lumbar intervertebral disc: ICD9: 847.2; Fusion Surgery: CPT: 22840, 22851, 22630, 22612, 22614; Other Back Surgery: CPT: 63001- 63051

Prescription drugs

Identify the low back pain-related medications by therapeutic class or generic/brand medication name during the measurement period (see Tables LBP-C and LBP-D): Analgesics: APAP/caffeine/dihydrocodeine, acetaminophen-codeine, acetaminophen-hydrocodone, acetaminophen-oxycodone, acetaminophen-pentazocine, acetaminophen-propoxyphene, acetaminophen-tramadol, buprenorphine, butorphanol, fentanyl , hydrocodone-ibuprofen, hydromorphone, ibuprofen-oxycodone , levorphanol , meperidine, meperidine-promethazine, methadone, morphine, nalbuphine, naloxone-pentazocine, oxycodone, oxymorphone, pentazocine, propoxyphene, tramadol, ziconotide, oxicontin; Corticosteroids: methylprednisolone , prednisolone, prednisone; Cox-2 inhibitors: celecoxib; Muscle relaxants: carisoprodol, chlorzoxazone, cyclobenzaprine, diazepam, metaxalone, methocarbamol, orphenadrine; NSAIDs: diclofenac, etodolac, fenoprofen, flurbiprofen, ibuprofen, diclofenac (volatren and flector), ketoprofen,

ketorolac, meclofenamate, mefenamic acid, meloxicam, nabumetone, naproxen, oxaprozin, sulindac, tolmetin; Other: lydoderm, duloxetine (cymbalta), gabapentin (neurontin), Pregabalin (lyrica); Additional J-codes to identify injections: J0592: Injection, buprenorphine hydrochloride, 0.1 mg; J0595: Injection, butorphanol tartrate, 1 mg; J1020: Injection, methylprednisolone acetate, 20 mg; J1030: Injection, methylprednisolone acetate, 40 mg; J1040: Injection, methylprednisolone acetate, 80 mg; J1094: Injection, dexamethasone acetate, 1 mg; J1100: Injection, dexamethasone sodium phosphate, 1mg; J1170: Injection, hydromorphone, up to 4 mg; J1885: Injection, ketorolac tromethamine, per 15 mg; J2175: Injection, meperidine hydrochloride, per 100 mg; J2180: Injection, meperidine and promethazine hcl, up to 50 mg; J2270: Injection, morphine sulfate, up to 10 mg; J2271: Injection, morphine sulfate, 100mg; J2275: Injection, morphine sulfate (preservative-free sterile solution),per10mg; 2300: Injection, nalbuphine hydrochloride, per 10 mg; J2310: Injection, naloxone hydrochloride, per 1 mg; J2410: Injection, oxymorphone hcl, up to 1 mg; J2650: Injection, prednisolone acetate, up to 1 ml; J2920: Injection, methylprednisolone sodium succinate, up to 40 mg; J2930: Injection, methylprednisolone sodium succinate, up to 125 mg; J0670: Injection, mepivacaine hydrochloride, per 10 ml

Rationale:
Rationale for cluster, grouping and assignment framework
Age: The measure includes individuals 18-64 years of age at the time of the qualifying event. Eliminating children follows from the fact there are relatively few cases under 18 and they are likely to be mainly due to injuries. The workgroup believed that people over age 65 may be treated differently and since the Marketscan data was limited to people under age 65, the measure was also limited to this group.
Trigger Outpatient Visit: To be included in this measure patient must have an ambulatory visit with a lower back pain diagnosis not indicative of radiculopathy. Patients with radiculopathy are put in a separate measure because it is believed it is more appropriate to treat such patients more aggressively. To qualify for a trigger visit, the visit must either have an E&M code or alternatively have a CPT code that has often been used to indicate chiropractic evaluation and management (CPT=98940, 98941, 98942) or physical therapist evaluation and management (CPT=97110,97112, 97113, 97124, 97140). These codes were determined by lower back pain workgroup, which included chiropractic and physical therapist input.
Measurement period: The expert panel of clinicians on the LBP workgroup initially specified that the measurement period include either 6 weeks or 90 days from the trigger visit and asked us to run our analysis on both time frames. After presenting results it was decided that 3 months was more appropriate based on the additional services used in the 90 day period. The workgroup also included 14 days prior to the trigger visit in the episode to include testing/imaging and medications that might be ordered after a telephone call, but before the initial visit.
Exclusion of individuals with prior LBP visits: Individuals were excluded if they had an outpatient E&M visit for LBP within the prior 90 days to define this as a new episode (consistent with the 90 day length of time for the episode). Also all patients with a radiculopathy diagnosis during the prior or measurement period were excluded from this measure (but could appear in the companion radiculopathy episode) to eliminate conditions which may call for more aggressive care.
Standard exclusions: We have several standard exclusions for each of our measures that are similar to the NCQA exclusions for their relative resource use measures. We exclude individuals with high resource use and high cost conditions that would likely be systematically different from the majority of individuals included in the analysis. These individuals are excluded to create a more homogeneous population included in the analysis.
Diagnostic codes to identify patients with LBP, simple non-specific (acute and subacute): Diagnostic codes to identify individuals with LBP were based on discussions of the LBP workgroup. Staff presented the workgroup the list of diagnsoses used by NCQA and then this was revised over a year of consultant telephone calls.

Rationale for assignment of specified codes
The LBP workgroup defined the triggering diagnostic codes for the episode as lumbago (724.2), backache, nos (724.5), other back disorders, NOS (724.9), somatic disfunction, lumbar (739.3) or sacral (739.4) region, and sprains and strains of lumbar (847.2). As noted above, E&M visits plus selected evaluation and management type codes used in chiropractic and physical therapy treatment were also required to define the trigger visit. Below, we will refer to this as the adjusted E&M list. Given a trigger, the scope of this measure was focused on 90 days of care for individuals with simple non-specific LBP (acute and subacute) following the initial ambulatory care visit with a primary or secondary diagnosis, plus the 14 days prior to the trigger visit for the episode. For this period, all services used that had a diagnosis of 721.xx, 722.xx, 724.xx, 739.3, 739.4, 847.2 in any diagnostic field were considered to be related to LBP care during this measurement period by the LBP clinical workgroup because these codes were LBP-related. That is, given the episode trigger codes, one wanted to include this more general list to cover coding practices that lacked the specificity of the initial trigger code. This was determined using a quasi-Modified Delphi Approach. Moreover, during the measurement testing and validation process, the workgroup refined the diagnostic codes, procedure codes, and medications included in the episode of care measure.

The overarching rationale for each of the codes included in the list is that the clinical workgroup considered the codes as potentially associated with the care of individuals during the measurement period following (and immediately preceding) the initial adjusted E&M visit. Importantly, this was not limited to appropriate care, but rather focused on resources that were likely to be associated with the condition.

In addition, surgical CPT codes to identify services/costs to be included regardless of diagnostic code (including associated hospitalization) were CPT codes for fusion surgery (22840, 22851, 22630, 22614) and other back surgery (63001 thru 63051 inclusive). Imaging CPT codes were considered for potential inclusion as well, but after analysis of most frequent CPT imaging codes for these individuals with and without an appropriate LBP diagnosis, it was determined by the workgroup that all imaging should be accompanied by an appropriate diagnostic code in order to be included as a LBP episode related service.

The medications selected for inclusion in the measure started with the list of medications developed by NCQA for their LBP resource use measure. These were reviewed by the clinical workgroup and a couple of additional medications were added to reflect current pain management – gabapentin. J-codes were included by the clinical workgroup to identify relevant injections.

In developing this measure, the following guidelines and papers on appropriateness of imaging were consulted:

References:

1. NCQA Guidelines for Relative Resource Use (RRU) for People with Acute Low Back Pain (RLB)
2.Chou R et al. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American college of physicians and the American pain society. Ann of Int Med. 2007; 147: 478-491.
3.Gandjour A, et al. European comparison of costs and quality in the treatment of acute back pain. Spine. 2005; 30: 969-975.
4.Pham HH, et al. Rapidity and modality of imaging for acute low back pain in elderly patients. Arch Intern Med. 2009; 169: 972-981.
5.Towers Perrin. Bridges to Excellence – NCQA: NCQA back pain recognition program analysis- part 2. Towers Perrin 2007.
6.Deyo RA. Imaging idolatry: the uneasy intersection of patient satisfaction, quality of care, and overuse. Arch Intern Med 2009; 169: 921-923.
7.Deyo RA, Weinstein JN. Low back pain. NEngl J Med. 2001: 344: 363-370.
8.Andersson GB. Epidemiologic features of chronic low-back pain. Lancet. 1999; 354: 581-585.
9.Bigos S, et al. Acute Low Back Problems in Adults: Clinical Practice Guideline No. 14. Rockville, MD: Agency for Health Care Policy and Research, December 1994. Publication 95-0642.

S8.3. Comorbid and interactions

Detail the treatment of co-morbidities & disease interactions and provide rationale for this methodology.

See risk adjustment details—Section S10.1

S8.4. Clinical hierarchies

Detail the hierarchy for codes or condition groups used and provide rationale for this methodology.

We do not provide specifications for clinical hierarchies.
The only clinical hierarchies used in the measure are associated with the identification of comorbid conditions that are used in risk adjustment. Details are provided in Section 10.1 of this submission form and in the risk adjustment section of the technical appendix in the written measure specification. In short, we use the CMS hierarchical condition categories (HCC) for assignment of comorbid conditions which utilizes a hierarchy of codes based on the ICD-9 codes present during the pre-index period. We rely on the HCC system for identifying comorbid conditions in our risk adjustment procedure. The hierarchies are important for our risk adjustment as they are intended to identify different levels of severity of conditions that may be differentially associated with resource use. We used the HCC system because it is a previously developed and validated system for use in resource use measures.

Within our episode measure there are no hierarchies assigned to any of the codes.

S8.5. Clinical severity levels
Detail the method used for assigning severity level and provide rationale for this methodology.

We do not provide specifications for clinical severity levels.
No severity level is defined for patients included in the episode. We attempt to create a relatively homogenous population through our inclusion and exclusion criteria.

S8.6. Concurrency of clinical events (that may lead to a distinct measure)
Detail the method used for identifying concurrent clinical events, how to manage them, and provide the rationale for this methodology.

We do not provide specifications for concurrency of clinical events.
Each of the measures developed as part of the ABMS measure set was intended as a standalone measure. The measures were not designed to be combined into a single composite measure of resource use for providers. Because the focus during the development of these measures was there eventual pairing with quality measures, each of the measures is considered as a unique measure. Therefore, the concurrency of events and the fact that events may be counted in more than one measure is not an issue. We were not trying to account for the overall resource use of a population but rather focused on resource use within specific cohorts of patients. The relative resource information produced is intended to result in actionable information which is not possible when all of the episodes are combined into a single composite measure.

S9. Measure Construction Logic (Resource Use Measure Module 3)
The measure's construction logic includes steps used to cluster, group or assign claims beyond those associated with the measure's clinical logic. For example, any temporal or spatial (i.e., setting of care) parameters used to determine if a particular diagnosis or event qualifies for the measure of interest.

Construction Logic Supplemental Attachment or URL:
If needed, attach supplemental documentation (Save file as: S9_Construction Logic). All fields of the submission form that are supplemented within the attachment must include a summary of important information included in the attachment and its intended purpose, including any references to page numbers, tables, text, etc.)

URL: http://www.healthqualityalliance.org/hvhc-project/cost-care-measurement-development
Please supply the username and password:
Attachment:

S9.1. Brief Description of Construction Logic
Briefly describe the measure's construction logic.

The following sequence is used to construct the measures:
1. Eligible population identification
2. Identification of related resources
3. Assignment of standardized prices
4. Creation of episode specific strata (if applicable)

S9.2. Construction Logic
Detail logic steps used to cluster, group or assign claims beyond those associated with the measure's clinical logic.

A minimum of 27 months of continuous data is necessary to calculate the measure.

An episode is defined by a trigger event observed over a 12-month identification year. In addition a prior utilization period, which is 12 months prior to the trigger event, is necessary to exclude individuals based on certain criteria. The episode is triggered by an initial ambulatory care visit for LBP; which was not preceded by another ambulatory care visit for LBP within 90 days. The end date is three months from the time of the trigger visit. The duration/measurement period is the three-month period subsequent to the trigger event plus a 14-day period prior to the trigger event.

Note that the identification year is a fixed 12-month period, while the prior year and the measurement period are both

defined relative to the trigger event. Thus, if trigger events occur on the first and last days of the identification year, 12 additional months of data prior to the identification year and three additional months of data subsequent to the identification year are needed.

The following steps are used to complete the construction sequence (for specific codes, see Section S8.2 clinical framework of this submission form as well as the written measure specification/technical appendix accessed via URL).

ELIGIBLE POPULATION IDENTIFICATION

The process of identifying patients to be included in the measure is divided into three separate steps, each with multiple sub-steps. The following steps are used for identifying the included population:

Step 1: Identify patients that meet episode inclusion criteria

Identify patients that have one of the diagnostic codes in Table LBP-A and LBP-A1 for an ambulatory care visit (including E&M visits and specified chiropractic and physical therapy visits) during the event identification year. These ICD-9 codes may be present in any diagnostic field, regardless of corresponding CPT and UB revenue codes.

Step 2: Identify patients that meet age, eligibility and continuous enrollment criteria

1. Age: Identify patients aged 18 to 64.

2. Eligibility
a. Identify benefits during both the identification year and the measurement year. To be included persons must have both of the following benefits in both years
i. Medical benefit
ii. Pharmacy benefit

3. Continuous enrollment
a. Determine enrollment during both the identification and measurement years. (To be eligible, persons must have both medical and pharmacy coverage for the measurement period and prior period (do not include persons whose pharmacy benefits are dropped partway through the identification or measurement period).
b. Identify (or estimate) total days of coverage in each year. (If precise information regarding persons' total days of coverage is not available, it is recommended that measure implementers estimate this information to the best of their ability using available data elements (e.g., monthly enrollment indicators).
c. To be eligible, persons must have at least 320 total days of coverage during each year

Step 3: Identify patients with exclusion criteria

Exclusion criteria:
- Patient had LBP diagnosis (see Table LBP-A) for ambulatory visit (see Table LBP-A1 for included CPT codes) within 90 days prior to potential trigger LBP visit
- Patient had fusion surgery, other back surgery, or fracture (see Table LBP-E) within 6 months prior to potential trigger LBP visit.
- Patient has radiculopathy diagnosis (see Table LBP-F1) during measurement or prior period
- Patient has coincident UTI or sacroiliatis diagnosis on trigger claim (see Table LBP-F3) – a claim with these diagnostic codes cannot be a trigger event
- Active cancer (excluding melanoma, skin, prostate, and chronic lymphocytic leukemia) during measurement or prior period (see Table LBP-F4)
- End stage renal disease (ESRD) during measurement or prior period (Table LBP-F5)
- HIV/AIDS during measurement or prior period (Table LBP-F7)
- Organ transplant during measurement or prior period (Table LBP-F6)
- IV drug abuse during measurement or prior period (Table LBP-F2)
- Neurological impairment during measurement or prior period (Table LBP-F2)
- Intraspinal abscess during measurement or prior period (Table LBP-F2)
- Thoracic or Lumbar Spondylosis with Myelopathy (progressive symptoms) during measurement or prior period (Table LBP-F2)
- Intervertebral disk disorder with myelopathy (progressive symptoms) during measurement or prior period

(Table LBP-F2)

Step 4: Combine prior steps to identify measure population

1. Identify LBP eligible population
2. Exclude those patients not meeting general inclusion criteria (e.g. age, continuous eligibility)
3. Exclude those patients meeting one or more measure exclusion criteria
4. The resulting collection of patients is the measure population

ELIGIBLE EVENT IDENTIFICATION

For each individual in the measure population, identify the low back pain-related claims for services rendered during the measurement year. Claims / encounters will be identified based on the presence of LBP-related diagnosis codes or procedure codes. These events will be used to determine the LBP-related resource use.

Inpatient hospitalization events

Referring to the codes listed in Section S8.2 above, identify all inpatient claims / encounters with a LBP-related diagnostic code appearing in the primary diagnosis field only (see Table LBP-B1). Also identify any inpatient claims with surgery CPT codes listed in Section S8.2 regardless of diagnostic code (see Table LBP-B2).

Outpatient events

Referring to the codes listed in Section S8.2 above, identify all outpatient claims / encounters with a LBP-related diagnostic code appearing in any position (see Table LBP-B1).

Prescription drugs

Referring to the codes listed in Section S8.2 above, identify the low back pain-related medications by therapeutic class or generic/brand medication name during the measurement period (see Tables LBP-C and LBP-D).

ASSIGNMENT OF STANDARDIZED PRICES

Standardized prices are calculated for all of the components of care used to treat or manage the patient's condition to ensure that comparisons can be made solely on the basis of differential practice patterns and resource use. Three separate methodologies are used to derive these standardized prices: for inpatient facility charges, for ambulatory pharmacy charges (i.e., prescriptions dispensed outside the inpatient hospital setting), and for all other charges. These standardized prices are then applied to the claims identified as LBP-related. For further details, see section S10.3 below.

CREATION OF EPISODE-SPECIFIC STRATA
Not applicable.

S9.3. Measure Trigger and End mechanisms
Detail the measure's trigger and end mechanisms and provide rationale for this methodology.

A minimum of 27 months of continuous data is necessary to calculate the measure.

An episode is defined by a trigger event observed over a 12-month identification year. In addition a prior utilization period, which is 12 months prior to the trigger event, is necessary to exclude individuals based on certain criteria. The episode is triggered by an initial ambulatory care visit for LBP; which was not preceded by another ambulatory care visit for LBP within 90 days. The end date is three months from the time of the trigger visit. The duration/measurement period is the three-month period subsequent to the trigger event plus a 14-day period prior to the trigger event.

Note that the identification year is a fixed 12-month period, while the prior year and the measurement period are both defined relative to the trigger event. Thus, if trigger events occur on the first and last days of the identification year, 12 additional months of data prior to the identification year and three additional months of data subsequent to the

identification year are needed.

Rationale:
The lower back pain workgroup, consistent with earlier formulations by NCQA, believed that the episode of care could be set at a fixed time length between six weeks and three months in duration. Data were presented for each of these two time periods and it was found that, while there were not great differences between the two episode time frames, 3 months sufficiently increased resource use that it was appropriate to include this period. The workgroup also believed that physicians would sometimes have imaging or perhaps other tests performed after telephone contact with the patient but before the initial visit. Therefore, it was decided to include lower back pain-related resource use in the episode if it occurred within 14 days prior to the initial visit. However, by the exclusion criteria these would not be included if accompanied by a prior visit. Because of the large role played by chiropractors and physical therapists, it was deemed important to include in the trigger visits E&M-like visits to chiropractors and physical therapists that were, however, not included in the traditional E&M procedure codes. Additional chiropractor and therapist codes were separately defined to allow these codes to be separately grouped (otherwise the betos software puts categorizes these as procedure or "other" codes).

S9.4.Measure redundancy or overlap
Detail how redundancy and overlap of measures can be addressed and provide rationale for this methodology.

We do not provide specifications for measure redundancy or overlap.
To avoid redundancy within episodes of low bsck pain, we have elected to create two distinct measures. One measure for simple, non-specific lower back pain and a separate measure for radiculopathy (with or without lower back pain). The two measures are designed not to overlap.

Beyond lower back pain, the other measures developed by ABMS REF were developed as standalone measures to address all relevant services associated with a particular health care condition. Collectively, the measures do not sum-up to a single total and there is the potential for overlap and redundancy to occur when multiple measures are applied simultaneously.

S9.5.Complementary services
Detail how complementary services have been linked to the measure and provide rationale for this methodology.

We do not provide specifications for linking complementary services.
All services included in the measure are included based on the presence of diagnosis codes, procedure codes, or medications.
Services are identified based on presence of qualifying codes. There is no effort to link complementary services to the episode. The strategy for all of our measures was to rely on the presence of codes to qualify for inclusion in the episode rather than to make assumptions about temporal or other associations between events.

S9.6.Resource Use Service Categories

Inpatient services: Inpatient facility services
Inpatient services: Evaluation and management
Inpatient services: Procedures and surgeries
Inpatient services: Imaging and diagnostic
Inpatient services: Lab services
Inpatient services: Admissions/discharges
Ambulatory services: Outpatient facility services
Ambulatory services: Emergency Department
Ambulatory services: Pharmacy
Ambulatory services: Evaluation and management
Ambulatory services: Procedures and surgeries
Ambulatory services: Imaging and diagnostic
Ambulatory services: Lab services
Durable Medical Equipment (DME)

S9.7.Identification of Resource Use Service Categories
For each of the resource use service categories selected above, provide the rationale for their selection and detail the method or algorithms to identify resource units, including codes, logic and definitions.

At the claim line level, the user should identify all relevant codes specified in the clinical framework Section 8.2 above (see also written measure specification). For inpatient services, these include all relevant ICD9, DRG v24, DRGv25, CPT codes; for ambulatory services, these in clued all relevant ICD9, and CPT codes; for procedures and laboratory these include all relevant ICD9 procedure codes, HCPCs, and CPT codes, and for prescription drugs, these include relevant HCPCs and NDCs.

The above categories were selected because they represent the vast majority of resource use for the episode and the measure developers examined the distribution of costs between categories to evaluate the face validity of the measure. Developers also reasoned that resource use variation between providers by category would be informative. Please refer to Section S8.2 Clinical Framework for the algorithms used to identify/assign some services.

Measure developers also applied the Berenson-Eggers Types of Service (BETOS) system which categorizes all HCPCS codes into resource use areas (e.g. Evaluation and Management, Procedures, Imaging, etc). In addition to the BETOS category there is an additional category included for medications related resource use that is determined using pharmacy data and HCPCs.

Rationale: The BETOS classification system is a widely used, publically available system for classifying healthcare services. These categories can be used to examine cost patterns across providers to identify differences across the different categories of service. This system provides a sufficient number of categories to make meaningful comparisons across patterns of resource use and yet is not too broad so as not to be able to draw conclusions based on differences. Furthermore, identification of important differences allows users to drill down within those categories to identify cost drivers within BETOS categories that may ultimately provide actionable information for providers.

If needed, provide specifications URL (preferred) or as an attachment:

URL:
Please supply the username and password:
Attachment:

S9.8. Care Setting; *provides information on which care settings the measure encompasses.*

Ambulatory Care : Ambulatory Surgery Center (ASC)
Ambulatory Care : Clinic/Urgent Care
Ambulatory Care : Clinician Office
Hospital/Acute Care Facility
Imaging Facility
Laboratory
Pharmacy

S10.Adjustments for Comparability (Resource Use Measure Module 4)
External factors can mingle and affect or confound a measure's result. Confounding occurs if an extraneous factor causes or influences the outcome (e.g., higher resource use) and is associated with the exposure of interest (e.g., episode of diabetes with multiple co-morbidities). Measure developers often include steps to adjust the measure to increase comparability of results among providers, employers, and health plans.

S10.1. Risk adjustment method
Define risk adjustment variables and describe the conceptual, statistical, or other relevant aspects of the model and provide rationale for this methodology.

Calculation of risk adjusted costs (see also the risk adjustment section in the technical appendix of the written measure specification).

The model developed for comorbidity adjustment uses Hierarchical Condition Categories (HCC) to identify comorbidities. This reflects the risk adjustment methodology used by CMS and recently evaluated by NCQA for their Relative Resource Use (RRU) measures. However, there is an important distinction between the use of HCCs by CMS and the model evaluated by NCQA and the risk adjustment model used to estimate expected costs. The CMS and NCQA model use HCCs to adjust TOTAL costs of care, whereas this model focuses on episode-specific costs of care. Because models developed to adjust total costs of care may not reflect the expected costs for episode-specific resource use, new models were developed from a sample of commercially insured patients for risk adjustment. The following process was completed to develop the models:

1. Utilized quasi-Modified Delphi approach with the condition-specific workgroup to categorize HCCs into three groups:
- Include in risk adjustment model;
- Exclude in risk adjustment model; and
- Test impact in risk adjustment model.

2. Identified HCCs in denominator population during the 12 months preceding the measurement year.

3. Tested 12 different model specifications (see Table LBP-RA1 in technical appendix of written measure specification), where the HCCs included in the model varied, and the distribution and link functions in the generalized linear models also varied. Models were developed in a stepwise manner as indicated. The first four models used a gamma distribution and a log link function. The first model included all HCCs identified by the condition-specific workgroup as "Include HCCs" with a prevalence in the population of >=1%. The second model was a reduction of the first model that only included HCCs where $p<0.1$. The third model extended the second model by including HCCs with prevalence >=1% identified as "Test HCCs" by the condition-specific workgroup. The fourth model was a reduction of the third model and included only those HCCs where $p<0.1$. The next set of four models (Models 5-8) repeated the process of the first four models but used a normal distribution and identity link function. Model 9 used all of the HCCs, with the exception of the HCC for the episode being evaluated (e.g., heart failure for the CHF post hospitalization episode), and a gamma distribution with log link function. Model 10 was a reduction of Model 9 where only the HCCs with $p<0.1$ were included. The final two models (Models 11-12) used the same process as Models 9 and 10 with a normal distribution and identity link function.

4. Models were developed in a split sample approach with 75% of the population randomly selected for model development and the remaining 25% used in model evaluation. Model performance was also evaluated in the full cohort.

5. The performance of each model was evaluated through comparisons of the observed and predicted distributions, comparisons of residuals, comparisons of absolute differences between observed and predicted, comparisons of observed-to-predicted ratios, and comparisons of mean squared errors across models. Summary information on model performance was presented to the condition-specific workgroup for selection of a risk adjustment model for the condition. Final model selection was based on the best performing model across metrics. Where model performance was similar, models using the normal distribution were preferentially chosen over the gamma distribution models for ease of implementation. More parsimonious models were also preferentially chosen.

The following is the model selected for estimating adjusted costs in the low back pain episode.

Risk Adjustment Model
Risk Adjusted LBP Episode Costs = \$295 + (Age*\$1) + (Major Depressive, Bipolar, and Paranoid Disorders*\$107) + (Septicemia/Shock*\$97) + (Diabetes with Ophthalmologic or Unspecified Manifestation*\$40) + (Diabetes without Complication*\$34) + (Protein-Calorie Malnutrition*\$112) + (Cirrhosis of Liver*\$81) + (Chronic Hepatitis*\$87) + (Intestinal Obstruction/Perforation*\$40) + (Pancreatic Disease*\$61) + (Inflammatory Bowel Disease*\$37) + (Bone/Joint/Muscle Infections/Necrosis*\$102) + (Rheumatoid Arthritis and Inflammatory Connective Tissue Disease*\$151) + (Severe Hematological Disorders*\$89) + (Drug/Alcohol Psychosis*\$113) + (Drug/Alcohol Dependence*\$86) + (Schizophrenia*\$56) + (Paraplegia*\$460) + (Spinal Cord Disorders/Injuries*\$133) + (Polyneuropathy*\$202) + (Multiple Sclerosis*\$154) + (Seizure Disorders and Convulsions*\$51) + (Coma, Brain

Compression/Anoxic Damage*$244) + (Respirator Dependence/Tracheostomy Status*$197) + (Cardio-Respiratory Failure and Shock*$81) + (Congestive Heart Failure*$60) + (Unstable Angina and Other Acute Ischemic Heart Disease*$58) + (Angina Pectoris/Old Myocardial Infarction*$36) + (Cerebral Hemorrhage*$234) + (Ischemic or Unspecified Stroke*$55) + (Hemiplegia/Hemiparesis*$117) + (Cerebral Palsy and Other Paralytic Syndromes*$208) + (Vascular Disease with Complications*$75) + (Vascular Disease*$73) + (Chronic Obstructive Pulmonary Disease*$81) + (Aspiration and Specified Bacterial Pneumonias*$207) + (Chronic Ulcer of Skin, Except Decubitus*$52) + (Vertebral Fractures without Spinal Cord Injury*$231) + (Major Complications of Medical Care and Trauma*$84) + (Artificial Openings for Feeding or Elimination*($54)) + (Amputation Status, Lower Limb/Amputation Complications*$452)

Measure implementers have two choices when calculating risk adjusted costs. The first is to follow the process specified above to create risk adjustment models that are specific to their population and their dataset. The second option is to follow the below steps and use the above estimates for calculating risk adjusted costs. While the latter is a straightforward calculation, caution is warranted as the risk adjusted equations were derived from a population that may be different from the population to which the measure is being applied.

To estimate risk adjusted costs using the above risk adjustment equations in the measurement population, use the following steps:

Step 1: Identify the presence of HCCs on any claim in the 12 months preceding the measurement year, utilizing both inpatient (primary diagnosis field only) and outpatient encounters (all diagnosis fields).

Step 2: Create a person level file that contains an indicator (yes/no) variable for each of the HCCs. These variables indicate whether or not the patient had evidence of each HCC during the previous 12 months.

Step 3: Calculate an adjustment factor of the average episode costs in the measure population and divide it by the average cost of the test episode (Table LBP-RA2). Apply the inflation factor to the risk adjustment coefficients to account for cost differences between datasets used in development of the risk adjustment models and those used in calculating episode costs.

Summary estimates of the average cost for the low back pain episode in the test episode: Average Cost: $395
Example: To calculate the inflation factor, determine the average episode cost for the population to which the measure is being applied. As an example, the average cost might be $700. Calculate the adjustment factor by dividing the costs from the current population by the average costs in Table LBP-RA2. That would result in an adjustment factor of 1.77 (700/395= 1.77). The adjustment factor is then applied to the estimated coefficients for the adjusted risk adjustment model.

Adjusted Risk Adjustment Model
Risk and Mean Adjusted LBP Episode Costs = 1.77 * Risk Adjusted LBP Episode Costs

Step 4: Use the equation for the appropriate age group to generate risk adjusted expected costs for each individual in the dataset.

Comorbidity Adjustment Strategy Rationale:

We acknowledge that risk adjustment is an important part of the development of an episode of care measure. Risk adjustment is intended to account for variation in episode costs that are not due to differences in practice patterns but rather are due to differences in the case mix of patients. When reporting episode costs at the provider level, risk adjustment attempts to account for differences in the case mix of patients across providers and minimizes the assertion that one providers patients are sicker than the comparator patients. An additional advantage of episode-based measurement is that focusing on costs related to care only for that episode may be a form of risk adjustment because we are not looking at the overall healthcare costs of the patients. Our risk adjustment strategy was not to attempt to account for all of the variation within an episode; however we want to be able to control for resource use variation that is attributed to the episode that may result from differences in patient case mix.

We selected to use Hierarchical Condition Categories (HCC) as our primary strategy for identification of comoribid conditions and for risk adjustment. We selected HCCs because of their use in risk adjustment methodology used by CMS and recently evaluated by NCQA for their Relative Resource Use (RRU) measures. We felt that many users of our episodes would be familiar with HCCs and the use of these measures in administrative data. Moreover, the analytic

programmers for generating HCCs are freely available on the CMS website and therefore we mitigate issues of access to code for creating the risk adjustment groups.

While we use HCC as the starting point for our risk adjustment models, there is an important distinction between the use of HCCs by CMS and the model evaluated by NCQA and our episode definitions. The CMS and NCQA model use HCCs to adjust for TOTAL costs of care whereas, we are focused on the episode-specific costs of care. Briefly, NCQA has created weights for each of the HCCs on total costs of care using data from a large population that has one of the conditions in their RRU measure. These weights can then be applied to different populations to adjust for the presence of comorbid conditions when estimating total costs. The primary concern with applying the adjustment factors available from either CMS or NCQA are the fact they are total costs and not related to the episode-specific costs of care. This would lead to very different risk adjustment models that would not account for as much of the variability within the episode as a risk adjustment model focused on episode-specific costs. We compared the use of the 'off the shelf' HCC values with a risk adjustment model developed specifically for our episode.

See attached supplemental document for illustrative example of comparison of "off the shelf" HCC values to the risk adjustment model developed specifically for our episode (note: diabetes is used for purposes of illustration).

Given the disparity in the means and distributions of the off the shelf HCC values, we felt this justified our approach to develop risk adjustment models for each of our episodes that were focused on episode specific costs.

If needed, provide supplemental information via a web URL (preferred) or attachment with the risk adjustment specifications.

URL:
Please supply the username and password:
Attachment: 10.1_Risk adjustment method-634349436185183385.pdf

S10.2. Stratification Method
Detail the stratification method including all variables, codes, logic or definitions required to stratify the measure and rationale for this methodology

This method is not straified.
There is no stratification, but two measures have been created to distinguish lower back pain with and without radiculopathy.
o Rationale for stratification method
Radiculopathy may require imaging, neurological testing, or other procedures within a shorter time frame, while absent this condition it is generally accepted that the back pain should be allowed to resolve before imaging or other tests are prescribed.

S10.3. Costing Method
Detail the costing method including the source of cost information, steps to capture, apply or estimate cost information, and provide rationale for this methodology.

Standardized prices are calculated for all of the components of care used to treat or manage the patient's condition to ensure that comparisons can be made solely on the basis of differential practice patterns and resource use. Three separate methodologies are used to derive these standardized prices: for inpatient facility charges, for ambulatory pharmacy charges (i.e., prescriptions dispensed outside the inpatient hospital setting), and for all other charges. These standardized prices are then applied to the claims identified as related.

Standard Cost Calculation

Step 1 Identify all claims paid for services rendered during the measurement period and with positive non-zero paid amounts for all patients, regardless as to whether they have been included in the measure population (rejected or unadjudicated claims should be dropped). Categorize these claims as follows (in accordance with the BETOS classification process):

- Inpatient Facility (services provided by a facility during an acute inpatient hospital stay, standard price includes room and board and ancillary services)
- Ambulatory Pharmacy (ambulatory prescriptions included in a member's pharmacy benefit)
- All other (E&M, procedures, imaging, tests, DME, other, and exceptions/unclassified)

Step 2 For each category identified, compute standardized prices. Refer to each service category's instructions (i.e., Calculating Standard Units of Service and Total Standard Cost) below.

Step 3 Combine standardized prices with eligible events (e.g., through a file merge as specified in each service category's instructions).

Step 4 For each individual claim, multiply the standardized price by the number of service units identified on the claim to determine the full cost of the service, hospitalization, or prescription.

Calculating Standard Units of Service and Total Standard Cost: Inpatient Facility

For inpatient facility costs, standardized prices are developed at the diagnosis-related group (DRG) level and – for those hospitalizations where DRG-level information is unavailable – at the ADSC level. Each is adjusted for length-of-stay (LOS) so as to more closely mirror the payment systems typically applied among commercial health plans. Both approaches use RRU HEDIS standardized daily price tables developed by NCQA. All inpatient facility costs are considered "acute" for this analysis.

Step 1 Identify all inpatient stays that occurred during the measurement period. Include stays that may have started before the measurement period or ended after the close of the measurement period. Define a single, unique record describing the member's inpatient stay.

Step 2. Identify the primary discharge DRG. Also identify the DRG version (e.g., CMS-DRG vs. MS-DRG). Care must be taken in using the standardized price tables (specified below) to insure the data and the tables use the same DRG version.

Step 3 Compute the stay's total LOS in days, using paid or expected-to-be-paid days only. Include all paid days in the LOS calculation, whether or not they fall outside the measurement period. Also identify the stay's LOS group based on the stay's LOS and the information below.

LOS (Days)	LOS GRP
1	A
2	B
3-4	C
5-6	D
7-8	E
9-15	F
16 or more	G

Step 4 Compute the LOS per diem multiplier. If the inpatient stay falls completely within the measurement period, use the total number of paid days as the per diem multiplier. If the inpatient stay does not fall completely inside the measurement period, count only the days within the measurement period (including the last day of the period) to compute the per diem multiplier.

Step 5 Download the HEDIS RRU standardized daily price tables from the NCQA website (http://www.ncqa.org/tabid/1092/Default.aspx) for the corresponding measurement periods. Note that there is a one period lag in the file and data periods (i.e. files designated 2007 are based on 2006 data). Some periods may have two sets of tables if there is a significant change in DRG versions. Note: The project staff worked in collaboration with NCQA in development of this methodology for purposes of testing the initial set of measures. Users of the measures may wish to implement their own methodology that does not rely on a price list from NCQA.

Step 6 Calculate the DRG-specific per-diem payment rate by adjusting the standard daily prices for inflation to a reference period using the medical care component of the Consumer Price Index (CPI).

Step 7 Combine DRG-specific per-diem payment rates with the dataset containing eligible inpatient hospital events for

the measure. For each event, multiply the per-diem payment rate by the event's LOS per diem multiplier to determine the event's total standard cost.

Total standard costs will not be computed using this approach for stays that have not been assigned a DRG, and for DRGs that are not assigned a standard price by HEDIS. These stays will be assigned a standard price using the ADSC method described below. (Note: Figures presented in this example are arbitrary and do not reflect any particular dataset or patient. Additionally, the DRG XXX is intended to be used as an illustrative example for calculating inpatient costs. Only DRGs related to the episode should be included in this calculation).

Example:

Assume the calculated DRG-specific per-diem payment rate for DRG XXX for FY 2007 is $900.17. An eligible member had an inpatient stay with the following characteristics:
- A principal diagnosis with an eligible ICD-9 code
- A DRG of XXX (DRG associated with an eligible inpatient stay for the episode)
- Date of admission of February 2, 2007 and date of discharge of February 9, 2007 (fiscal period 2007)
- A LOS of 8 days, and therefore a LOS per diem multiplier of 8 days

This event has a calculated total standard cost of $900.17 x 8 = $7,201.36.

Example:

Again assume the calculated DRG-specific per-diem payment rate for DRG XXX for FY 2007 is $900.17. An eligible member had an inpatient stay with the following characteristics:
- A principal diagnosis with an eligible ICD-9 code
- A DRG of XXX (DRG associated with an eligible inpatient stay for the episode)
- Date of admission of December 28, 2006 and date of discharge of January 2, 2007 (fiscal period 2007)
- A LOS of 6 days, and a LOS per diem multiplier of 2 days (January 1-2).

This event has a calculated total standard cost of $900.17 x 2 = $1,800.34.

Step 8 If DRG information is not available for a given inpatient hospitalization a method must be used that assigns prices to those hospitalizations. The methodology used in testing the initial development of the measures was to assign an Aggregate Diagnostic Service Category (ADSC) for the stay using the principal discharge diagnosis. To assign ADSC, download the ADSC Table (Table SPT-INP-ADSC) from the NCQA Web site (http://www.ncqa.org/tabid/1092/Default.aspx) and match the principal ICD-9-CM Diagnosis code from the discharge claim to an ADSC. If the claim does not contain a DRG and the primary ICD-9-CM Diagnosis code is invalid or missing, map the inpatient stay to the ADSC Table's MISA category. An alternative would be to create average prices from the dataset the measures are being implemented for each of the ADSC categories and discharge ICD-9-CM codes and assign those prices to missing hospitalizations.

Step 9 Determine if the member underwent major surgery during the inpatient stay. If this information is not available within the dataset, this may be determined using the list of codes included in a table from the NCQA Web site (Maj-Surg Table). Flag eligible members if one procedure code in the Maj-Surg-Table is present from any provider during the time period defined by the admission and discharge dates.

Step 10 Match each ADSC, LOS per diem multiplier, and major surgery flag assignment for the stay to a value in the Table SPT-INP-ADSC to obtain the assigned standard price. For each event, multiply the per-diem payment rate by the event's LOS per diem multiplier to determine the event's total standard cost. As with the DRG method, the ADSC standard prices must be adjusted for inflation to a reference period using the CPI. Between this ADSC methodology and the previously described DRG-based methodology, each inpatient hospital stay should now have an associated standardized price.

Example:

An eligible member had an inpatient stay with the following characteristics:
- A principal diagnosis for an eligible event assigned to ADSC category Respiratory-C (RESC)
- No available valid DRG information
- Date of admission of February 2, 2007 and date of discharge of February 9, 2007
- A LOS of 8 days, and therefore LOS group E

- A major surgery event during the stay

Using Sample Table SPT-INP-ADSC, we determine this event has a standard per-diem payment rate of $1,474.00. Therefore this event has a calculated total standard cost of $1,474 x 8 = $11,792.

Calculating Standard Units of Service and Total Standard Cost: Ambulatory Pharmacy

For ambulatory pharmacy-related costs, standardized prices are developed at the NDC level, adjusted for days supply.

Step 1 Identify all pharmacy services that occurred during the measurement period. The following pharmacy services should also be included:

- Prescriptions that may have been dispensed before the measurement period and had days supply that extended into the measurement period (e.g., a prescription with a dispensed date of December 15, 2007 and 30 days supply would extend 13 days into the measurement period beginning January 1, 2008)
- Prescriptions that may have been dispensed during the measurement period and had days supply that extended into the following period (e.g., a prescription with a dispensed date of December 20, 2008).

Define a single, unique record describing the pharmacy service.

Step 2 Identify the NDC code and the days supply for each prescription, whether or not some days fall outside the measurement period.
If the days supply is not available for a given pharmacy claim, set the claim's standard cost to be equal to its listed payment amount.

Step 3 Compute the days supply per diem multiplier. If the prescription's days supply fall completely within the measurement period, use the claim's listed days supply as the per diem multiplier. If the prescription's days supply do not fall completely inside the measurement period, count only the days within the measurement period (including the last day of the period) to compute the per diem multiplier.

Step 4 For each NDC, calculate the total NDC-specific payments and the total days supply across all pharmacy claims within that NDC during the measurement period. Using these totals, calculate NDC-specific per-day-supply payment rates by dividing total NDC-specific payments by total days supply for each NDC.

Step 5 Combine NDC-specific per-day-supply payment rates with the dataset containing eligible pharmacy events for the measure. For each event, multiply the per-day-supply payment rate by the event's days supply per diem multiplier to determine the event's total standard cost.

Calculating Standard Units of Service and Total Standard Cost: All Other

For all non-inpatient hospital, non-pharmacy costs, standardized prices are developed at the procedure code and modifier level.

Step 1 Identify all non-inpatient hospital, non-pharmacy services that occurred during the measurement period.

Step 2 Identify the primary procedure code (CPT, HCPCs, ICD-9, etc.) and the first modifier code for each service.

Step 3 For each procedure-modifier combination, calculate the total procedure/modifier-specific payments across all non-inpatient-hospital, non-pharmacy claims with that procedure-modifier combination as well as the frequency of the procedure-modifier combination during the measurement period. Calculate procedure/modifier-specific payment rates by dividing total procedure/modifier-specific payments by the frequency for each procedure-modifier combination.

Example:
Assume that there are 3 non-inpatient-hospital, non-pharmacy claims during the measurement period with the following characteristics:
Patient: 1111, Procedure (CPT-4): 71010, Modifier: Date: 2/1/2007, Payment: $21
Patient: 1111, Procedure (CPT-4): 72240, Modifier: TC, Date: 2/18/2007, Payment: $90
Patient: 2222, Procedure (CPT-4): 71010, Modifier: Date: 1/5/2007, Payment: $25

For the procedure/modifier combination: 71010

The total payment is $21 + $25 = $46
The total frequency is 2
Therefore the procedure/modifier-specific payment rate is $46/2 = $23
For the procedure/modifier combination: 72240/TC
The total payment is $90
The total frequency is 1
Therefore the procedure/modifier-specific payment rate is $90/1 = $90

Step 4 Combine procedure/modifier-specific payment rates with the dataset containing eligible non-inpatient-hospital, non-pharmacy events for the measure so that each procedure-modifier combination is paired with its corresponding payment rate. This payment rate is the event's total standard cost.

Calculation of total individual episode costs

The resource use identified as diabetes-related– and to which standardized prices have been applied (i.e., the collection of eligible events) – is used to calculate individual level episode costs. The following steps are used in the calculation of total individual level costs.

Step 1: For each individual included in the episode, sum all of the total standard costs linked to diabetes-related events occurring during the measurement period at the BETOS service category level. This will provide an estimate of the costs of each category of service over the measurement period.

Step 2: For each individual in the episode, sum ALL total standard costs linked to diabetes-related events to calculate TOTAL episode costs.

Step 3: Exclude individuals that do not have positive, non-zero costs (e.g. outpatient visit, hospitalization, medication use) during the measurement period.

Rationale for costing method

We used standardized prices to estimate the costs for all components of care in the claims data that a patient received data during the measurement period. Because costs in claims data reflect both the quantity and mix of services delivered as well as the prices paid for those services, some of the cost variation is due to price differences across providers (Thomas et al., 2005). Variations in cost data among organizations and over time can obscure real cost differences (Ritzwoller, et al., 2004) and impede comparisons across providers. To ensure that comparisons are made on the basis of differences in practice patterns and resource use, we developed standardized prices, such that a given service would have the same price across all providers (Thomas et al., 2005). We used separate methods to estimate standardized price that were used to calculate for inpatient facility costs, pharmacy costs, and cost for all other care.
For the inpatient facility use, we developed standardized prices using diagnosis-related group (DRG) information. For hospitalizations without DRG-level information, we used aggregate diagnostic service category (ADSC) level information. In each case, we adjusted for length-of-stay (LOS) during the measurement period so as to more closely mirror the payment systems typically applied among commercial health plans. Both approaches use relative resource use (RRU) HEDIS standardized daily price tables developed by NCQA. We worked in collaboration with NCQA in development of this methodology; however, users of the measure may need to implement their own methodology that does not rely on a price list from NCQA.
For pharmacy use, we determined the days supply for each medication that was dispensed during the measurement period identified by a unique national drug code (NDC). We calculated a standardized price per diem for each NDC in our data by dividing the total payments in the claims data by the total days supply in the claims data for that NDC. We then estimated patient's pharmacy costs by multiplying the standardized price per diem for each NDC by the patient's days supply during the measurement period for that NDC. Standardized prices for pharmacy was estimated using this approach rather than an average whole price (AWP) because the AWP is not defined by law or regulation and does not reflect discounts obtained by most purchasers. As a result, the ultimate price paid by purchasers is often significantly lower than the AWP (Pereira, 2005).
For all other use, we identify the primary procedure code (CPT, HCPCs, ICD-9, etc.) and the first modifier code for each service. We calculated a standardized price for each procedure/modifier by dividing the total procedure/modifier-specific payments by the frequency for each procedure/modifier combination in the claims data. We then applied this standardized price to each patient's procedure/modifier combination that occurred during the measurement period. This approach allowed for a consistent methodology to be applied to each procedure/modifier combination in the claims data

to achieve the same price for a service across all providers.

References:
Pereira BJG. Medicare Prescription Drug, Improvement and Modernization Act: Average Wholesale Price (AWP) Medscape Nephrology.2005;2(1)

Ritzwoller DP, Goodman MJ, Maciosek MV, Lafata JE, Meenan R, Hornbrook MC, Fishman PA. Creating Standard Cost Measures Across Integrated Health Care Delivery Systems. J Natl Cancer Inst Monogr 2005;35:80 – 87

Thomas JW, Grazier KL, Ward K. Economic Profiling of Primary Care Physicians: Consistency among Risk-Adjusted Measures. Health Services Research. 2004;39(4):985- 1004

S11. Measure Reporting (Resource Use Measure Module 5)
The measure developer must determine which of the following Measure Reporting functions: attribution approach, peer group, outliers and thresholds, sample size, and benchmarking and comparative estimates, are submitted as measure specifications or as guidelines. Specifications limit user options and flexibility and must be strictly adhered to; whereas guidelines are well thought out guidance to users while allowing for user flexibility. If the measure developer determines that the requested specification approach is better suited as guidelines, please select and submit guidelines, otherwise specifications must be provided.

S11.1. Detail attribution approach
Detail the attribution rule(s) used for attributing costs to providers and rationale for this methodology (e.g., a proportion of total measure cost or frequency of visits during the measure's measurement period) and provide rationale for this methodology.

Resource use and costs for low back pain episodes are attributed to one or more physicians on a hierarchical basis. The total counts of E&M codes by unique provider ID are used for provider attribution. For each episode identify all such E&M codes occurring during the measurement year. The E&M codes are used to assign attribution using the following hierarchy:

1. Costs and resource use assigned to a single provider if that physician has at least 70% of the E&M codes during the measurement year ("single attribution"); OR
2. If no provider has more than 70% of the E&M codes, costs and resource use are assigned to each of the providers that have at least 30% of the E&M codes for a patient during the measurement year ("multiple attribution"); OR
3. If no provider has at least 30% of the E&M codes during the measurement year, the costs and resource use for that patient are not attributed to any provider ("no attribution").

To identify the attributable provider, the following steps will be used:

Step 1: Identify qualifying E&M codes for the episode (see also Table LBP-A1): Office or Other Outpatient Services 99201–99215; Hospital Observation Services 99217–99220; Hospital Inpatient Services99221–99239; Consultations99241–99275; Critical Care and Intensive Care Services 99289–99298; Nursing Facility, Domiciliary and Home Services 99301–99350; Case Management Services and Care Plan Oversight Services 99361–99380; Preventive Medicine Services 99381–99429; Other E&M Services99450–99456, 99354–99357

Step 2: For each individual included in the episode, sum the total qualifying E&M codes by each provider for that individual.

Step 3: Calculate the proportion of E&M codes for each provider that had a claim for each of the patients:

- Proportion of Care = Total count of provider's E&M qualifying codes divided by total count of all qualifying E&M codes

Step 4: Assign attribution based on the hierarchical attribution model described above.

Rationale:

A minimum of 30% of physician visits or physician costs has often been used as a minimum before an episode has been attributed to a physician (1,2). Similar to these previous efforts, our physician workgroup believed that this was a reasonable cutoff to define the minimum number of E&M visits before a physician received attribution. By the same token until a physician was responsible for 70% of E&M visits, it was believed by the physician workgroup that more than one physician shared responsibility for the costs of the episode and therefore multiple attribution was appropriate. Further, an advantage of multiple attribution is that it increases the number of cases attributed to physicians – a factor that is important given the generally acknowledged problem of many physicians having too limited number of cases to allow them to be included in a comparison with other physicians. As to the use of E&M visits rather than payments to define attribution cutoff levels, the use of visits appears to be more transparent to physicians, especially given the use of standardized rather than actual payments and the fact that many expensive aspects of care resulting from physician decisions are not billed by that physician. Further, when primary physicians are involved in the episode, their physician-related payments are likely to be lower due to lower visit fees, yet it is more likely that they were responsible for referrals to specialists.

1. Merotra A, Adams JL, Thomas W, McGlynn A. The effect of different attribution rules on individual physician cost profiles. Annals of Internal Medicine 2010; 152:649-654.
2. Adams JL, Mehrotra A, Thomas JW, McGlynn EA. Physician cost profiling – reliability and risk of misclassification. N England J Med; 362: 1014-21.

S11.2.Identify and define peer group
Identify the peer group and detail how peer group is identified and provide rationale for this methodology

Guidelines : Peer group comparisons should be based on physician specialty as providers should only be compared to those of the same specialty.

Focusing on comparing physicians of the same specialty is another mechanism to ensure the severity of patients is similar across providers. It is quite possible that patients predominantly seen by specialists may be more complex or sicker patients than those seen by primary care physicians. Additionally, research has shown differences in the care provided by specialists versus generalists (1.2). Therefore, comparisons should be made to providers of similar specialties.

References:

1. Nash IS, Corrato RR, Dlutowski MJ, O´Connor JP, Nash DB. Generalist versus specialist care for acute myocardial infarction. Am J Cardiol. 1999 Mar 1;83(5):650-4.
2. Schreiber TL, Elkhatib A, Grines CL, O´Neill WW. Cardiologist versus internist management of patients with unstable angina: treatment patterns and outcomes. J Am Coll Cardiol. 1995 Sep;26(3):577-82.

S11.3. Level of Analysis:

Clinician : Individual

S11.4.Detail measure outliers or thresholds
Detail any threshold or outlier rules and decisions based on measure resource use and provide rationale for this methodology

Guidelines : For the physician reports, total observed episode costs are winsorized at the 2nd and 98th percentile, but claim line outliers are not removed and the use of risk adjusted results are intended to correct for any extreme outliers. The only exception is inpatient admissions. Extremely high admissions costs are winsorized at the 99th percentile (i.e. any value higher than the 99th percentile are set to the 99th percentile cost).
Rationale: Winsorizing and risk adjustment limits the influence of outliers. Episodes with extremely high admission costs skews mean costs for the entire episode. Winsorizing admissions at the 99th percentile reduces this effect without eliminating information on the distribution of total episode costs.

S11.5.Detail sample size requirements
Detail the sample size requirement including rules associated with the type of measure

We do not provide specifications or guidelines for sample size requirements : The ABMS REF episode-based resource use measures do not randomly sample enrollees nor do we recommend that implementers construct measures from a random sample. Regarding the issue of sample size determination. It is well known that the nature of resource use measurement at the level of individual providers will often lead to unstable estimations. There have been a number of efforts to derive a single number for which such measures might be stable enough for comparison of providers or individual providers over time. Yet to date there is no commonly accepted minimum. At this time we have not attempted to derive a minimal sample size for measure use.

S11.6.Define benchmarking or comparative estimates

Detail steps to produce benchmarking and comparative estimates and provide rationale for this methodology

Guidelines : Creation of provider summaries

The provider summaries are a report of the resource use for an attributable unit (hospital or provider) compared to their peer group, their non-peer group and all episodes in the dataset. Creation of the provider summaries uses the summary episode costs combined with the attributable provider data and the risk adjusted episode costs.

Step 1: Create a dataset that includes the following information: patient ID, total episode cost, attributable provider ID (or ID for the attributable unit if at the hospital level), attributable provider specialty type and episode expected costs from the risk adjustment model.

Step 2: Calculate the observed-to-expected ratio for each of the episodes by dividing observed costs for the episode by expected (predicted) costs for the episode.

O-to-E = Sum of Observed Costs / Expected Costs from Risk Adjustment Model

Step 3: If applicable, create indicators for the strata the episodes fall into so that separate summaries can be created for each of the strata.

Step 4: Summarize the observed, expected and observed-to-expected ratio for each attributable provider. Report minimum, maximum, median and mean values of the observed-to-expected ratio for all episodes attributed to the provider.

Step 5: Summarize the observed, expected and observed-to-expected ratio for each provider type, overall, and within each strata (if applicable). Report summary statistics for each of the provider types so the data are summarized for all providers of the same type. For example, report the summary statistics for the observed-to-expected ratio for all of the family practice physicians to facilitate peer group comparisons.

Step 6: Summarize the observed, expected, and observed-to-expected ratio for all of the episodes.

Step 7: For each of the individual attributable units (hospital or provider), determine the proportion of O-to-E ratios that are greater than or equal to the 75th percentile of the O-to-E ratio for the peer group. Calculate the 95% confidence interval for the proportion. For example, if the provider for which summary statistics are being calculated is a general internist and it is Dr. Y, the 75th percentile of O-to-E ratios for all episodes attributable to general interests is determined. The proportion of Dr. Y´s O-to-E ratio that are above the 75th percentile for all general interest episodes is determined and a 95% confidence interval is calculated for that proportion.

Step 8: Create provider summary reports for each attributable provider in the dataset

S12.Type of Score:

Ratio

If available, please provide a sample report:

S12_sample score report LBP.pdf

S12.1. Interpretation of Score.
(Classifies interpretation of score (s) according to whether higher or lower resource use amounts is associated with a higher or lower score, a score falling within a defined interval, or a passing score, etc)

The summary score calculated for the measure is the ratio of the observed cost to the expected cost or the O-to-E ratio. The O-to-E ratio is calculated for each patient for the attributable provider and summary statistics are calculated for the O-to-E ratio. The O-to-E ratio provides an estimate of the observed cost for a patient to the expected cost based on the patient's mix of chronic conditions. Expected costs for each patient are the calculation of their risk adjusted costs. A value of 1 for the O-to-E ratio indicates that the observed costs are equal to the expected costs. A value greater than 1 indicates that observed costs are more than what would be expected based on the patient's mix of chronic conditions. A value less than 1 indicates that the observed costs are less than what would be expected based on the patient's mix of chronic conditions. Calculation of the O-to-E ratio incorporates our approach to risk adjustment by determining the expected costs from the risk adjustment model. A summary O-to-E ratio is calculated for each of the attributable providers which combines all the episodes for that provider. Summary statistics are calculated for each provider for the raw (unadjusted) costs for the episode, expected costs and the O-to-E ratio. Each summary measure includes minimum, maximum, median, and mean values.

S12.2. Detail Score Estimation
Detail steps to estimate measure score.

Creation of provider summaries
The provider summaries are a report of the resource use for an attributable unit (hospital or provider) compared to their peer group, their non-peer group and all episodes in the dataset. Creation of the provider summaries uses the summary episode costs combined with the attributable provider data and the risk adjusted episode costs.

Step 1: Create a dataset that includes the following information: patient ID, total episode cost, attributable provider ID (or ID for the attributable unit if at the hospital level), attributable provider specialty type and episode expected costs from the risk adjustment model.

Step 2: Calculate the observed-to-expected ratio for each of the episodes by dividing observed costs for the episode by expected (predicted) costs for the episode.
O-to-E = Sum of Observed Costs / Expected Costs from Risk Adjustment Model
Step 3: If applicable, create indicators for the strata the episodes fall into so that separate summaries can be created for each of the strata.
Step 4: Summarize the observed, expected and observed-to-expected ratio for each attributable provider. Report minimum, maximum, median and mean values of the observed-to-expected ratio for all episodes attributed to the provider.

Step 5: Summarize the observed, expected and observed-to-expected ratio for each provider type, overall, and within each strata (if applicable). Report summary statistics for each of the provider types so the data are summarized for all providers of the same type. For example, report the summary statistics for the observed-to-expected ratio for all of the family practice physicians to facilitate peer group comparisons.

Step 6: Summarize the observed, expected, and observed-to-expected ratio for all of the episodes.
Step 7: For each of the individual attributable units (hospital or provider), determine the proportion of O-to-E ratios that are greater than or equal to the 75th percentile of the O-to-E ratio for the peer group. Calculate the 95% confidence interval for the proportion. For example, if the provider for which summary statistics are being calculated is a general internist and it is Dr. Y, the 75th percentile of O-to-E ratios for all episodes attributable to general interests is determined. The proportion of Dr. Y´s O-to-E ratio that are above the 75th percentile for all general interest episodes is determined and a 95% confidence interval is calculated for that proportion.
Step 8: Create provider summary reports for each attributable provider in the dataset

S12.3. Describe discriminating results approach
Detail methods for discriminating differences (reporting with descriptive statistics--e.g., distribution, confidence intervals)

Summary reports are generated at the attribution level that includes a summary estimate for the provider or hospital, the

peer group, the non-peer group and the overall summary for the episode in the entire population. For each attributable provider / hospital the observed, expected and O-to-E ratio are summarized. The summaries are created to facilitate comparisons for the attributable provider or hospital with other providers in the same peer group and overall. The most meaningful comparisons are likely those between the provider or hospital and the peer group. Even though the results are risk adjusted, this may help to further balance the case mix or severity of the patients being compared. The summary statistics for the O-to-E ratios can be compared in order to provide a sense of the relative performance of the provider or hospital compared to peers. In addition, the proportion of O-to-E ratios about thresholds of 2.0 and 2.5 are provided for comparisons. Finally, for the attributable unit (hospital or provider) the proportion of O-to-E ratios that are greater than or equal to the 75th percentile of the O-to-E ratio for the peer group is determined and the 95% confidence interval calculated. The expectation would be that 25% of the estimates for the attributable provider would fall about this value if the distribution of O-to-E ratios is similar to the peer group. A statistically significant difference would be found between the groups if the 95% confidence interval did not include 25% in the range. For example, if the proportion at or above the 75th percentile of the peer group is 38% and the 95% confidence interval ranges from 28% to 48% than this provider would have significantly more O-to-E ratios at the upper end of the distribution than the peer providers. Alternatively, if the proportion at or above the 75th percentile was 8% and the 95% confidence interval ranged from 3% to 16% then the provider would have significantly fewer O-to-E ratios in the upper end of the distribution than the peer group. The 75th percentile in our testing was selected as an illustrative cut-point and it will be important to evaluate this threshold for comparing providers.	

TESTING/ANALYSIS	
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. See guidance on measure testing.	**Eval Rating**
TESTING ATTACHMENT (5MB or less) or URL: *If needed, attach supplemental documentation (Save file as: SA_Reliability_Validity Testing) All fields of the submission form that are supplemented within the attachment must include a summary of important information included in the attachment and its intended purpose, including any references to page numbers, tables, text, etc.* URL: Please supply the username and password: Attachment: SA_Reliability_Validity Testing LBP Simple.pdf	
SA1. Reliability Testing *For each module tested or for the overall measure score:* **SA1.1. 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)* Thomson Reuter´s Marketscan Dataset was used in the testing of the ABMS REF episode-based resource use measures. The MarketScan Commercial Database provides a rich, comprehensive source of longitudinal administrative claims data, offering the largest convenience sample available in proprietary databases with over 30 million covered lives in each of the three most current years of data. The MarketScan Commercial Claims and Encounters (Commercial) Database is constructed from data contributed from over 100 medium and large size employers and health plans, representing over 130 unique carriers. The MarketScan Databases' large sample size constitutes a nationally representative data sample of the U.S. population under the age of 65 with employer-sponsored health insurance. The stability of MarketScan data sources provides superior continuity of patients over multiple years, generally longer than other claims databases because the majority of the MarketScan data are sourced from large employers. As long as individuals remain with the same employer, they can be tracked across health plans.	**2a2** **H☐** **M☐** **L☐** **I☐**

Features of the MarketScan Research Databases include:
- Fully paid and adjudicated claims including inpatient, outpatient, and prescription drug claims
- Complete payment/charge information, including amount of patient responsibility
- Validated diagnosis, procedure, and other standard codes on claims where applicable (CPT, ICD-9, DRG, NDC, etc)
- Demographic information on enrollees including age, gender, and geographic information (three-digit zip codes and MSA)
- Plan-type identifiers in the database include major medical, comprehensive, PPO, EPO, HMO, consumer-driven health plan, capitated or part-capitated POS and non capitated POS
- Standardized data elements and definitions, ensuring accurate comparisons
- Clinical data enhancements, such as Therapeutic Class and Generic Product Identifiers on drug records, and Major Diagnostic Categories and Diagnosis Related Groups on inpatient and outpatient records
- Case records linking all of the hospital, physician, and ancillary services provided during an inpatient stay, allowing for comparisons based on such statistics as average length of stay, cost per admission, etc.

These data reflect the real world of treatment patterns and costs by tracking millions of patients as they travel through the healthcare system, offering detailed information about all aspects of care. Data from individual patients are integrated from all providers of care, maintaining all healthcare utilization and cost record connections at the patient level.

SA1.2. Analytic Methods
(Describe method of reliability testing and rationale)

Reliability refers to the reproducibility of results (Bannigan and Watson, 2009). To investigate the reliability of the measure, we examined the distribution of costs across categories of care (inpatient facility charge, evaluation and management, procedures, etc.) for all simple non-specific acute/subacute LBP episodes in the Marketscan data that met inclusion/exclusion criteria and for a subsample of this cohort. After applying inclusion criteria to the Marketscan data, we identified 6,600,621 potential LBP trigger visits (see attached data summary Slide 4). This was reduced to 3,081,900 eligible visits after eliminating people lacking medical or pharmacy coverage over the prior and observation periods. After applying the exclusion criteria, there were 511,312 episodes for the measure. The major reasons for these exclusions were prior unspecified LBP claims (meaning it was not an index event in 73.1% of cases) and presence of a radiculopathy diagnosis during the prior to measurement period (23.7% of cases). For these 511,312 individual episodes, we examined the distribution of costs across categories of care for the entire cohort and the subsample as well as across geographic regions. For those individuals with physician unique physician identifiers we were able to define variation across physicians (by specialty), chiropractors, and physical therapists and compare costs across specialties. Rationale: Our investigation of reliability allowed us to leverage on analyses that were being done to examine overall resource use and attribution of care.

Reference: Bannigan K, Watson R. Reliability and validity in a nutshell Journal of Clinical Nursing. 2009;18: 3237–3243

SA1.3.Testing Results
(reliability statistics, assessment of adequacy in the context of norms for the test conducted)

For all LBP episodes in the Marketscan data that met the inclusion/exclusion criteria (i.e., 511,312 episodes), procedures comprised the largest portion of costs, at 26% of total costs, followed by evaluation and management (22% of total costs)(see attached data summary Slide 6), with total medical care costs averaging $385 per episode. A problem with this analysis is that it uses the betos classification scheme that places chiropractic and physical therapy CPT codes in either procedures or "other services" categories. The clinical workgroup therefore defined chiropractic and physical therapy codes and these were taken out of "procedures" and "other services" and defined separately. As shown in Table 27, the largest portion of costs (out of the total average costs of $383 per episode) were E&M OP at $81, Chiropractic at $63, PT at $62, and drug costs at $61.

SA1.4.Finding statement(s)—*(i.e., is the measure deemed reliable, limitations identified)*

The results of our comparison would suggest that the measure could be deemed reliable. It should be noted that this investigation highlighted a limitation of the data regarding the portion of missing provider identifiers. In the Marketscan data about half of the data are lacking physician identifiers. However, this is due to confidentiality

agreements with insurers and would not be expected to be a problem were the measure to be implemented by the insurers to monitor provider episode costs.

SA2.Validity Testing
For each module tested or for the overall measure score:

SA2.1. 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 section SA1.1 for description of Thomson Reuters Marketscan dataset

SA2.2.Analytic Method
(Describe method of validity testing and rationale; if face validity, describe systematic assessment)

Validity testing focused primarily on face validity. Initial testing included:
Level 1 analyses
- Examined impact of inclusion/exclusion criteria on episode denominator
- Examined total episode spending by type of service
- Identified top 20 "condition-related" and "non-condition-related" E&M, procedures, imaging, tests, inpatient admissions (by ICD-9 and DRG) and drugs, by service counts and dollar volume
- Tested proposed attribution logic, examined variability in per-episode resource use at individual provider level (as relevant) and by provider specialty.

Level 2 analyses
- Incorporated risk adjustment
- Produced sample physician-level reports in which observed-to-expected ratios are computed and the distribution of each physician's episodes is compared to the peer group's distribution.
- Examined specific drivers of resource use variation
- Examined variability in per-episode resource use across regions, states and the specialties of attributed providers.

Throughout the process of empirically testing the measures, summary analyses were presented to the workgroups for review and discussion. The workgroups reviewed denominator attrition diagrams to assess how the measure's inclusion and exclusion criteria affected the episode's denominator. They also reviewed summaries of costs by type of service (inpatient hospital care, outpatient care, procedures, imaging, tests, and prescription drugs) and were asked to assess whether the distributions matched the clinical expectations for the condition's treatment. The clinicians were also presented with analyses of diagnosis and procedure level details in order to ensure that appropriate services were being captured and grouped to the episodes. At each step in the process, the measure specifications were revised based on workgroup feedback.
In addition to workgroup feedback results of the preliminary testing were also shared with a Technical Advisory Committee and the QASC Episodes Work Group and the measures revised according to feedback.
By presenting our results to the clinical workgroups and others to examine the distributions of resource use and costs to determine if these results meant their clinical expectations, we were able to access the face validity of our results.

SA2.3.Testing Results
(statistical results, assessment of adequacy in the context of norms for the test conducted; if face validity, describe results of systematic assessment)

2b2

We have developed a measure specification to measure resource use associated with a non-specific acute/subacute LBP episode of care. The measure includes resource use related to management of LBP over a 90-day period (plus 14 days prior to the index visit) in order to capture all LBP-related costs of treating these patients during the episode. For the Level 1 analysis, we found that there were 511,312 individuals after applying our exclusion criteria (see attached data summary Slide 4). We found that the average total cost of a non-specific LBP episode was $385. As part of the Level 2 analyses, we examined variability in per episode resource use by specialties of the attributed providers. The highest volume specialty was chiropractor, followed by family practice and internal medicine (see attached data summary Slide 27). Because of the substantial use of chiropractors and physical therapists in treating LBP, in addition to E&M OP visits, chiropractic and PT visit codes were separated out (these included codes beyond the evaluation and management-

H☐ M☐ L☐ I☐

like codes used by these specialties and included when defining episode trigger visits – see tables) and when this was done it was found that OP E&M visits accounted for $81, chiropractic code visits accounted for $63, and PT code visits accounted for $62, while drug costs were the next highest category at $61. It would be expected that E&M, chiropractic and PT visits would be a large component of costs for patients with LBP. It would also be expected that chiropractors, family medicine, internal medicine and physicians not otherwise described, would account for most of the resource use since individuals might be expected to see either their personal physicians or a chiropractor for initial treatment. These results were presented to the clinical workgroup who concurred that these results met their clinical expectations and had face validity.

SA2.4. Finding statement(s)—(*i.e., is the measure deemed reliable, limitations identified)*

Based on the results of our investigations and concurrence from the clinical workgroup, our measure should be deemed to have face validity.

SA3.Testing for Measure Exclusions

SA3.1. Describe how the impact of exclusions (if specified) is transparent as required in the criteria

In the attached data summary, we have detailed how the exclusions impacted the resulting size of the cohort (see attached data summary Slide 4).

SA3.2. 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)

See section SA1.1 for description of Thomson Reuters Marketscan dataset.

SA3.3. Analytic Method
(Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference)

We examined the impact of several types of exclusions. In order to ensure that data are available for assessing the episode of care, we excluded individuals without continuous insurance coverage including medical and pharmacy benefits. We also excluded individuals who met standard NCQA exclusions for conditions that are resource intensive, which could potentially have a larger impact on resource use than the condition being studied (i.e., end stage renal disease, active cancer management, etc.) There were also exclusion criteria that were specified for this condition by the clinical workgroup: age 18-64, a UTI or sacroiliitis concomitant with LBP trigger codes, neurological impairment, progressive symptoms, drug abuse, intraspinal abscess. We examined the impact of these exclusions on the resulting cohort size.

2b3

SA3.4. Results
(statistical results for analysis of exclusions, e.g., frequency, variability, sensitivity analyses)

The exclusion of individuals without continuous enrollment in health insurance with medical and pharmacy benefits reduced the sample by more than 50%. Among the 6,600,621 episodes that met the inclusion criteria for the measure, 3,081,900 or 46.7% remained after the continuous enrollment exclusion criteria were applied (see attached data summary Slide 4). Of these, 511,312 remained after implementing the other exclusion criteria. Most were excluded due to prior ambulatory LBP visits within 90 days prior to a potential trigger visit, i.e., it was a follow-up rather than initial visit in the LBP episode. 23.7% of individuals were excluded because they at some point had a radiculopathy diagnosis.

SA3.5. Finding statement(s)-- *(i.e., is the measure deemed reliable, limitations identified)*

Based on the results of our analyses and feedback from the clinical workgroup, we would deem the measure to be reliable. Our investigation did find that a substantial portion of individuals were excluded due to the continuous enrollment criteria, which is related to the data itself rather than the clinical characteristics of the individuals.

H☐
M☐
L☐
I☐

SA4. Testing Population *Which populations were included in the testing data? (Check all that apply)* Commercial	
SA5. Risk adjustment strategy *Refer to items S10.1 and S10.2 to rate this criterion.*	**2b4** **H☐** **M☐** **L☐** **I☐**
SA6. Data analysis and scoring methods *Refer to items S12-S12.3 to rate this criterion.*	**2b5** **H☐** **M☐** **L☐** **I☐**
SA7. Multiple data sources *Refer to S7 & all SA1 items to evaluate this criterion.*	**2b6** **H☐** **M☐** **L☐** **I☐** **NA☐**
SA6. Stratification of Disparities (if applicable) *Refer to item S10.2 to rate this criterion.*	**2c** **H☐** **M☐** **L☐** **I☐**
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Scientific Acceptability of Measure Properties*?	
Steering Committee: Overall, was the criterion, *Scientific Acceptability of Measure Properties*, **met?** **Rationale:**	**Y☐** **N☐**
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.	**Eval Rating**
Meaningful, Understandable, and Useful Information **U1. Current Use:** Public reporting (disclosure to performance results to the public at large) Quality improvement with external benchmarking **U1.1. Use in Public Reporting Initiative 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 plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement)*	**3a** **H☐** **M☐** **L☐**

The ABMS REF has only recently completed the development and testing of its Episode-based Resource Use Measures. The Robert Wood Johnson Foundation (RWJF) has provided follow-up funding in the form of technical assistance to Aligning Forces for Quality communities for continued testing of the measures—a 15-month award to Brookings Institute with a subcontract to ABMS REF for continued field testing of select measures in up to four Aligning Forces for Quality (AF4Q) communities toward the goal of public reporting and quality improvement benchmarking. **U1.2. Use in QI** *(If used in improvement programs, provide name of program(s), locations, Web page URL(s)).* See Section U1.1 **U1.3. 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).* See Section U1.1	**I**☐
U2. Testing of Interpretability *(Provide a rationale for why the measure performance results are meaningful, understandable, and useful to the intended audience(s) for both public reporting and quality improvement).* **U2.1. If understanding or usefulness was demonstrated** *(e.g., through systematic feedback from users, focus group, cognitive testing, analysis of quality improvement initiatives) describe the data, methods, and results.* The ABMS REF measures have not yet been tested for usefulness or interpretability. They are currently undergoing continued testing in up to four RWJF AF4Q communities.	**3b** **H**☐ **M**☐ **L**☐ **NA**☐
U2.2. Resource use data and result can be decomposed for transparency and understanding. *Refer to items S11 -S12.3.*	**3c** **H**☐ **M**☐ **L**☐ **I**☐
U3. If there are similar or related measures (either same measure focus or target population) measures (both the same measure focus and same target population), list the NQF # and title of all related and/or similar measures. **U3.1. 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?** **U3.2. If the measure specifications are not completely harmonized identify the differences, rationale, and impact on interpretability and data collection burden.** *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.)*	**3d** **H**☐ **M**☐ **L**☐ **I**☐ **NA**☐
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for *Usability*?	

Steering Committee: Overall, to what extent was the criterion, *Usability*, met? **Rationale:**	**H☐** **M☐** **L☐**
FEASIBILITY	
Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement.	Eval Rating
F1. Data Elements Generated as Byproduct of Care Processes *How are the data elements needed to compute measure scores generated? Data used in the measure are:* Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)	**4a** **H☐** **M☐** **L☐** **I☐**
F2. Electronic Sources *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 in electronic claims **F2.1. 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.**	**4b** **H☐** **M☐** **L☐** **I☐**
F3. Susceptibility to Inaccuracies, Errors, or Unintended Consequences *Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to minimize or prevent. If audited, provide results.* • The majority of measures developed for this project are of 12 months duration or less with identification of the population in one year and measurement in the following. This resulted in eligibility criteria requiring a minimum of 24 months of continuous data (full medical and pharmacy benefit enrollment). Often, clinical workgroup members expressed a desire to extend the duration of a measure to encompass more longitudinal clinical outcomes (e.g. cardiac complications for diabetes) however this was not practical due to the typical enrollment patterns in the commercial population. • Sample size may be of concern for implementers seeking to measure resource use at the level of the individual provider. Many of the measures, when tested on commercial datasets, resulted in small sample sizes that may prohibit meaningful attribution. Discontinuous medical coverage and missing pharmacy coverage were responsible for significant (often greater than 50%) decreases in eligible populations, emphasizing the trade-offs between ensuring adequate sample size and achieving specificity/homogeneity in the measure denominator. If users are unable to achieve adequate sample size at the level of the individual provider, the measures specifications may still provide valuable information at the level of group, system or region. • Administrative claims lack the detail necessary to fully understand appropriateness of resource use in relation to severity of disease (e.g. bundled hospital payments, absence of cancer staging information, absence of cardiac severity indicators, Type 1 v. Type 2 diabetes). Future efforts should consider the integration of administrative claims with other sources of clinical information such as registries and electronic health records. • Resource use is only one component of efficiency measurement. The measures created in this project are not intended to be used in isolation to evaluate physician performance; rather they are intended to complement quality measures as an important component of performance evaluation. • The measures developed in this project represent a small subset of clinical conditions, and do not address the full range of patient and provider experience. Each measure was developed independently and, as such, they are not summative. Efforts to sum multiple measures will result in double counting of services. • The standardized pricing algorithms used for testing the measures were developed for use in the Marketscan dataset. The technical appendices accompanying the measures provide a guide to assist users in developing their own	**4c** **H☐** **M☐** **L☐** **I☐**

set of standardized prices unique to their datasets. Until a national list of standardized prices is made available to the general public, the methods employed in the testing phase of this project do not allow for national benchmarking.

F4. Data Collection Strategy

Describe what you have learned/modified as a result of testing regarding barriers to operational use of the measure (e.g., availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, cost of proprietary measures).

Administrative claims lack the detail necessary to fully understand appropriateness of resource use in relation to severity of disease (e.g. bundled hospital payments, absence of cancer staging information, absence of cardiac severity indicators, Type 1 v. Type 2 diabetes). Future efforts should consider the integration of administrative claims with other sources of clinical information such as registries and electronic health records.

There were several lessons learned throughout the development and testing of the ABMS REF episode-based resource use measures. First, was the importance of garnering a diverse range of clinical input in a transparent manner to foster face validity and acceptance in the clinical community. Second was the importance of adequate resources for data acquisition, preparation and analyses (time and personnel). Not all datasets are formatted the same which can lead to significant amounts of programmer time for re-formatting code or datasets. It is also important to allow 2-6 months lead time to negotiate data use agreements as use of health care data–even de-identified data--often involves complex contract negotiations.

4d

H☐ M☐ L☐ I☐

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

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

H☐ M☐ L☐

RECOMMENDATION

Steering Committee: Do you recommend for endorsement?
Comments:

Y☐ N☐ A☐

CONTACT INFORMATION

Co.1 Measure Steward (Intellectual Property Owner)

Co.1 Organization

American Board of Medical Specialties Research and Education Foundation, 222 N. LaSalle St, Suite 1500, Chicago, Illinois, 60601

Co.2 Point of Contact

Kevin, Weiss, MD, kweiss@abms.org, 312-436-2600-

Measure Developer If different from Measure Steward

Co.3 Organization

American Board of Medical Specialties Research and Education Foundation, 222 N. LaSalle St, Suite 1500, Chicago, Illinois, 60601

Co.4 Point of Contact

Kevin, Weiss, MD, kweiss@abms.org, 312-436-2600-

Co.5 Submitter If different from Measure Steward POC

Robin, Wagner, rwagner@abms.org, 312-436-2605-, American Board of Medical Specialties Research and Education Foundation

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

Development of the ABMS REF Episode-based Resource Use Measures was supported by the Robert Wood Johnson Foundation under the High Value Healthcare Project: Characterizing Episodes and Costs of Care. Grant number 63609.

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.

Low Back PainWorkgroup Members
Larry Benz, DPT, American Physical Therapy Association
Zoher Ghogawala, MD, American Association of Neurological Surgeons
Debra Gordon, RN, American Nurses Association
Maureen Hanlon, MD, Kaiser Permanente
Thomas James III, MD, America's Health Insurance Plans, Humana
George McClelland, MD, American Chiropractic Association
C. Douglas Phillips, MD, American College of Radiology
Richard Snow, MD, American Osteopathic Association
Jeffrey Susman, MD, American Academy of Family Physicians
William Watters III, MD, American Academy of Orthopedic Surgery
David Wong, MD, North American Spine Society

Workgroups consisting of a panel of experts were assembled for each condition. In collaboration with the AMA PCPI, a formal call for nominations was issued to the PCPI membership. This process was supplemented with direct outreach to relevant organizations in an effort to achieve representation from a wide range of clinical expertise (medical, nursing, pharmacy, other allied health professionals). Workgroup members were selected based on their clinical knowledge and administrative experience—many also had significant experience in developing quality measures. Where possible, groups also included technical expertise from the health plan perspective.
The measure development process involved a series of deliberate steps where participating clinicians took into account the natural progression of a condition and existing best practices before carefully considering how to best use administrative claims data to construct the episode.

Each clinical workgroup initially convened for a two-day in-person meeting that began with an introduction to the concepts of episodes of care and resource use measurement-- including a review of the NQF framework for evaluating efficiency across episodes of care. The groups were then asked to conceptualize one or more episodes based on the phases of the NQF model. They aimed to identify clinically homogenous populations so that the measures would be sensitive to provider decisions and existing practice protocols for like patients. Workgroup members were then asked to conceptualize the measure specifications based on their combined knowledge of guidelines, evidence, and clinical experience. The workgroups helped to define the denominator, duration, clinically relevant services and attribution of each episode as related to the clinical progression and treatment of the condition.

Throughout the months following the in-person meeting, project staff then worked to translate the concepts into detailed written measure specifications. The workgroups subsequently re-convened via a series of conference calls to review data analyses, share expert opinions, consider additional evidence-based literature, revise and finalize the measure specifications.

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released:

2010

Ad.3 Month and Year of most recent revision:

12, 2010

Ad.4 What is your frequency for review/update of this measure?

every 3 years

Ad.5 When is the next scheduled review/update for this measure?

12, 2013

Ad.6 Copyright statement/disclaimers:

Ad. 7 Date of Submission (*MM/DD/YY*)**:**

04/21/2011

Variable Name	Variable Description	Required Data Sources*
admdate	Date of Admission	A
age	Age	E
billtyp	Facility Bill Type Code	C
days	Length of Stay	A
daysupp	Day's Supply	D
disdate	Date of Discharge	A
drg	Diagnosis related group	A,B
dstatus	Discharge status	A
egeoloc	Geographic Location	E
enrolid	Enrollee ID	All
fachdid	Facility Header Record ID	C
facprof	Professional/Facility Indicator	C
gennme	Generic Drug Name	D
mastfrm	Master Form Code	D
memdays	Member Days	E
ndcnum	National Drug Code (ndc_code in Redbook)	D
pay	Payment	A,B,C,D
pdx,dx1,dx2,…,dxn	Diagnosis Codes	A,B,C
physid	Physician ID	A,B
pproc, pproc1,…, pprocn	Procedure/Service Codes	A,B,C
procmod	Procedure Code Modifier	A,C
proctyp	Procedure Code Type	B,C
prodnme	Product Name	D
provid	Provider ID	A
qty	Quantity of Services	A,B,C,D
region	Region	E
revcode	Revenue Code	C
rx	Cohort Drug Indicator	D
sex	Gender	E
stdplac	Place of Service	C
stdprov	Provider Type	C
svcdate	Service Date	A,B,C,D
thercls	Therapeutic Class	D
tsvcdat	Date Service Ending	C

Data Sources*

A. Administrative claims data – inpatient (facility)
B. Administrative claims data – inpatient (professional)
C. Administrative claims data – outpatient/ambulatory (professional and facility)
D. Administrative claims data – pharmacy
E. Enrollment/coverage data (2 or more years)

Measure Component	**Required Variables**
Standardized Prices*	enrolid, ndcnum, pay, qty, drg, pproc,…,pprocn.
Exclusions and standard coverage definition	enrolid, pdx,dx1,…,dxn, age, svcdate, pproc, pproc1,…, pprocn, pay, qty, revcode, memdays, rx, stdplac, proctyp.
Cohort Definition	enrolid, svcdate, pdx, pdx1,…,pdxn, pproc1,…, pprocn, pay, qty, sex, age, thercls, dstatus, stdplac, billtyp, fachdid, revcode.
Related Resource Use	enrolid, facprof, pay, qty, pproc1,…, pprocn, svcdate, admdate, disdate, pdx, dx1,…, dxn, drg, ndcnum, thercls, gennme, prodnme, daysupp, procmod, mastfrm.
Output and Attribution	enrolid, svcdate, standardized price variables*, BETOS**, pproc1,…,pprocn, pdx, dx1,…,dxn, egeoloc, region, provid, stdprov, age, sex, physid.

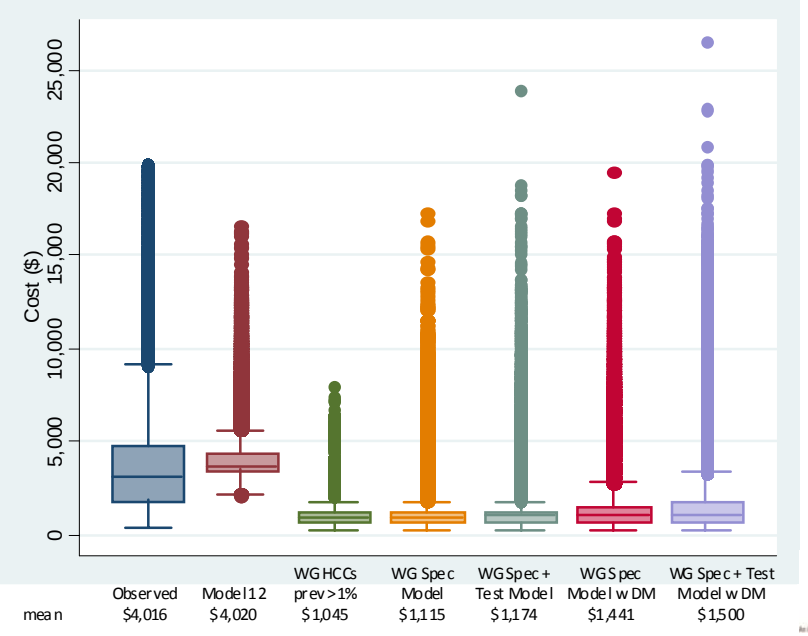
* For internal testing and validation purposes, drug prices were calculated by taking the average of 2006 and 2007 Marketscan prices, inpatient facility prices were computed by calculating average daily price by DRG from 2007, and outpatient and service prices were constructed by calculating the mean price by procedure code within the Marketscan dataset.
** Berenson-Eggers Type of Service – Categorizes Health Care Procedure Coding System (HCPCS) procedure codes in order to analyze health care expenditures. See link for full description.
http://www.cms.hhs.gov/hcpcsreleasecodesets/20_betos.asp

Condition (Workgroup)	***Measure Name***	***Abbreviation***
Acute Myocardial Infarction (AMI)	Episode-of-Care for 30 days Following Onset	AMI1
Acute Myocardial Infarction (AMI)	Episode-of-Care for Post-Acute Period (Days 31-365 Days Post-Event)	AMI2
Asthma	Episode-of-Care for Patients with Asthma over a 1-year Period	ASTH
Breast Cancer	Episode-of-Care for 60-Day Period Preceding Breast Biopsy	BB
Breast Cancer	Episode-of-Care for Treatment in Newly Diagnosed Cases of Breast Cancer over a 15-month Period	BCT
Chronic Obstructive Pulmonary Disease (COPD)	Episode-of-Care for Patients with Stable COPD over a 1-year Period	COPD1
Chronic Obstructive Pulmonary Disease (COPD)	Episode-of-Care for Patients with Unstable COPD over a 1-year Period	COPD2
Colon Cancer	Episode-of-Care for 21-Day Period Around Colonoscopy	COL
Colon Cancer	Episode-of-Care for Treatment of Localized Colon Cancer	CCT
Congestive Heart Failure (CHF)	Episode-of-Care for Management of CHF Over 1-Year Period	CHF1
Congestive Heart Failure (CHF)	Episode-of-Care for Post Hospitalization Management of CHF over 4-Month Period	CHF2
Coronary Artery Disease (CAD)	Episode-of-Care for Management of Chronic CAD Over 1-Year Period	CAD1
Coronary Artery Disease (CAD)	Episode-of-Care for Management of CAD Post Revascularization Over 1-Year Period	CAD2
Diabetes	Episode-of-Care for Diabetes Over 1-Year Period	DIAB
Low Back Pain	Episode-of-Care for Simple Non-Specific Lower Back Pain (Acute and Sub-Acute)	LBP1
Low Back Pain	Episode-of-Care for Acute/Sub-Acute Lumbar Radiculopathy With or Without Lower Back Pain	LBP2
Pneumonia	Episode-of-Care for Community-Acquired Pneumonia Hospitalization	PN1
Pneumonia	Episode-of-Care for Ambulatory Pneumonia Episode	PN2

Comparison 'off the shelf' HCC Values with Episode-specific Risk Adjustment Model

Below we show the figure for the comparison of the diabetes risk adjustment model with diabetes risk adjustment models if we had used HCC values. The first box plot in the figure shows the observed costs in for the episode. The second box plot shows the risk adjustment model that we developed for our diabetes episode that is focused on diabetes-related costs. The final five box plots show the distribution of predicted costs including different HCCs for our diabetes episode if we had relied on the off the shelf HCC values. The mean predicted value for all of the off the shelf HCCs models is $1500 or less, while the observed episode costs were slightly more than $4,000. Given the disparity in the means and distributions of the off the shelf HCC values we felt this justified our approach to develop risk adjustment models for each of our episodes that were focused on episode specific costs

Observed and Predicted Values – Diabetes Episode with "off the shelf HCCs"



For this reason, we have developed separate risk adjustment models for each of our episodes that are based on episode-specific costs. We realize this increases the complexity of implementing our measures; however, we feel it is a more appropriate approach for risk adjustment within our episodes. Within our risk adjustment approach, we control for different comorbidities for each condition because patients with each of the measurement conditions often had very different risk profiles.

We used the following risk adjustment strategy in the development of our risk adjustment models:

1. Utilized quasi-Modified Delphi approach with the condition-specific workgroup to categorize HCCs into three groups:

- Include in risk adjustment model;
- Exclude in risk adjustment model; and
- Test impact in risk adjustment model.

2. Identified HCCs in denominator population during the 12 months before the measurement year.

3. Tested 12 different model specifications shown in Table 1 (below), where the HCCs included in the model varied, and the distribution and link functions in the generalized linear models also varied. Models were developed in a stepwise manner as indicated. The first four models used a gamma distribution and a log link function. This functional form of the model was selected as cost data are typically skewed and we wanted to account for that in the analysis. The first model included all HCCs identified by the condition-specific workgroup as "Include HCCs" with a prevalence in the population of >=1%. The second model was a reduction of the first model that only included HCCs where $p<0.1$. The third model extended the second model by including HCCs with prevalence >=1% identified as "Test HCCs" by the condition-specific workgroup. The fourth model was a reduction of the third model and included only those HCCs where $p<0.1$. The next set of four models (Models 5-8) repeated the process of the first four models but used a normal distribution and identity link function. We opted to include this functional form of the model so that the model output could be interpreted in dollars without requiring a transformation. We followed this strategy as we felt it would be easier for those implementing our measure to create their own risk adjustment models using this functional form of the model if they decided to create their own models. Finally, we opted to evaluate models that included all of the HCCs in case the work group may have failed to include HCCs that were influential on the overall episode costs. Model 9 used all of the HCCs, with the exception of the HCC for the episode being evaluated (e.g., diabetes for the diabetes episode; however HCCs for complications of diabetes were included), and a gamma distribution with log link function. Model 10 was a reduction of Model 9 where only the HCCs with $p<0.1$ were included. The final two models (Models 11-12) used the same process as Models 9 and 10 with a normal distribution and identity link function.

Table 1. Risk Adjustment Model Specifications

Model #	Independent Variables						Distri-bution	Link function
	WG Specified (> 1%)	WG specified (> 1%) p < 0.1	Test conditions (> 1%)	Test conditions (> 1%) p < 0.1	All HCCs	All HCCs p < 0.1		
1	X						Gamma	Log
2		X					Gamma	Log
3		X	X				Gamma	Log
4		X		X			Gamma	Log
5	X						Normal	Identity
6		X					Normal	Identity
7		X	X				Normal	Identity
8		X		X			Normal	Identity
9					X		Gamma	Log
10						X	Gamma	Log
11					X		Normal	Identity
12						X	Normal	Identity

4. Models were developed in a split sample approach with 75% of the population randomly selected for model development and the remaining 25% used in model evaluation. Model performance was also evaluated in the full cohort.

5. The performance of each model was evaluated through comparisons of the observed and predicted distributions, comparisons of residuals, comparisons of absolute differences between observed and predicted, comparisons of observed-to-predicted ratios, and comparisons of mean squared errors across models. Summary information on model performance was presented to the condition-specific workgroup for selection of a risk adjustment model for the condition. Final model selection was based on the best performing model across metrics. Where model performance was similar, models using the normal distribution were preferentially chosen over the gamma distribution models for ease of implementation. More parsimonious models were also preferentially chosen.

Example Episode Report

LBP Unspecified Episode

Report for Physician #18310983

Provider type = Chiropractor

	MD	Peer Group	Non-Peer Group	National Avg
Episodes	16	44,545	118,910	163,471
Observed Costs*				
Average	$ 265	$ 762	$ 1227	$ 1100
Min	$ 58	$ 58	$ 58	$ 58
Median	$ 212	$ 411	$ 590	$ 514
Max	$ 687	$ 8289	$ 8289	$ 8289
Predicted Costs				
Average	$ 1061	$ 1082	$ 1101	$ 1096
Min	$ 994	$ 446	$ 473	$ 446
Median	$ 1062	$ 1074	$ 1083	$ 1082
Max	$ 1136	$ 2220	$ 2619	$ 2619
Observed-to-Expected Ratio				
Average	0.25	0.71	1.12	1.00
Min	0.05	0.03	0.03	0.03
Median	0.19	0.38	0.54	0.47
Max	0.63	8.63	11.16	11.16
% ≥ 2.0	0.0%	7.5%	16.0%	13.7%
% ≥ 2.5	0.0%	4.6%	11.6%	9.7%
% ≥ 75th percentile peers	0.0%	(0.0%, 20.6%)		

* Observed costs adjusted for outliers (windsorized)

Notes:
- Use Model 12

American Board of Medical Specialties

Higher standards. Better care.®

Research and Education Foundation

Analytic Findings: Low Back Pain – Simple Non-Specific Episode of Care

NQF Submission

Overview of Analyses Presented for LBP Episode*

- Denominator Attrition
- Related and Non-related Services
- Resource Use, Attribution and
- Risk Adjustment

** The following results are based on the measure specification at different points in time, so the numbers are not always consistent, but they are not substantively different.*

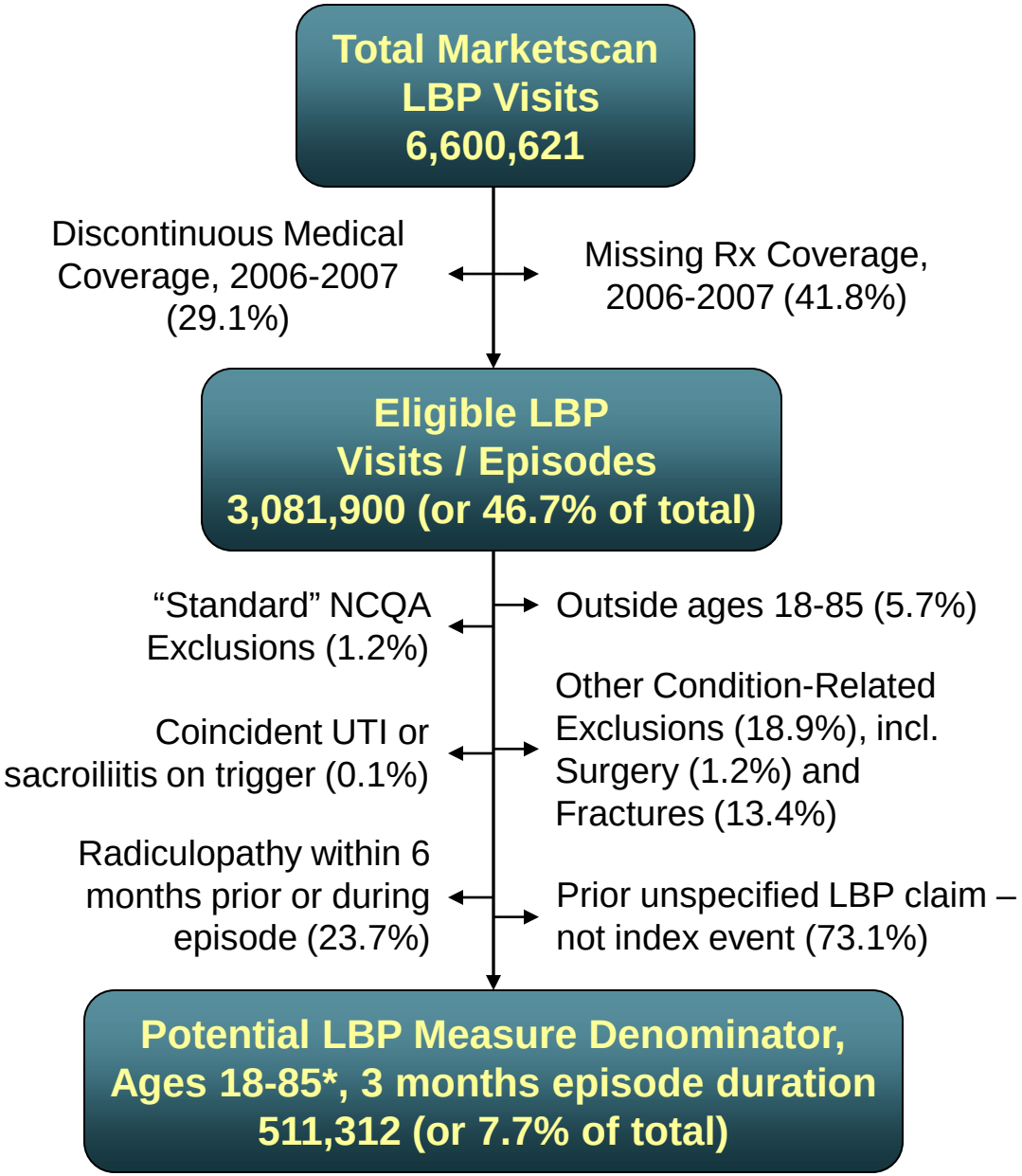
Denominator Attrition

- Summarizes the initial denominator based on the workgroup's specifications
- Describes the percentage of enrollees removed from the analysis due to NCQA exclusions or other criteria.

Unspecified LBP Measure Denominator

- 3 mos. (or 6 wks.) of care for patient with 1+ ambulatory care visits with non-specific LBP ICD-9 code (expanded list): 724.2, 724.5, 724.8, 724.9, 739.3, 739.4 or 847.2
- Ages 18-85*, to be stratified into elderly and non-elderly groups
- Episode triggers between July 1, 2006 and August 31, 2007
- Test data: Marketscan 2006-2007
- Note: exclusions are not additive (double-counting occurs often)

* Marketscan database contains data for non-elderly population only (under 65); as such, no data are presented for the elderly (65+) group



Related and Non-Related Services

- Examines most frequent related and non-related resource use by BETOS category
  - Evaluation and Management Visits, Procedures, Imaging, Tests, Admissions and Medications.
- Results are presented to the workgroup to examine the face validity of episodes.

Resource Use by Type of Service: LBP Unspecified - 3MO

- All episode triggers

Description	Mean	% of Total	5th %	25th %	50th %	75th %	95th %
Inpatient facility charge	$5	1%	$0	$0	$0	$0	$0
OP Facility Costs	$29	8%	$0	$0	$0	$0	$0
Evaluation and Management	$87	22%	$0	$44	$65	$107	$248
Procedures	$99	26%	$0	$0	$0	$47	$534
Imaging	$29	8%	$0	$0	$0	$0	$103
Tests	$4	1%	$0	$0	$0	$0	$6
Durable Medical Equipment	$2	1%	$0	$0	$0	$0	$0
Other Services	$66	17%	$0	$0	$0	$70	$342
Unclassified	$1	0%	$0	$0	$0	$0	$0
Drug Costs	$64	17%	$0	$0	$3	$43	$281
Sum of charges	**$385**	**100%**	**$55**	**$99**	**$183**	**$395**	**$1,357**

Note: results do not reflect the specialty of the physician providing/ordering the care (e.g., chiropractors may bill using a “standard” E&M CPT code)

Top 20 Related E&M, LBP Unspecified Episode – 3MO

- 24% of total episode costs

CPT	Svcs.	Cost	% of Svcs	% of Cost	Description
99213	184,668	$15,200,553	43.5%	39.9%	Office or other outpatient visit, established patient
99214	68,411	$6,514,817	16.1%	17.1%	Office or other outpatient visit, established patient
99203	31,900	$3,398,088	7.5%	8.9%	Office or other outpatient visit, new patient
97001	22,829	$1,788,780	5.4%	4.7%	Physical therapy evaluation
99204	10,914	$1,637,406	2.6%	4.3%	Office or other outpatient visit, new patient
99212	31,016	$1,408,460	7.3%	3.7%	Office or other outpatient visit, established patient
99283	10,591	$1,240,821	2.5%	3.3%	Emergency department visit for E&M care
99202	16,515	$1,232,452	3.9%	3.2%	Office or other outpatient visit, new patient
99284	6,615	$1,215,655	1.6%	3.2%	Emergency department visit for E&M care
99215	5,667	$887,257	1.3%	2.3%	Office or other outpatient visit, established patient
99244	3,648	$730,892	0.9%	1.9%	Office consultation for a new or established patient
99243	3,348	$500,873	0.8%	1.3%	Office consultation for a new or established patient
99285	1,525	$427,679	0.4%	1.1%	Emergency department visit for E&M care
99205	2,007	$380,574	0.5%	1.0%	Office or other outpatient visit, new patient
97002	5,137	$271,477	1.2%	0.7%	Physical therapy re-evaluation
99245	994	$259,198	0.2%	0.7%	Office consultation for a new or established patient
99201	5,245	$236,431	1.2%	0.6%	Office or other outpatient visit, new patient
99211	6,338	$164,348	1.5%	0.4%	Office or other outpatient visit, established patient
99242	720	$81,339	0.2%	0.2%	Office consultation for a new or established patient
99396	578	$72,486	0.1%	0.2%	Period comprehensive preventive medicine re-evaluation
Total	424,564	$38,072,149	100.0%	100.0%	

Non-related E&M, Top 20 ICD-9 Codes, LBP Unspecified Episode – 3MO

ICD-9 Code	Related	Not Related	Related Costs	Non-Related Costs
V7231-Routine Gyn Examination	150	23,061	$18,413	$2,858,056
V700 -Routine Medical Exam	395	15,825	$49,040	$1,947,045
4011 -Benign Hypertension	819	16,518	$77,548	$1,585,112
4019 -Hypertension NOS	613	12,424	$51,803	$1,250,367
25000-Dm II wo Cmp Nt St Uncntr	203	11,075	$18,579	$1,175,041
78650-Chest Pain NOS	333	7,665	$42,802	$1,170,419
4770 -Rhinitis Due to Pollen	44	17,441	$3,479	$1,029,982
78900-Abdmnal Pain Unspcf Site	496	7,790	$54,689	$1,008,500
4619 -Acute Sinusitis NOS	235	11,340	$20,636	$1,001,635
4779 -Allergic Rhinitis NOS	204	14,514	$18,282	$934,229
30928-Adjust Dis w Anxiety/Dep	1	9,731	$94	$918,494
7231 -Cervicalgia	1,821	9,080	$162,986	$837,038
7840 -Headache	347	6,079	$33,331	$738,396
4660 -Acute Bronchitis	154	7,816	$13,885	$728,295
2724 -Hyperlipidemia NEC/NOS	350	8,197	$34,106	$726,640
4659 -Acute URI NOS	222	8,236	$22,371	$712,052
5990 -Urin Tract Infection NOS	54	6,988	$4,807	$689,972
3004 -Dysthymic Disorder	17	7,345	$1,890	$686,442
311 -Depressive Disorder NEC	51	6,631	$3,948	$649,378
29632-Recurr Depr Psychos-Mod	1	6,863	$77	$628,576

Top 20, Related Imaging, LBP Unspecified Episode – 3MO

- 8% of total episode costs

CPT	Svcs.	Cost	% of Svcs	% of Cost	Description
72148	9,734	$4,496,824	6.3%	30.5%	MRI, spinal canal and contents, lumbar; without contrast
72100	53,964	$2,402,954	35.2%	16.3%	Radiologic exam, spine, lumbosacral; two or three views
72110	24,555	$1,278,500	16.0%	8.7%	Radiologic exam, spine, lumbosacral; minimum of four views
72141	2,175	$919,176	1.4%	6.2%	MRI, spinal canal and contents, cervical; without contrast
72158	1,030	$763,649	0.7%	5.2%	MRI, spinal canal and contents, with & w/out contrast; lumbar
72146	1,356	$599,375	0.9%	4.1%	MRI, spinal canal and contents, thoracic; without contrast
72040	9,558	$417,978	6.2%	2.8%	Radiologic examination, spine, cervical; two or three views
72070	7,983	$333,344	5.2%	2.3%	Radiologic examination, spine; thoracic, two views
72010	2,389	$201,624	1.6%	1.4%	Radiologic examination, spine, entire, survey study
72131	760	$160,237	0.5%	1.1%	Computed tomography, lumbar spine; without contrast
77003	1,979	$157,075	1.3%	1.1%	Fluoroscopic guidance and localization of needle or catheter
72156	208	$154,875	0.1%	1.0%	MRI, spinal canal and contents, with & w/out contrast; cervical
72050	2,668	$149,039	1.7%	1.0%	Radiologic examination, spine, cervical; minimum of 4 views
27096	518	$147,816	0.3%	1.0%	Injection procedure for sacroiliac joint
76005	1,305	$140,111	0.9%	0.9%	Fluoroscopic guidance and localization of needle or catheter
72114	1,799	$138,314	1.2%	0.9%	Radiologic exam, spine, lumbosacral; complete
72157	185	$120,969	0.1%	0.8%	MRI, spinal canal and contents, with & w/out contrast; thoracic
72170	3,309	$116,041	2.2%	0.8%	Radiologic examination, pelvis; one or two views
72192	628	$108,239	0.4%	0.7%	Computed tomography, pelvis; without contrast material
72072	2,893	$103,941	1.9%	0.7%	Radiologic examination, spine; thoracic, three views
Total	153,373	$14,764,280	100.0%	100.0%	

Common Non-related Imaging, LBP Unspecified Episode – 3MO

CPT	Label	Related	Not Related	Related Costs	Non-Related Costs
78465	Myocardial perfusion imaging; tomographic (SPECT), multiple stu	6	5,793	$1,663	$2,675,099
70553	Magnetic resonance (eg, proton) imaging, brain (including brain st	43	3,112	$26,909	$2,133,955
73721	Magnetic resonance (eg, proton) imaging, any joint of lower extren	101	4,573	$48,910	$2,049,889
93307	Echocardiography, transthoracic, real-time with image documenta	28	8,223	$6,160	$1,659,797
74150	Computed tomography, abdomen; without contrast material	612	9,116	$100,298	$1,468,468
72192	Computed tomography, pelvis; without contrast material	628	8,881	$108,239	$1,468,259
74160	Computed tomography, abdomen; with contrast material(s)	132	6,449	$26,012	$1,374,975
72193	Computed tomography, pelvis; with contrast material(s)	153	6,998	$28,504	$1,293,694
74170	Computed tomography, abdomen; without contrast material, follov	96	4,433	$27,884	$1,244,779
72141	Magnetic resonance (eg, proton) imaging, spinal canal and conter	2,175	2,401	$919,176	$1,073,262
76092	Screening mammogram	9	10,347	$743	$1,056,255
76830	Ultrasound, transvaginal	157	9,407	$18,404	$1,050,566
71020	Radiologic examination, chest, two views, frontal and lateral;	2,669	29,710	$98,731	$994,418
73221	Magnetic resonance (eg, proton) imaging, any joint of upper extrer	52	2,222	$23,921	$994,218
77057	Screening mammography, bilateral (2-view film study of each brea	16	13,039	$1,268	$983,038
93325	Doppler echocardiography color flow velocity mapping (List separa	27	8,649	$3,151	$886,374
G0202	Screening mammography, producing direct digital image, bilateral	4	7,994	$535	$815,571
93320	Doppler echocardiography, pulsed wave and/or continuous wave v	29	8,753	$2,913	$798,406
72148	Magnetic resonance (eg, proton) imaging, spinal canal and conter	9,734	1,515	$4,496,824	$728,391
76075	Dual energy x-ray absorptiometry (DXA), bone density study, one	110	3,759	$21,659	$686,959

Top 20, Related Outpatient Facility, LBP Unspecified Episode – 3MO

- 8% of total episode costs

CPT	Svcs.	Cost	% of Svcs	% of Cost	Description
72148	1,549	$1,645,076	1.8%	11.2%	MRI, spinal canal and contents, lumbar; without contrast
99283	3,321	$953,333	3.9%	6.5%	Emergency department visit
97110	10,187	$925,459	12.0%	6.3%	Therapeutic procedure, one or more areas, each 15 minutes
72110	3,517	$743,728	4.1%	5.0%	Radiologic exam, spine, lumbosacral; min. of four views
99284	1,242	$577,805	1.5%	3.9%	Emergency department visit
72100	3,367	$485,832	4.0%	3.3%	Radiologic exam, spine, lumbosacral; two or three views
74150	641	$482,781	0.8%	3.3%	Computed tomography, abdomen; without contrast material
72192	611	$439,777	0.7%	3.0%	Computed tomography, pelvis; without contrast material
99282	2,284	$431,772	2.7%	2.9%	Emergency department visit
72141	338	$368,395	0.4%	2.5%	MRI, spinal canal and contents, cervical; without contrast
72158	211	$311,243	0.2%	2.1%	MRI, spinal canal and contents, with and without contrast
97001	2,219	$279,461	2.6%	1.9%	Physical therapy evaluation
97140	3,591	$229,946	4.2%	1.6%	Manual therapy techniques, one or more regions, 15 min.
62311	369	$224,649	0.4%	1.5%	Injection, single, of diagnostic or therapeutic substance(s)
72146	192	$213,875	0.2%	1.4%	MRI, spinal canal and contents, thoracic; without contrast
90772	2,229	$200,706	2.6%	1.4%	Therapeutic, prophylactic or diagnostic injection
72193	184	$147,935	0.2%	1.0%	Computed tomography, pelvis; with contrast material(s)
64475	192	$142,793	0.2%	1.0%	Injection, anesthetic agent and/or steroid
74160	161	$134,896	0.2%	0.9%	Computed tomography, abdomen; with contrast material(s)
72131	171	$126,325	0.2%	0.9%	Computed tomography, lumbar spine; with and w/out contrast
Total	85,074	$14,750,872	100.0%	100.0%	

Select Non-related Outpatient Facility, LBP Unspecified Episode – 3MO

CPT	Label	Related	Not Related	Related Costs	Non-Related Costs
99284	Emergency department visit for the evaluation and management of a	1,242	3,681	$577,805	$1,695,343
99283	Emergency department visit for the evaluation and management of a	3,321	4,564	$953,333	$1,311,886
50590	Lithotripsy, extracorporeal shock wave	1	331	$2,221	$1,197,724
74150	Computed tomography, abdomen; without contrast material	641	1,669	$482,781	$1,169,557
72192	Computed tomography, pelvis; without contrast material	611	1,610	$439,777	$1,116,891
72193	Computed tomography, pelvis; with contrast material(s)	184	1,341	$147,935	$1,104,569
74160	Computed tomography, abdomen; with contrast material(s)	161	1,191	$134,896	$1,025,807
97110	Therapeutic procedure, one or more areas, each 15 minutes; therap	10,187	8,430	$925,459	$888,265
99285	Emergency department visit for the evaluation and management of a	154	1,244	$105,383	$857,539
74170	Computed tomography, abdomen; without contrast material, followe	90	813	$95,199	$799,629
71020	Radiologic examination, chest, two views, frontal and lateral;	719	4,371	$87,542	$555,852
73721	Magnetic resonance (eg, proton) imaging, any joint of lower extremit	27	485	$34,407	$539,715
99282	Emergency department visit for the evaluation and management of a	2,284	2,684	$431,772	$508,885
71260	Computed tomography, thorax; with contrast material(s)	68	571	$58,976	$491,767
88305	Level IV - Surgical pathology, gross and microscopic examination Al	5	2,699	$575	$456,251
90774	Therapeutic, prophylactic or diagnostic injection (specify substance (	820	2,942	$92,956	$344,287
76700	Ultrasound, abdominal, real time with image documentation; comple	72	979	$25,953	$326,963
72194	Computed tomography, pelvis; without contrast material, followed by	39	344	$34,102	$314,850
72141	Magnetic resonance (eg, proton) imaging, cervical (w/o)	338	295	$368,395	$312,755
47563	Laparoscopy, surgical; cholecystectomy with cholangiography		174		$303,058
97140	Manual therapy techniques (eg, mobilization/ manipulation, manual	3,591	3,352	$229,946	$245,721
97001	Physical therapy evaluation	2,219	1,577	$279,461	$217,735
72148	Magnetic resonance (eg, proton) imaging, lumbar (w/o)	1,549	144	$1,645,076	$161,833
90772	Therapeutic, prophylactic or diagnostic injection (specify substance (	2,229	1,985	$200,706	$155,504
72110	Radiologic examination, spine, lumbosacral; minimum of four views	3,517	424	$743,728	$94,072
72158	Magnetic resonance (eg, proton) imaging, lumbar (w, w/o)	211	45	$311,243	$67,738
72100	Radiologic examination, spine, lumbosacral; two or three views	3,367	445	$485,832	$66,238

Top 20, Related Procedures, LBP Unspecified Episode – 3MO

- 26% of total episode costs

CPT	Svcs.	Cost	% of Svcs	% of Cost	Description
97110	197,806	$10,849,193	17.3%	27.1%	Therapeutic exercises to develop strength, each 15 minutes
97140	171,848	$7,160,363	15.0%	17.9%	Manual therapy techniques (e.g., mobilization, manipulation)
97014	172,990	$2,918,742	15.1%	7.3%	Electrical stimulation (unattended)
97112	62,634	$2,599,984	5.5%	6.5%	Neuromuscular reeducation of movement, balance, etc.
97530	47,725	$2,537,561	4.2%	6.3%	Therapeutic activites; dynamic activities to improve performanc
97012	144,717	$2,483,722	12.6%	6.2%	Application of traction, mechanical
97035	78,967	$1,243,867	6.9%	3.1%	Application of ultrasound, each 15 minutes
97124	32,751	$1,185,067	2.9%	3.0%	Massage, including effleurage, petrissage and/or tapotement
97032	56,179	$1,182,374	4.9%	3.0%	Electrical stimulation (manual), each 15 minutes
97010	82,937	$809,316	7.2%	2.0%	Application of hot or cold packs
63075	257	$407,672	0.0%	1.0%	Discectomy, anterior, with decompression of spinal cord
22554	272	$335,807	0.0%	0.8%	Arthrodesis, anterior interbody technique, including min. discec
G0283	21,466	$334,762	1.9%	0.8%	Electrical stimulation (unattended) for indication, part of therapy
97810	5,692	$286,567	0.5%	0.7%	Acupuncture without electrical stimulation, initial 15 minutes
90772	9,709	$262,895	0.8%	0.7%	Therapeutic, prophylactic or diagnostic injection
22845	240	$258,536	0.0%	0.6%	Anterior instrumentation; 2 to 3 vertebral segments
97813	3,913	$217,653	0.3%	0.5%	Acupuncture with electrical stimulation, initial 15 minutes
97113	2,149	$203,575	0.2%	0.5%	Aquatic therapy with therapeutic exercises
62311	736	$201,186	0.1%	0.5%	Injection, single; epidural or subarachnoid; lumbar, sacral
00670	111	$199,283	0.0%	0.5%	Anesthesia for extensive spine and spinal cord procedures
Total	1,146,645	$39,982,737	100.0%	100.0%	

Highlighted rows feature procedures more common in 3-month episodes, relative to 6-week episodes

Common Non-related Procedures, LBP Unspecified Episode – 3MO

CPT	Label	Related	Not Related	Related Costs	Non-Related Costs
97110	Therapeutic procedure, one or more areas, each 15 minutes; t	197,806	71,223	$10,849,193	$3,972,926
97140	Manual therapy techniques (eg, mobilization/ manipulation, n	171,848	59,180	$7,160,363	$2,525,038
59400	Routine obstetric care including antepartum care, vaginal deliv	0	1,014	$0	$2,362,526
45378	Colonoscopy, flexible, proximal to splenic flexure; diagnostic, v	2	4,505	$864	$1,983,616
59510	Routine obstetric care including antepartum care, cesarean de	0	560	$0	$1,378,461
00840	Anesthesia for intraperitoneal procedures in lower abdomen i	0	1,464	$0	$1,256,056
45380	Colonoscopy, flexible, proximal to splenic flexure; with biopsy	0	2,465	$0	$1,199,589
00790	Anesthesia for intraperitoneal procedures in upper abdomen i	1	1,360	$905	$1,189,945
00810	Anesthesia for lower intestinal endoscopic procedures, endos	0	2,615	$0	$1,123,773
01967	Neuraxial labor analgesia/anesthesia for planned vaginal deliv	0	1,042	$0	$1,071,500
43239	Upper gastrointestinal endoscopy including esophagus, stoma	0	3,281	$0	$1,054,734
45385	Colonoscopy, flexible, proximal to splenic flexure; with remov	0	1,643	$0	$965,324
20610	Arthrocentesis, aspiration and/or injection; major joint or burs	1,120	8,999	$104,797	$813,550
47562	Laparoscopy, surgical; cholecystectomy	0	842	$0	$784,033
27447	Arthroplasty, knee, condyle and plateau; medial AND lateral c	0	407	$0	$745,274
97112	Therapeutic procedure, one or more areas, each 15 minutes; r	62,634	16,766	$2,599,984	$722,909
97530	Therapeutic activities, direct (one-on-one) patient contact by t	47,725	13,260	$2,537,561	$709,434
19318	Reduction mammaplasty	90	379	$138,973	$655,941
90772	Therapeutic, prophylactic or diagnostic injection (specify subst	9,709	17,825	$262,895	$627,005
29881	Arthroscopy, knee, surgical; with meniscectomy (medial OR la	2	694	$2,011	$624,784

Top 20, Related "Other" Services, LBP Unspecified Episode – 3MO

- 15% of total episode costs

CPT	Svcs.	Cost	% of Svcs	% of Cost	Description
98941	333,199	$13,061,206	53.1%	57.0%	Chiropractic manipulative treatment; spinal, 3-4 regions
98940	186,058	$5,898,197	29.7%	25.7%	Chiropractic manipulative treatment; spinal, 1-2 regions
98942	40,154	$1,911,296	6.4%	8.3%	Chiropractic manipulative treatment; spinal, 5 regions
98943	38,484	$1,051,524	6.1%	4.6%	Chiropractic manipulative treatment; extraspinal, 1+ regions
A0427	326	$171,980	0.1%	0.8%	Ambulance service, advanced life support, emergency
A0429	414	$164,109	0.1%	0.7%	Ambulance service, basic life support, emergency
J1885	8,769	$107,647	1.4%	0.5%	Injection, ketorolac tromethamine, per 15 mg
A0425	715	$84,281	0.1%	0.4%	Ground mileage, per statute mile
J3490	320	$68,880	0.1%	0.3%	Unclassified drugs
90471	1,094	$60,670	0.2%	0.3%	Immunization administration
J3301	2,167	$36,152	0.3%	0.2%	Injection, triamcinolone acetonide, per 10mg
J1040	2,515	$33,789	0.4%	0.1%	Injection, methylprednisolone acetate, 80 mg
J1745	8	$23,942	0.0%	0.1%	Injection infliximab, 10 mg
J1030	2,043	$17,959	0.3%	0.1%	Injection, methylprednisolone acetate, 40 mg
J1100	2,988	$15,517	0.5%	0.1%	Injection, dexamethasone sodium phosphate, 1mg
J0696	235	$14,642	0.0%	0.1%	Injection, ceftriaxone sodium, per 250 mg
J0702	957	$13,182	0.2%	0.1%	Injection, betamethasone
J2275	26	$13,108	0.0%	0.1%	Injection, morphine sulfate
90658	763	$11,735	0.1%	0.1%	Influenza virus vaccine
A0999	54	$10,505	0.0%	0.0%	Unlisted ambulance service
Total	627,470	$22,924,533	100.0%	100.0%	

Common Non-related "Other" Services, LBP Unspecified Episode – 3MO

CPT	Label	Related	Not Related	Related Costs	Non-Related Costs
J1745	Injection infliximab, 10 mg	8	649	$23,942	$2,361,808
98941	Chiropractic manipulative treatment (CMT); spinal, three to fo	333,199	50,320	$13,061,206	$1,976,146
90471	Immunization administration (includes percutaneous, intrader	1,094	14,442	$60,670	$1,247,632
98940	Chiropractic manipulative treatment (CMT); spinal, one to two	186,058	29,454	$5,898,197	$933,775
A0427	Ambulance service, advanced life support, emergency transpo	326	1,613	$171,980	$859,007
J1567	Injection, immune globulin, intravenous, non-lyophilized (e.g.	0	149	$0	$460,619
A0425	Ground mileage, per statute mile	715	2,289	$84,281	$354,822
J1566	Injection, immune globulin, intravenous, lyophilized (e.g. pow	0	98	$0	$344,535
98943	Chiropractic manipulative treatment (CMT); extraspinal, one o	38,484	12,087	$1,051,524	$328,606
98942	Chiropractic manipulative treatment (CMT); spinal, five region	40,154	6,848	$1,911,296	$327,064
J3490	Unclassified drugs	320	1,367	$68,880	$316,574
J0696	Injection, ceftriaxone sodium, per 250 mg	235	3,660	$14,642	$247,941
A0429	Ambulance service, basic life support, emergency transport (b	414	595	$164,109	$239,286
J9310	Rituximab, 100 mg	0	28	$0	$184,308
90658	Influenza virus vaccine, split virus, when administered to indiv	763	11,645	$11,735	$179,962
J0256	Injection, alpha 1 - proteinase inhibitor - human, 10 mg	0	38	$0	$179,839
A0999	Unlisted ambulance service	54	521	$10,505	$169,628
J0585	Botulinum toxin type a, per unit	4	200	$2,736	$155,619
J7188	Injection, von willebrand factor complex, human, iu	0	65	$0	$148,936
A0431	Ambulance service, conventional air services, transport, one w	1	22	$6,479	$139,565

Related Inpatient Admissions, LBP Unspecified Episode – 3MO

- 4% of total episode costs

ICD-9 Diagnosis	N	Amount
7220 -Cervical Disc Displacmnt	101	$538,805
72402-Spinal Stenosis-Lumbar	36	$308,261
7242 -Lumbago	45	$186,588
7210 -Cervical Spondylosis	30	$171,124
28262-Hb-Ss Disease w Crisis	6	$134,735
7245 -Backache NOS	30	$121,130
7224 -Cervical Disc Degen	18	$108,336
7211 -Cerv Spondyl w Myelopath	14	$106,999
7384 -Acq Spondylolisthesis	10	$101,495
0380 -Streptococcal Septicemia	1	$90,020
5849 -Acute Renal Failure NOS	7	$84,575
5770 -Acute Pancreatitis	8	$83,142
78659-Chest Pain NEC	19	$82,443
41071-Subendo Infarct, Initial	7	$81,957
8472 -Sprain Lumbar Region	15	$73,291
Top 10	291	$1,867,493
Grand Total	888	$6,137,169

DRG	DRGlabel	N	Amount
552	Medical back problems w/o MCC	77	$326,360
460	Spinal fusion except cervical w/o MCC	65	$321,266
243	Permanent cardiac pacemaker implant w C	56	$260,891
473	Cervical spinal fusion w/o CC/MCC	54	$241,887
392	Esophagitis, gastroent & misc digest disord	22	$183,502
498	Local excision & removal int fix devices of h	16	$160,458
520	(blank)	22	$154,321
491	Back & neck proc exc spinal fusion w/o CC/	16	$148,421
103	Headaches w/o MCC	7	$135,220
497	Local excision & removal int fix devices exc	8	$107,638
578	Skin graft &/or debrid exc for skin ulcer or c	2	$100,904
781	Other antepartum diagnoses w medical con	8	$100,358
812	Red blood cell disorders w/o MCC	12	$92,481
500	Soft tissue procedures w MCC	9	$77,022
811	Red blood cell disorders w MCC	1	$74,760
Top 10		343	$2,039,964
Grand Total		888	$6,137,169

Non-related Inpatient Admissions, LBP Unspecified Episode – 3MO

ICD-9 Diagnosis	N	Amount
41401-Crnry Athrscl Natve Vssl	340	$2,903,536
71536-Loc Osteoarth NOS-L/Leg	171	$1,595,691
486 -Pneumonia, Organism NOS	164	$1,548,562
65421-Prev C-Delivery-Delivrd	232	$1,322,376
56211-Dvrtcli Colon wo Hmrhg	131	$1,192,349
5770 -Acute Pancreatitis	97	$929,927
78659-Chest Pain NEC	228	$929,199
66411-Del w 2 Deg Lacerat-Del	229	$853,684
V5789-Rehabilitation Proc NEC	46	$840,797
27801-Morbid Obesity	157	$791,234
65971-Abn Ftl Hrt Rate/Rhy-Del	131	$706,785
41071-Subendo Infarct, Initial	87	$697,998
71535-Loc Osteoarth NOS-Pelvis	80	$671,718
71596-Osteoarthros NOS-L/Leg	68	$609,263
5921 -Calculus of Ureter	139	$572,681
Top 10	1,795	$12,907,355
Grand Total	9,955	$67,064,578

DRG	DRGlabel	N	Amount
775	Vaginal delivery w/o complicating diagnoses	685	$2,560,957
470	Major joint replacement or reattachment of l	274	$2,524,980
766	Cesarean section w/o CC/MCC	290	$2,215,898
392	Esophagitis, gastroent & misc digest disord	237	$2,123,288
373	Major gastrointestinal disorders & peritonea	415	$1,415,600
765	Cesarean section w CC/MCC	146	$1,350,505
359	Uterine & adnexa proc for non-malignancy v	183	$1,015,854
371	Major gastrointestinal disorders & peritonea	194	$926,146
544	Pathological fractures & musculoskelet & co	120	$904,446
885	Psychoses	182	$883,394
743	Uterine & adnexa proc for non-malignancy v	339	$753,469
313	Chest pain	183	$681,328
249	Perc cardiovasc proc w non-drug-eluting ste	110	$625,147
103	Headaches w/o MCC	45	$606,158
288	Acute & subacute endocarditis w MCC	58	$571,935
Top 10		2,726	$15,921,068
Grand Total		9,955	$67,064,578

Related Drug Costs, LBP Unspecified Episode – 3MO

- Notes: Drugs compose 13% of total episode costs

Therapeutic Class	N	Amount	% of N	% of Amount
060-Anal/Antipyr, Opiate Agonists	219,964	$8,989,603	36.0%	35.8%
059-Analg/Antipyr, Nonsteroid/Antiinflam	140,374	$5,357,770	23.0%	21.3%
029-Muscle Relax, Skeletal Central	142,013	$4,741,180	23.2%	18.9%
069-Psychother, Antidepressants	12,708	$1,816,753	2.1%	7.2%
068-Anticonvulsants, Misc	9,833	$1,423,793	1.6%	5.7%
062-Analgesics/Antipyretics, NEC	34,089	$1,212,979	5.6%	4.8%
077-CNS Agents, Misc.	5,059	$738,103	0.8%	2.9%
166-Adrenals & Comb, NEC	38,294	$301,108	6.3%	1.2%
030-Muscle Relax, Skeletal, Misc	4,148	$223,487	0.7%	0.9%
061-Anal/Antipyr, Opiate Part Agonist	1,373	$185,152	0.2%	0.7%
058-Analg/Antipyr, Salicylates	2,514	$92,398	0.4%	0.4%
138-Antiinflam Agents EENT, NEC	754	$13,941	0.1%	0.1%
137-Antiinfect, Antiinflam EENT	156	$6,898	0.0%	0.0%
999-Other/unavailable	26	$3,393	0.0%	0.0%
133-Antiinfect, Antibiotics, EENT	1	$92	0.0%	0.0%
Grand Total	611,306	$25,106,650	100.0%	100.0%

Common Non-related Drug Costs, LBP Unspecified Episode – 3MO

Therapeutic Class	N	Amount	% of N	% of Amount
053-Antihyperlipidemic Drugs, NEC	125,361	$16,376,472	7.3%	11.5%
162-Gastrointestinal Drugs Misc, NEC	83,259	$15,796,086	4.9%	11.1%
069-Psychother, Antidepressants	147,478	$12,576,033	8.6%	8.8%
234-Unclassified Agents, NEC	57,887	$9,927,611	3.4%	6.9%
174-Antidiabetic Agents, Misc	43,407	$4,989,290	2.5%	3.5%
068-Anticonvulsants, Misc	17,090	$4,083,993	1.0%	2.9%
046-Cardiac Drugs. NEC	48,421	$3,996,910	2.8%	2.8%
001-Antihistamines & Comb, NEC	54,870	$3,810,934	3.2%	2.7%
075-Anxiolytic/Sedative/Hypnotic NEC	43,874	$3,386,318	2.6%	2.4%
052-Cardiac, Calcium Channel	41,799	$3,057,932	2.4%	2.1%
166-Adrenals & Comb, NEC	19,129	$2,979,060	1.1%	2.1%
047-Cardiac, ACE Inhibitors	66,661	$2,686,892	3.9%	1.9%
032-Vascular 5HT1 Agonist, NEC	13,030	$2,675,733	0.8%	1.9%
170-Estrogens & Comb, NEC	47,723	$2,603,375	2.8%	1.8%
138-Antiinflam Agents EENT, NEC	28,246	$2,394,531	1.7%	1.7%
051-Cardiac, Beta Blockers	65,241	$2,304,788	3.8%	1.6%
168-Contraceptive, Oral Comb, NEC	51,675	$2,303,314	3.0%	1.6%
014-Antivirals, NEC	14,451	$2,245,131	0.8%	1.6%
070-Psychother, Tranq/Antipsychotics	7,661	$2,199,629	0.4%	1.5%
022-Interferons, NEC	1,058	$1,976,892	0.1%	1.4%

LBP Unspecified Provider Attribution

- Identify the provider or providers “responsible” for the patient’s care during the course of an episode
- Support a comparison across providers rather than simply across all episodes, which may be reflective of a normal distribution of costs population-wide

Proposed Attribution Model – Unspecified Pneumonia Episode

- Episodes will be attributed to the provider who billed for the visit acting as the episode's trigger or index event
  - These providers will be identified using the Provider ID number, where valid
- Providers of any specialty may be attributed the episode (though we recommend that comparisons of resource use between providers be made only within a single specialty)

Identifying Variability in LBP-specific Resource Use

- Analyses intended to identify trends in the observed variability in resource use for episodes of LBP management
- Variability measured at the following levels:
  - Region
  - State
  - Specialty

Unspecified LBP (3 mos.): Mean Resource Use by Type of Service, All Episodes

Description	Mean	% of Total	5th %	25th %	50th %	75th %	95th %
Evaluation and Management - OP	$81	21.2%	$0	$0	$65	$107	$227
Chiropractic	$63	16.5%	$0	$0	$0	$67	$327
Physical Therapy	$62	16.1%	$0	$0	$0	$27	$357
Drug Costs	$61	15.8%	$0	$0	$0	$38	$268
Procedures	$43	11.3%	$0	$0	$0	$0	$214
Outpatient Facility Costs	$28	7.3%	$0	$0	$0	$0	$0
Imaging	$28	7.2%	$0	$0	$0	$0	$98
Inpatient Facility Costs	$9	2.2%	$0	$0	$0	$0	$0
Tests	$4	0.9%	$0	$0	$0	$0	$6
Other Services	$3	0.7%	$0	$0	$0	$0	$9
Durable Medical Equipment	$2	0.6%	$0	$0	$0	$0	$0
Unclassified	$1	0.1%	$0	$0	$0	$0	$0
Evaluation and Management - IP	$0	0.0%	$0	$0	$0	$0	$0
Total Costs	**$383**	**100.0%**	**$50**	**$96**	**$178**	**$389**	**$1,348**

n = 517,177

Unspecified LBP (3 mos.): Resource Use by Type of Service vs. Overall Mean, by Region

Description	Mean	South	North Central	West	Northeast
N	517,177	228,026	135,143	102,224	43,167
Evaluation and Management - OP	$81	1.09	0.88	1.00	1.02
Chiropractic	$63	0.93	1.17	0.90	1.23
Physical Therapy	$62	1.01	0.86	1.13	1.21
Drug Costs	$61	1.09	0.92	0.95	0.80
Procedures	$43	0.97	0.90	1.17	1.20
Outpatient Facility Costs	$28	1.15	1.01	0.60	1.21
Imaging	$28	1.24	0.90	0.72	0.84
Inpatient Facility Costs	$9	0.99	1.13	0.86	0.42
Tests	$4	1.41	0.70	0.57	0.94
Other Services	$3	1.21	0.77	0.98	0.83
Durable Medical Equipment	$2	1.19	0.98	0.82	0.63
Unclassified	$1	1.27	0.91	0.76	0.55
Evaluation and Management - IP	$0	0.00	1.28	3.37	0.00
Total Costs	**$383**	**1.06**	**0.95**	**0.96**	**1.05**

Unspecified LBP (3 mos.): Resource Use by Type of Service vs. Overall Mean, by State

Description	Mean	TX	CA	GA	MI	TN	OH	SC	IL	FL	MO
N	517,177	53,914	50,589	32,555	27,562	24,969	21,195	20,108	17,260	16,650	14,483
E&M - OP	$81	1.19	1.12	1.07	1.01	0.98	1.03	1.10	0.87	1.29	0.86
Chiropractic	$63	0.78	0.67	0.93	1.11	1.42	0.84	0.73	1.35	0.71	1.14
PT	$62	1.07	1.22	0.92	0.93	1.15	0.81	0.72	1.35	1.12	0.75
Drug Costs	$61	0.92	1.01	0.82	1.24	1.33	1.04	1.13	0.62	1.41	1.02
Procedures	$43	1.04	1.16	0.98	0.96	0.97	0.96	0.94	1.28	1.33	0.91
OP Facility	$28	1.80	0.46	0.58	0.90	0.82	1.34	0.67	1.04	1.29	1.20
Imaging	$28	1.13	0.62	0.98	0.98	1.48	0.77	1.15	1.04	1.90	1.27
IP Facility	$9	0.97	0.62	0.87	1.35	0.64	1.45	1.30	1.27	1.24	0.60
Tests	$4	1.76	0.53	1.50	0.88	1.58	0.65	1.53	1.05	1.50	0.69
Other Services	$3	1.37	1.07	1.30	0.68	1.12	0.82	1.40	0.74	1.33	0.91
DME	$2	1.12	0.57	1.27	0.95	1.15	1.33	0.36	0.97	1.44	0.91
Unclassified	$1	1.44	0.54	2.87	0.73	0.42	0.45	0.96	0.92	1.30	1.20
E&M - IP	$0	0.00	6.80	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Total Costs	**$383**	**1.09**	**0.94**	**0.93**	**1.04**	**1.16**	**0.97**	**0.94**	**1.07**	**1.24**	**0.97**

Unspecified LBP (3 mos.): Resource Use by Type of Service vs. Overall Mean, by Specialty

- Results presented for high-volume specialties: Top 1-5

Description	Mean	Chiropractor	Family Practice	Internal Medicine	Medical Doctor NEC	Multi-Specialty Group
N	517,177	176,489	136,655	61,902	33,206	14,129
E&M - OP	$81	0.50	1.23	1.26	1.12	1.21
Chiropractic	$63	2.60	0.09	0.07	0.58	0.15
PT	$62	1.58	0.34	0.37	0.68	0.57
Drug Costs	$61	0.30	1.34	1.38	1.06	1.27
Procedures	$43	0.78	0.63	0.65	0.80	1.01
OP Facility	$28	0.17	1.03	1.15	1.36	0.47
Imaging	$28	0.55	1.07	1.23	0.87	1.06
IP Facility	$9	0.27	0.92	1.07	1.14	1.14
Tests	$4	0.16	1.37	1.86	0.87	0.85
Other Services	$3	0.20	0.96	0.84	1.18	2.35
DME	$2	1.19	0.62	0.55	0.75	0.41
Unclassified	$1	0.82	1.10	0.73	0.99	0.78
E&M - IP	$0	0.00	0.00	0.00	0.00	0.00
Total Costs	**$383**	**0.99**	**0.81**	**0.85**	**0.91**	**0.85**

Unspecified LBP (3 mos.): Resource Use by Type of Service vs. Overall Mean, by Specialty

- Results presented for high-volume specialties: 6-10

Description	Mean	Emergency Medicine	Physical Therapy	Orthopaedic Surgery	Physical Rehab	Mental Health
N	517,177	14,100	12,376	11,866	6,848	3,737
E&M - OP	$81	2.21	0.39	1.63	1.69	1.39
Chiropractic	$63	0.10	0.10	0.05	0.08	0.08
PT	$62	0.27	6.12	0.95	1.86	0.19
Drug Costs	$61	0.93	0.90	1.14	3.33	0.91
Procedures	$43	0.47	6.66	2.33	3.17	0.46
OP Facility	$28	7.30	0.55	1.84	1.39	0.66
Imaging	$28	1.18	0.82	3.94	2.23	0.77
IP Facility	$9	4.01	1.04	3.42	0.70	1.01
Tests	$4	0.97	0.54	0.74	3.45	0.81
Other Services	$3	7.18	0.36	0.97	1.68	1.09
DME	$2	0.51	2.93	2.78	2.79	0.53
Unclassified	$1	1.61	1.52	1.20	1.44	2.36
E&M - IP	$0	36.68	0.00	0.00	0.00	0.00
Total Costs	$383	1.50	2.13	1.48	1.90	0.68

Risk Adjustment

- Testing of risk adjustment models

- Apply risk adjusted results to produce a provider specific summary report.

Risk Adjustment Model Specification

- Test 12 different model specifications
  - Logged GLM model using gamma distribution
    - Full list of recommended comorbidities (> 1% prevalence)
    - Only recommended comorbidities that are statistically significant
    - Only recommended comorbidities that are statistically significant + additional comorbidities flagged for “empirical analysis” (all, significant only)
    - All HCCs & all statistically significant HCCs (regardless of prevalence)
  - Normal GLM model (estimates in dollars)
    - Same tweaks as above
- Fit models for the entire cohort, then for each of the age strata separately (total of 48 risk adjustment models)

LBP Episode Risk Adjustment Matrix

Model #	Independent Variables						Distribution	Link function
	WG Specified (> 1%)	WG specified (> 1%) p < 0.1	Test conditions (> 1%)	Test conditions (> 1%) p < 0.1	All HCCs	All HCCs p < 0.1		
1	X						Gamma	Log
2		X					Gamma	Log
3		X	X				Gamma	Log
4		X		X			Gamma	Log
5	X						Normal	Identity
6		X					Normal	Identity
7		X	X				Normal	Identity
8		X		X			Normal	Identity
9					X		Gamma	Log
10						X	Gamma	Log
11					X		Normal	Identity
12						X	Normal	Identity

Example Episode Report

LBP Unspecified Episode

Report for Physician #18310983

Provider type = Chiropractor

	MD	Peer Group	Non-Peer Group	National Avg
Episodes	16	44,545	118,910	163,471
Observed Costs*				
Average	$ 265	$ 762	$ 1227	$ 1100
Min	$ 58	$ 58	$ 58	$ 58
Median	$ 212	$ 411	$ 590	$ 514
Max	$ 687	$ 8289	$ 8289	$ 8289
Predicted Costs				
Average	$ 1061	$ 1082	$ 1101	$ 1096
Min	$ 994	$ 446	$ 473	$ 446
Median	$ 1062	$ 1074	$ 1083	$ 1082
Max	$ 1136	$ 2220	$ 2619	$ 2619
Observed-to-Expected Ratio				
Average	0.25	0.71	1.12	1.00
Min	0.05	0.03	0.03	0.03
Median	0.19	0.38	0.54	0.47
Max	0.63	8.63	11.16	11.16
% ≥ 2.0	0.0%	7.5%	16.0%	13.7%
% ≥ 2.5	0.0%	4.6%	11.6%	9.7%
% ≥ 75th percentile peers	0.0%	(0.0%, 20.6%)		

* Observed costs adjusted for outliers (windsorized)

Notes:
- Use Model 12