

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 #: 1611 NQF Project: Endorsing Resource Use Standards- Phase II

BRIEF MEASURE INFORMATION

Measure Title: ETG Based PNEUMONIA cost of care measure

Measure Steward (IP Owner): Ingenix, 950 Winter Street, Suite 3800, Waltham, Massachusetts, 02154

Brief description of measure: The measure focuses on resources used to deliver episodes of care for patients with pneumonia. Pneumonia episodes are defined using the Episode Treatment Groups (ETG) methodology and describe the unique presence of the condition for a patient and the services involved in diagnosing, managing and treating pneumonia. A number of resource use measures are defined for pneumonia episodes, including overall cost of care, cost of care by type of service, and the utilization of specific types of services. Each resource use measure is expressed as a cost or a utilization count per episode and comparisons with internal and external benchmarks are made using risk adjustment to support valid comparisons.

As requested by NQF, the focus of this submission is for pneumonia episodes and will cover both measures at the pneumonia base and severity level and also a pneumonia composite measure where pneumonia episode results are combined across pneumonia severity levels. At the most detailed level, the measure is defined as the base condition of pneumonia and an assigned level of severity (e.g., resources per episode for pneumonia, severity level 1 episodes). Composite measures can then be created using these measurement units to meet a specific need. For example, a composite measure for pneumonia is derived by combining pneumonia episode results across pneumonia severity levels. Appropriate risk adjustment is applied to support comparisons (e.g., for physician measurement, adjusting for a physician's mix of pneumonia episodes by severity level when supporting a pneumonia composite comparison).

The focus of this measure is on pneumonia. However, pneumonia episode results could also be included in a "pulmonary" or other clinical composite for a physician, combining episodes in clinical areas similar to pneumonia. Further, an "overall" composite for a physician can be created, again by aggregating episode results across appropriate conditions and severity levels and applying proper risk adjustment when making comparisons.

Resource use service categories: Inpatient services: Inpatient facility 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

Brief description of measure clinical logic: This measure identifies patients with pneumonia and creates pneumonia episodes of care using the ETG methodology described in the ETG Construction Logic attached in our response to S.2. Each episode of pneumonia is characterized by an ETG Base class ID that specifies the type of condition; the ETG Base class ID representing pneumonia is 437400.

An episode of pneumonia will contain all clinically relevant information related to the condition. In addition to this information, certain diagnoses are considered co-morbidities or condition status factors for pneumonia. For example, for pneumonia, Streptococcal pneumonia is a condition status factor and Septicemia is a comorbidity.

Each episode is assigned a severity level based on age, gender and the observed comorbidity and condition status factors. The severity level is an indicator of the relative resources expected to be required for the given episode of pneumonia.

The pneumonia episode clinical framework is defined by the services, or claim lines, that can begin an episode, the primary and incidental diagnosis relationships involved, how records group to an episode, including relative strength of relationship, and the severity logic employed.

If included in a composite or paired with another measure, please identify composite or paired measure:

Subject/ Topic Areas: Pulmonary/Critical Care : Pneumonia
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
<p>A. Measure Steward Agreement. <i>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.</i></p> <p>A.1. Do you attest that the measure steward holds intellectual property rights to the measure? (If no, do not submit)</p> <p>Yes</p> <p>A.2. Please check if either of the following apply:</p> <p>Proprietary measure</p> <p>A.3. Measure Steward Agreement.</p> <p>Agreement signed and submitted</p> <p>A.4. Measure Steward Agreement attached:</p> <p>NQF Resource Use Addendum FINAL-634371615693359178.pdf</p>	<p>A</p> <p>Y <input type="checkbox"/> N <input type="checkbox"/></p>
<p>B. Maintenance. <i>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)</i></p> <p>Yes, information provided in contact section</p>	<p>B</p> <p>Y <input type="checkbox"/> N <input type="checkbox"/></p>
<p>C. Purpose/ Use (All the purposes and/or uses for which the measure is specified and tested:</p> <p>Payment Program Public Reporting Quality Improvement (Internal to the specific organization) Quality Improvement with Benchmarking (external benchmarking to multiple organizations)</p>	<p>C</p> <p>Y <input type="checkbox"/> N <input type="checkbox"/></p>
<p>D. Testing. <i>The measure is fully specified and tested for reliability <u>and</u> validity (See guidance on measure testing).</i></p> <p>Yes, reliability and validity testing completed</p>	<p>D</p> <p>Y <input type="checkbox"/> N <input type="checkbox"/></p>
<p>E. Harmonization and Competing Measures. <i>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)</i></p> <p>Yes</p>	<p>E</p> <p>Y <input type="checkbox"/> N <input type="checkbox"/></p>

<p>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)</p> <p>Yes</p> <p><i>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? Yes</i></p>	
<p>F. Submission Complete. <i>The requested measure submission information is complete and responsive to the questions so that all the information needed to evaluate all criteria is provided.</i></p>	<p>F</p> <p>Y <input type="checkbox"/> N <input type="checkbox"/></p>
<p>Have all conditions for consideration been met? Staff Notes to Steward (if submission returned):</p>	<p>Y <input type="checkbox"/> N <input type="checkbox"/></p>
<p>Staff Notes to Reviewers (issues or questions regarding any criteria):</p>	
<p>File Attachments Related to Measure/Criteria: Attachment: ETG Construction Logic Pneumonia.doc Attachment: S5_Pneumonia_DataDictionary.xls Attachment: S5_Pneumonia_DataDictionary-634413047004559687.xls Attachment: S6_DataProtocol-634413048365653437.xls Attachment: S7.2_Data Source Reference-634413052967372187.xls Attachment: S8_Pneumonia_ClinicalLogic.xls Attachment: ETG Construction Logic Pneumonia-634413056088309687.doc Attachment: S9.7_RU_Categories-634413225228606337.xls Attachment: S10_Risk Adjustment Method Example-634413225817356337.xls S12_sample_score_report_EPI-634413291076418837.pdf Attachment: SA_Reliability_VValidity Testing_Pneumonia.xls</p>	

IMPORTANCE TO MEASURE AND REPORT	
<p>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.</p> <p>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.</p>	
<p>High Impact</p> <p>IM1. Demonstrated high impact aspect of healthcare:</p> <p>Affects large numbers A leading cause of morbidity/mortality High resource use</p> <p>IM1.1. Summary of evidence of high impact:</p> <p>Pneumonia is an inflammation of the lung most often caused by infection with bacteria, viruses, and other organisms, although there are also non-infectious causes. Approximately 50 percent of pneumonia cases are believed to be caused by viruses and tend to result in less severe illness than bacteria-caused pneumonia. Most pneumonia in the very young is caused by viral infection. Streptococcus pneumoniae or pneumococcal pneumonia is the most common cause of bacterial pneumonia acquired outside of hospitals. It is estimated that 175,000 cases of pneumococcal pneumonia occur each year, with a fatality rate of 5-7%, or even much higher among the elderly.⁵ Mycoplasmas are responsible for</p>	<p>Eval Rating</p> <p>1a</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/></p>

approximately 15-50 percent of all adult cases of pneumonia and an even higher rate in school-aged children⁶. Pneumonia is often a complication of a pre-existing condition/infection and is triggered when a patient's defense system is weakened, most often by a simple viral respiratory tract infection or a case of influenza, especially in the elderly. There are no generally effective treatments for most types of viral pneumonia, which usually heal on their own. Early treatment with antibiotics can cure bacterial pneumonia and speed recovery from mycoplasma pneumonia. However, the disease has become more resistant to these drugs, making treatment of pneumococcal infections more difficult. Together, pneumonia and influenza represented a cost to the U.S. economy in 2005 of \$40.2 billion, \$6 billion due to indirect mortality I costs and \$34.2 billion in direct II costs.¹

Pneumonia and influenza together are ranked as the eighth leading cause of death in the United States.² Pneumonia consistently accounts for the overwhelming majority of deaths between the two. In 2006, 55,477 people died of pneumonia.³

There were an estimated 589,000 hospital discharges in males (40.2 per 10,000) and 643,000 discharges in females (42.4 per 10,000) attributable to pneumonia in 2006. The highest pneumonia discharge rate that year was seen in those 65 and over at 189.0 per 10,000.⁴

The pneumococcal polysaccharide vaccine (PPSV) is recommended for anyone over 65 years of age and all those over 2 years of age who have asthma, long-term health problems, lowered infection resistance (from disease, condition, or treatment), are a candidate for or recipient of a cochlear implant, or are a smoker.¹⁰

Overall in 2008, 67 percent of those 65 years of age or older received the pneumonia vaccine (median for the fifty states and DC).⁷ In 2008, only 53.6 percent of African Americans 65 years of age or older received the pneumonia vaccine, compared to 69.5 percent of whites in the same age group (median for the fifty states and DC).⁸ This is despite it being covered by Medicare and available to seniors.⁹

Analyses of Ingenix healthcare benchmark data for a large population of individuals can support an understanding of the importance of pneumonia and the measurement of resource use. Using a 12-month sample population of more than 7 million individuals (primarily non-elderly) from 9 health care organizations, patients with pneumonia were identified using diagnosis codes assigned to medical administrative claim records. The percentage of costs for these patients related to pneumonia and other conditions was also estimated using ETG grouped data for the identified pneumonia patients. Using this benchmark data, 0.4% of the total population was identified as having pneumonia. Total cost per member per month for these individuals was \$4,905. Approximately 14% of the total costs for the members identified with pneumonia were identified as being related to pneumonia (based on total costs grouped to those condition episodes for those patients).

Analyses of the Ingenix healthcare benchmark data described above for episodes attributed to internal medicine physicians can further support an understanding of the relative financial importance of resource use measures for the condition. As shown below, across all physician episodes, the average total cost per episode is more than \$4,000. Hospital Services comprise the largest component of costs for these episodes.

Pneumonia

of Episodes 12,448

Cost per Episode:

Total Cost per Episode	\$4,054
Primary Care Core Cost per Episode	\$161
Specialty Care Cost per Episode	\$677
ER Cost per Episode	\$160
Radiology Cost per Episode	\$116
Pharmacy Cost per Episode	\$152
Laboratory Cost per Episode	\$44
Hospital Services Cost per Episode	\$2,745

Utilization per 1,000 Episodes:

Specialist Visits per 1000 Episodes	3,586
Radiology Encounters per 1000 Episodes	1,781
Laboratory Encounters per 1000 Episodes	356
ER Visits per 1000 Episodes	169
Admission Days per 1000 Episodes	1,357
Number of Admissions per 1000 Episodes	215
Number of Prescriptions per 1000 Episodes	1,942

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

- 1 Centers for Disease Control. MMWR Prevention and Control of Influenza: Recommendations of the Advisory Committee on Immunization Practices (ACIP), 2007. Vol. 56: June 2007.
- 2 National Center for Health Statistics. National Vital Statistics Report. Deaths: Final Data for 2006. Vol. 57, 14 April 2009.
- 3 Ibid.
- 4 National Center for Health Statistics. National Hospital Discharge Survey, 1988 2006.
- 5 Immunization Action Committee. Ask the Experts. Pneumococcal polysaccharide vaccine (PPSV). July 2009. Available at http://www.immunize.org/askexperts/experts_ppv.asp. Accessed on August 19, 2009.
- 6 US National Library of Medicine and US National Institutes of Health. Medline Plus. Medical Encyclopedia: Mycoplasma pneumonia. July 2007. Available at <http://www.nlm.nih.gov/medlineplus/ency/article/000082.htm>. Accessed on August 12, 2009.
- 7 Centers for Disease Control and Prevention. Behavioral Risk Factor Surveillance System. Prevalence and Trends Data, 2008. Available at <http://apps.nccd.cdc.gov/brfss/display.asp?cat=IM&yr=2008&qkey=4408&state=UB>. Accessed on August 19, 2009.
- 8 Ibid
- 9 Centers for Disease Control and Prevention. Public health and aging: Influenza vaccination coverage among adults aged >50 years and pneumococcal vaccination coverage among adults aged >65 years – United States, 2003. Morbidity and Mortality Weekly Report. October 2003;52(41): 987-992.
- 10 Centers for Disease Control and Prevention. Vaccines and Preventable Diseases: Pneumococcal Disease in Short. January 2009. Available at <http://www.cdc.gov/vaccines/vpd-vac/pneumo/in-short-both.htm>. Accessed on August 12, 2009.
- 12 American Lung Association. pneumonia Fact Sheet: national estimates and general information on pneumonia in the United States, 2011 [Internet]. Available at <http://www.lungusa.org/lung-disease/influenza/in-depth-resources/pneumonia-fact-sheet.html>. Accessed on February 1, 2011.

IM2. Opportunity for Improvement**IM2.1. Briefly explain the benefits envisioned by use of this measure:**

Benefits envisioned by this set of measures relates to identifying opportunities and measuring value. In particular, the measure and its components can support:

- The understanding of opportunities to improve the efficiency of healthcare, in particular for patients with selected conditions. Reducing unwarranted variation will provide an opportunity to decrease resources expended without a significant impact on quality of care and outcomes. In some cases, outcomes may improve due to the decrease in the provision of unnecessary services and
- Measurement of the value delivered by individual providers, provider groups, and delivery systems – in particular the resources expended in care delivery. A number of current initiatives require a valid and robust approach to resource measurement, including medical homes, value-based payment and accountable care organizations (ACOs). The ETG episode methodology described in this submission provides a solid foundation to support such measurements. The resource cost and use measures included in this submission provide actionable insights into relative performance and opportunities for improvement.

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

The variation in resource use across providers can be demonstrated using actual measures of physician performance for the condition episodes.

Data to explore this question were extracted from the Ingenix National health care services benchmark database. This database describes enrollment, medical and pharmacy services, and providers for a population of more than 25 million covered lives. The data used for this analysis was primarily for commercial non-elderly individuals and covered the years 2009 thru 2010. In particular, data for 9 health care organizations including 7 million members were selected. The information was processed to produce pneumonia episodes. Incomplete and low cost outlier episodes were excluded. High cost outlier episodes were truncated at the high outlier threshold level. Episodes were attributed to providers in relevant specialties (peer groups).

The observed and expected costs for pneumonia episodes were computed, with expected costs based on averages for a provider's peers, adjusted to reflect the providers mix of pneumonia episodes by severity level. In particular, the

1b

H ☐
M ☐
L ☐
I ☐

following steps were performed:

- Computed the observed experience for the provider being measured, across all episodes to be included in the comparison;
- Computed the experience for the provider's peers. Compute this experience at the level of the risk adjustment, in this case ETG base condition and severity level. For a peer benchmark, average cost per episode across all peers for the ETG base condition and episode level can be computed.;
- Compared the observed experience to the expected result. This expected result is based on the peers average level of performance, adjusted to reflect the provider's own case mix of episodes by condition and level of severity. The ratio of observed to expected results can be termed the relative cost ratio (O/E ratio) and is a risk adjusted measure. A ratio above 1.00 indicates greater resource use than peers, less than 1.00 lower resource use.

Variation in the O/E ratio across providers was assessed. In this way comparisons or relative resource use can be made, removing differences in the underlying mix of episodes included. Providers with greater than 20 pneumonia episodes were selected. For pneumonia 187 providers and 5,241 episodes were included covering the specialties of internal medicine, family practice and pediatrics. The providers in each specialty were compared with their peers only (same specialty and same enrolled population for the healthcare organization). However, OE results were aggregated across healthcare organizations and specialties to summarize variation.

The observed variation in cost of care performance can be summarized using the inter-quartile range for the O/E ratio (the difference between the 25th and 75th percentile physician OE ratios). The results showed variation in performance across these measure physicians. In particular, the inter-quartile range for the O/E ratio for the following key measures was approximately: (e.g., 0.60 can be interpreted as 40 percent below peers, 1.40 as 40 percent above peers)

- Total Cost per Episode – 0.48 to 1.32
- Specialty Care Cost per Episode – 0.25 to 1.21
- Pharmacy Prescriptions per Episode – 0.80 to 1.15

As shown, the variation observed across providers is significant.

IM2.3. Citations for data on variation:

Variations in per capita spending - Inpatient-based and specialist-oriented pattern of practice

Regional differences in Medicare spending are largely explained by the more inpatient-based and specialist-oriented pattern of practice observed in high-spending regions. Neither quality of care nor access to care appear to be better for Medicare enrollees in higher-spending regions.

Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care. *Ann Intern Med* . 2003 138(4): 273-287. The Dartmouth Atlas shows a more than two-fold variation in per capita Medicare spending in different regions of the country. Adjusting for price differences leads to only a modest decline in overall variations. It is utilization -- the amount of care delivered to patients -- that explains most of the regional variation in Medicare spending. Most spending variation was due to differences in use of the hospital as a site of care (versus, say, hospice, nursing home, or the doctor's office) and to discretionary specialist visits and tests.

Reflections on variations, The Dartmouth Atlas Of Health Care. Available at: <http://www.dartmouthatlas.org/keyissues/issue.aspx?con=1338>. Accessed on February 12, 2011.

Variations in clinical decision making – ambulatory care-sensitive conditions

Clinicians have identified a group of diagnoses referred to as “ambulatory care-sensitive” conditions – such as poorly controlled pneumonia or worsening heart failure – which can be treated in either the inpatient or the outpatient setting, and for which hospitalization can often be prevented by better outpatient management. The variations among regions in admission rates of patients with these conditions can be ascribed to differences in clinical decision-making, rather than to differences in underlying illness rates. Hospitalization rates for these – and for most medical conditions – are also highly correlated with the local supply of hospital beds.

Hospital Discharges for Ambulatory Care-Sensitive Conditions Per 1,000 Medicare Enrollees, By Gender And Type Of Admission, The Dartmouth Atlas Of Health Care (2005) Available at: <http://www.dartmouthatlas.org/data/topic/topic.aspx?cat=20> Accessed on February 12, 2011.

Variations in the use of diagnostic tests and discretionary services

Variations in ECG ordering are not explained by patient characteristics. The tremendous nonclinical variations in ECG test ordering suggest a need for greater consensus about use of screening ECGs in primary care.

Randall SS, Bismruta M. Variation in routine electrocardiogram use in academic primary care practice. *Arch Intern Med.* 2001;161:2351-2355

Physicians in high-spending regions see patients back more frequently and are more likely to recommend screening tests of unproven benefit and discretionary interventions compared with physicians in low-spending regions; however, both appear equally likely to recommend guideline-supported interventions.

Physicians in higher-spending regions were much more likely than those in lower-spending regions to recommend discretionary services, such as referral to a subspecialist for typical gastroesophageal reflux or stable angina or, in another vignette, hospital admission for an 85-year-old patient with an exacerbation of end-stage congestive heart failure. And they were three times as likely to admit the latter patient directly to an intensive care unit and 30% less likely to discuss palliative care with the patient and family. Differences in the propensity to intervene in such gray areas of decision making were highly correlated with regional differences in per capita spending.

Sirovich B, Gallagher PM, Wennberg DE, Fisher ES. Discretionary decision making by primary care physicians and the cost of U.S. health care. *Health Aff (Millwood)*, 2008; 27:813-823

Widely varying levels of health care spending across the United States are strongly correlated with the tendency of local physicians to recommend discretionary interventions. Physicians in regions of differing spending appear to differ only in their discretionary decision making. For decisions that are informed by evidence or practice guidelines (such as screening mammography and standard exercise tolerance testing), physicians were equally likely to recommend interventions regardless of local spending levels

Sirovich B, Gallagher PM, Wennberg DE, Fisher ES. Discretionary Decision Making By Primary Care Physicians And The Cost Of U.S. Health Care. *Health Aff (Millwood)*. 2008; 27(3): 813–823.

Supply sensitive care

Supply-sensitive care accounts for more than half of all Medicare spending. In regions where there are more hospital beds per capita, patients will be more likely to be admitted to the hospital. In regions where there are more intensive care unit beds, more patients will be cared for in the ICU. More specialists will result in more visits to specialists. And the more CT scanners are available, the more CT scans patients will receive. The Dartmouth Atlas has consistently demonstrated these relationships.

Patients do not experience improved survival or better quality of life if they live in regions with more care. In fact, the care they receive appears to be worse. They report being less satisfied with their care than patients in regions that spend less, and having more trouble getting in to see their physicians.

Supply sensitive care, The Dartmouth Atlas Of Health Care (2005) Available at:

<http://www.dartmouthatlas.org/keyissues/issue.aspx?con=2937> Accessed on February 14, 2011.

Numerous studies have found that higher bed supply is associated with more hospital use for conditions where outpatient care is a viable alternative. This includes most medical causes of hospitalization. In 2006, bed supply remained an important determinant of medical discharges.

The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care. *Annals of Internal Medicine.* Feb 18 2003;138(4):273-287.

Fisher ES, Wennberg DE, Stukel TA, Gottlieb DJ, Lucas FL, Pinder EL. The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care. *Annals of Internal Medicine.* Feb 18 2003;138(4):288-298.

By far, the most significant factor associated with how much Medicare spends in any given region is the availability of medical resources. Studies from the Dartmouth Atlas Project have shown that the frequency with which physicians admit patients with chronic diseases to the hospital is highly correlated

with the number of beds per capita in the region. The frequency of visits to medical specialists is correlated with the number of specialists available. And the frequency with which chronically ill patients undergo many diagnostic tests and procedures also varies. We call such procedures and tests, along with the rates of hospitalization and physician visits, “supply-sensitive” care, or care that varies with the local availability of such medical resources as physicians, hospital beds, intensive care unit (ICU) beds, and diagnostic imaging equipment. The volume of supply-sensitive care that is delivered to the chronically ill is a powerful force driving Medicare spending. The utilization of supply-sensitive services for treating the chronically ill varies dramatically across different regions of the country, and it is responsible for much of Medicare spending. Local capacity, or the local supply of medical resources per capita, varies widely, and this local

capacity bears directly on how much care is used to treat the chronically ill.

Wennberg JE, Fisher ES, Goodman DC, Skinner JS. "Tracking the care of patients with severe chronic illness." The Dartmouth Atlas of Health Care 2008. Available at: http://www.dartmouthatlas.org/downloads/atlas/2008_Chronic_Care_Atlas.pdf Accessed on February 14, 2011.

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

Health disparities are defined as differences in the occurrence, frequency, death and burden of diseases and other unfavorable health conditions that exist among specific population groups¹. Examining health care differences or gaps experienced by one population compared to another is an integral part of understanding and improving health care quality². The quality of healthcare delivered within the United States also differs from population to population due to differences in access to care, healthcare utilization and other factors².

Measures of healthcare utilization allow for a broader understanding of access to care². Barriers to care that are associated with differences in healthcare utilization may have a more significant impact on healthcare quality than other factors². Several studies on disparities have relied upon measures of healthcare utilization and the data demonstrates some of the most significant differences in care among diverse groups². Current efforts to improve healthcare delivery continue to rely upon measures of health care utilization to fully understand the complexities surrounding disparate health care outcomes. For example, greater utilization of services does not necessarily indicate better care. In fact, high use of some inpatient services may reflect compromised access to outpatient health services².

In 2006, the Nation's 14 million health service workers provided approximately 960 million office visits, 673 million hospital outpatient visits, treated 37 million hospitalized patients and 1.4 million nursing home residents². Approximately 70% of the non-institutionalized civilian population visited a provider's medical office or outpatient facility and about 60% received a prescription medication². National health expenditures totaled over \$2 trillion dollars in fiscal year 2006 with 5% of the population accounting for 55% of total costs². Additionally, almost one-third of all healthcare expenditures are estimated to be the result of low-quality care, including overuse, misuse and waste². Utilization resource measures provide a mechanism to better understand healthcare delivery patterns in order to improve the health of all population groups.

The cost and use measures included in this submission will provide an approach to assessing disparities. For example, episode-based measures of cost and use can be employed to create severity-adjusted comparisons of the resources expended in treating cardiovascular conditions, including supporting a focus on the condition-related resources.

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

1. Health Disparities in the United States: Facts and Figures, American Society of Clinical Oncology, 2009
2. National Healthcare Disparities Report, U.S. Department of Health & Human Services, Agency for Healthcare Research and Quality, 2008

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

As noted in IM2.1, the intent of the measure and its components is to support:

-- The understanding of opportunities to improve the efficiency of healthcare, in particular for patients with selected conditions. Reducing unwarranted variation will provide an opportunity to decrease resources expended without a significant impact on quality of care and outcomes. In some cases, outcomes may improve due to the decrease in the provision of unnecessary services and

-- Measurement of the value delivered by individual providers, provider groups, and delivery systems – in particular the resources expended in care delivery. A number of current initiatives require a valid and robust approach to resource measurement, including medical homes, value-based payment and accountable care organizations (ACOs). The ETG episode methodology described in this submission provides a solid foundation to support such measurements. The resource cost and use measures included in this submission provide actionable insights into relative performance and opportunities for improvement.

1c

H ☐
M ☐
L ☐
I ☐

IM4. Resource use service categories are consistent with measure construct <i>Refer to IM3.1. & all S9 items to evaluate this criteria.</i>	1d H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/>
TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Importance to Measure and Report</i> ?	
Steering Committee: Was the threshold criterion, <i>Importance to Measure and Report</i> , met? Rationale:	Y <input type="checkbox"/> N <input type="checkbox"/>

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

<p>S1. Measure Web Page: <i>Do you have a web page where current detailed measure specifications can be obtained?</i></p> <p>No</p> <p>S2. General Approach <i>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.</i></p> <p>All of our submitted measures for pneumonia rely on a foundational “episodes of care” concept that uses the Ingenix Episode Treatment Groups (ETG) methodology. Episode-based resource use measurement provides a representation of a patient’s course of treatment for a specific condition. The attached ETG General Methods Construct Logic provides a high level explanation of our ETG concept and a summary of the ETG approach to creating episodes of care for pneumonia.</p> <p>Attachment: ETG Construction Logic Pneumonia.doc</p> <p>S3. Type of resource use measure: Per episode</p> <p>S4. Target Population: Adult/Elderly Care</p> <p>S4.1. Subject/Topic Areas: Pulmonary/Critical Care : Pneumonia</p> <p>S4.2. Cross Cutting Areas (<i>HHS or NPP National health goal/priority</i>) Care Coordination Overuse</p> <p>S5. Data dictionary or code table <i>Please provide a web page URL or attachment if exceeds 2 pages. NQF strongly prefers URLs. Attach</i></p>	<p>Eval Rating 2a1/2b1</p>
--	------------------------------------

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_Pneumonia_DataDictionary.xls

Code Table:

URL:

Please supply the username and password:

Attachment: S5_Pneumonia_DataDictionary-634413047004559687.xls

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:

Please supply the username and password:

Attachment: S6_DataProtocol-634413048365653437.xls

S6.1. Data preparation for analysis

Detail (specify) the data preparation steps and provide rationale for this methodology.

Guidelines : Administrative medical and pharmacy claims, member enrollment and demographic information and provider characteristics describe the primary data sources used in creating ETG pneumonia episodes of care and measures of resource use per episode. The key data elements required to support ETG processing and the creation of resource use per episode measures for pneumonia are detailed in attachment S6_DataProtocol.

General recommendations for preparing data for ETG processing and the creation of resource use sub-measures are as follows:

- The data for all required elements should be complete, valid and consistently populated. In particular:
- Only final claims should be included in processing. Adjustments and pended/non-fully adjudicated claims should be removed;
- All recorded diagnosis, procedure and NDC codes should be included and conform to standard ICD-9, HCPCS, CPT, NUBC revenue code and NDC coding conventions. Any non-standard, or "local" codes should be cross-walked to a valid code;
- An assessment of the relative validity of diagnosis and procedural coding should be made. If significant differences in the prevalence or validity of diagnosis and procedural coding are observed across populations, data sources or administrative claims systems, these discrepancies should be validated and addressed, if relevant. If systematic discrepancies and data issues are the result of incomplete data, the members impacted by the incomplete information should be excluded from processing and measurement. An example is a defined population with significant evidence of missing or invalid coding or a population where primary care capitation is in place and claims or encounters for those

services are not available;

-- Financial fields should be complete and valid, reflecting the actual payment or costs associated with the service or a standard-priced resource cost amount. As a guideline, the financial amount used in resource measurement should reflect all payments for a service, including those made to the provider by payer, patient and other entities. The allowed or equivalent payment is an example;

-- An assessment of the relative validity of the financial information should be made. Systematic gaps in financial data should be validated and if resulting from incomplete data, the members impacted by the incomplete information should be excluded from processing. An example is a defined population with significant evidence of missing or invalid financial data where options are not available to estimate the financial amounts;

-- Inpatient facility claims should accurately represent the admission and discharge dates for the inpatient stay. Interim facility bills where the patient has not been discharged should reflect the time period of the services rendered and captured on the interim bill.

-- The member IDs used to identify a member should be unique – describing an individual member. The member ID field across claims and membership should follow the same format. Duplicate IDs for a member are not recommended;

-- Each member enrollment record should describe a unique enrollment span, that is, the input data includes one row per member for each continuously enrolled period where the member has consistent attributes. A member may have multiple enrollment records reflecting a gap in enrollment or a change to their member attributes (i.e. PCP or Pharmacy Benefit) over time.

-- It is recommended that member enrollment span overlaps are reconciled prior to processing;

-- A member's pharmacy benefit status should be noted and reflects whether or not the member has pharmacy data generally available for use in measurement. Examples of populations where pharmacy data may not be available include the individual not have pharmacy coverage for the defined enrollment period or pharmacy services managed by a pharmacy benefits manager (PBM) and the PBM data has not been integrated with the medical claims;

-- The provider IDs used to identify a provider should be unique – describing an individual physician or other provider. The provider ID field across claims and membership (Assigned PCP) should follow the same format. Duplicate IDs for a provider are not recommended;

-- Each provider ID should be assigned a specialty that reflects the primary specialty of the provider. This information is used to support valid episode grouping and also to assign providers to an appropriate peer group to support episode analysis;

-- A place of service crosswalk table that maps each native place of service code to a standard format is required.

Ingenix valid values include:

-- 11 – Office

-- 12 – Home

-- 21 – Inpatient Hospital

-- 22 – Outpatient Hospital

-- 23 – Emergency Room, Hospital

-- 24 – Ambulatory Surgical Center

-- 31 – Skilled Nursing Facility

-- 39 – Nursing Home, Custodial, Hospice

-- 49 – Ambulance

-- 51 – Inpatient Psychiatric Facility

-- 59 – Psychiatric Facility

-- 61 – Comprehensive Inpatient Facility

-- 69 – Rehab Facility

-- 81 – Independent Lab

-- 99 – Unknown or Other (this POS value should represent a small portion of the data for optimal results)

-- Provider Specialty on claims should accurately reflect the service category of the claim and support assignment of ETG Type of Provider for each claim. Type of Provider values used to support ETG processing include:

- 0 – Clinician
- 1 – Facility
- 2 – Other

- Place of Service, Provider Specialty, CPT/HCPC Procedure Codes and Revenue codes should be accurate and support assignment of ETG Type of Service for each claim. Type of Service values used to support ETG processing include:

- 0 – Ancillary
- 1 – Medical/Surgical
- 2 – Room and Board

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)

Specifications : In creating pneumonia episodes of care, ETG includes all claims for initial processing provided the input format is correct and required fields are provided (refer to section S6.1 for data preparation details and considerations). The ETG methodology does not truncate or eliminate service records based on any cost or other criteria. The identification of financial cost outliers, non-standard diagnosis or procedural coding and other invalid information at the service level is performed by the organization preparing the input data. As noted in S6.1, financial amounts on individual service records should be validated prior to their use in measurement.

In terms of resource use measure construction following ETG grouping, no additional data inclusion criteria are applied. Only pneumonia episodes are included in the measurement of pneumonia episode-based resource use, including the individual services that ETG groups to those episodes. As noted below in section 6.3, it is recommended that incomplete episodes be excluded from resource measurement and outlier episodes be treated appropriately.

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)

Specifications : As described in the submission for S6.2, for the application of ETG episode logic for pneumonia, ETG accepts all claims for initial processing provided the input format is correct and required fields are provided (refer to section S6.1 for data preparation details and considerations). The ETG methodology does not truncate or eliminate service records based on any cost or other criteria. The identification of financial cost outliers, non-standard diagnosis or procedural coding and other invalid information at the service level is performed by the organization preparing the input data. As noted in S6.1, financial amounts on individual service records should be validated prior to their use in measurement.

ETG does include logic to identify high or low cost outliers at the episode level. Although this is not the same as detailed service level data exclusions, inappropriately high individual claims or mispriced claims, in general, will impact the outlier treatment of the pneumonia episodes the claim is grouped to.

In terms of resource use measure construction following ETG grouping, no additional data exclusion criteria are applied. Only pneumonia episodes are included in the measurement of pneumonia episode-based resource use, including the individual services that ETG groups to those episodes. It is recommended that incomplete episodes be excluded from resource measurement and outlier episodes be treated appropriately.

S6.4. Missing Data

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

Specifications : Missing provider specialty assignment will impact the ability to assign record type to a claim line. In addition invalid and incomplete diagnosis and procedure coding, will impact the results of the episode grouping and the measures for pneumonia. For example, inaccurate coding may result in a service record not grouping to a pneumonia episode – due to the miscoding of a pneumonia diagnosis or the procedure code assigned to the service. ETG will attempt to group these services. However, invalid data may prevent this grouping to happen in an appropriate way. In this way, ETG handles data quality issues through the rigor of the logic designed to create appropriate episodes.

In terms of working with missing information during the episode grouping process, ETG uses the following approaches:

-- Missing Diagnosis Codes: If all four diagnosis codes are missing from a non-pharmaceutical claim the ETG application will use the procedure code to group, except when the procedure code requires a valid diagnosis code to be present. This requirement is per the ETG eligibility table. In cases where all diagnosis codes are missing and the procedure requires a valid diagnosis code to also be present, the service record will not group to a pneumonia episode and will be assigned to an error ETG.

-- Missing Procedure Codes: If there is no procedure code on a service record then the record will group based on the diagnosis codes or NDC drug code. If there is no diagnosis, procedure or pharmacy code on the claim, then the claim will not group to a pneumonia episode and will have an error code assigned to it.

--Missing Provider Specialty: If the provider specialty is not available on a service record then the record will be assigned an error ETG code and will not group to a pneumonia episode.

The services not assigned to an episode and noted as “errors” based on missing data are marked with an error ETG number. Services with these ETG numbers would not be included in a pneumonia episode or be used in episode-based resource measurement for pneumonia.

-- Missing Pharmacy Data: For some members and populations, pharmacy data can be missing generally, due to the different factors, including not having a pharmacy benefit with the entity collecting the data used in measurement or pharmacy services being managed by a pharmacy benefits manager (PBM) for the measurement entity. Where pharmacy data are not generally available for a member, adjustments are required to ensure valid comparisons.

The ETG grouping methodology for pneumonia itself does not require pharmacy data. Pharmacy services are treated as ancillary records and can never start an episode for pneumonia. Pharmacy services will join pneumonia episodes. However, missing pharmacy records will impact the observed cost of an episode – which will be underestimated, on average, where pharmacy data are missing. It is recommended that pharmacy benefit/data status be used as a separate category in risk adjusting pharmacy and total costs per episode. For example, the expected or “peer” results for a physician should reflect their mix of members with and without pharmacy benefits/data.

Eval
Rating
2a1

H ☐
M ☐
L ☐
I ☐

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.)

Both medical and pharmacy administrative service records (claims or encounters) are used to support the measures. Member enrollment span, pharmacy benefit status and age and gender are also required. Provider characteristics, including specialty and unique provider identifier also have importance to support episode grouping, attribution and definition of peers.

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:

Please supply the username and password:

Attachment: S7.2_Data Source Reference-634413052967372187.xls

Eval
Rating
2b1

H ☐
M ☐
L ☐
I ☐

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:

Please supply the username and password:

Attachment: S8_Pneumonia_ClinicalLogic.xls

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.

This measure identifies patients with pneumonia and creates pneumonia episodes of care using the ETG methodology described in the ETG Construction Logic attached in our response to S.2. Each episode of pneumonia is characterized by an ETG Base class ID that specifies the type of condition; the ETG Base class ID representing pneumonia is 437400.

An episode of pneumonia will contain all clinically relevant information related to the condition. In addition to this information, certain diagnoses are considered co-morbidities or condition status factors for pneumonia. For example, for pneumonia, Streptococcal pneumonia is a condition status factor and Septicemia is a comorbidity.

Each episode is assigned a severity level based on age, gender and the observed comorbidity and condition status factors. The severity level is an indicator of the relative resources expected to be required for the given episode of pneumonia.

The pneumonia episode clinical framework is defined by the services, or claim lines, that can begin an episode, the primary and incidental diagnosis relationships involved, how records group to an episode, including relative strength of relationship, and the severity logic employed.

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 pneumonia measure's episodes are defined using the Episode Treatment Group (ETG) methodology. Please note that this specification will reference different attachments included with the submission for these measures, including:

- S2_ETG_Construction_Logic_Pneumonia. This attachment provides an overview of ETGs and a summary of the methodology used for pneumonia episodes.
- S5_Peumonia_DataDictionary (Excel workbook attachment). This attachment describes the clinical relationships between diagnosis and procedure codes and the episode condition.
- S8_Peumonia_ClinicalLogic (Excel workbook attachment). This attachment includes Worksheets that describe the details around the components of pneumonia methodologies that relate to co-morbidities, condition status factors, and severity adjustment.

The individual Worksheets in these attachments that relate to the specific components of the methodology are referenced in the following specification.

The pneumonia ETG episode building process that supports pneumonia resource use measures has four important steps:
 Step 1: Identify Records; Assign Record Type and Anchor Records, Classify Diagnoses and Procedures
 Step 2: Build Episodes from Anchor Records
 Step 3: Group Non-Anchor Records to Episodes
 Step 4: Finalize the Episodes (identify co-morbidities and complicating factors, and assign episode severity)

This section (S8.2 Clinical Framework) describes the first three steps in the episode building process. Sections S8.3 and S8.5 describe episode co-morbidities and condition status factors and episode severity.

Step 1- Identify Records; Assign Record Type and Anchor Records, Classify Diagnoses and Procedures
Assign services to record types, identify anchor records and classify diagnoses and procedures on service records to support the creation of pneumonia and other episodes.

Step 1A: Assign Record Type to each Service:

Assign each service to one of the following 5 record types:

- Facility: A claim record submitted by a treatment facility for room & board charges (F)
- Surgery: A claim record submitted by a provider for surgical or related procedure (S)
- Management: A claim record submitted by a provider related to the evaluation of a patient's condition (M)
- Ancillary: A claim record submitted by any provider for laboratory, radiological or similar services (A)
- Pharmaceutical: A claim record for a prescription drug claim (P)

Assign record type based upon servicing provider type and the nature of the service procedure.

- Assign provider type based on the specialty of the service provider. The "ExTypeOfProvider" worksheet of the attachment S5_Pneumonia_DataDictionary includes an example mapping of specialty to provider type. Based upon the specialty of the service provider on the claim record the provider type recognized by ETG is assigned. For example, using the "ExTypeOfProvider" worksheet a provider specialty code of 100 on the claim would be assigned the ETG provider type of Facility.

- Type of service is based on the service procedure code (CPT, HCPCS, Revenue, NDC). The worksheet "ProcToRecordType" in the attachment S5_Pneumonia_DataDictionary includes the information required to assign record type based upon the procedure code on the claim record.

- Use the combination of type of provider and type of service to determine record type. The worksheet "ProcToRecordType" in the attachment S5_Pneumonia_DataDictionary provides a mapping of provider type and type of service to record type. For example, procedure code 99025 (Initial surgical evaluation) is assigned a record type of Management (M) when the provider type is either clinician (see column "Clinician Record Type" where procedureCode=99025) or a facility (see column "Facility Record Type" where procedureCode=99025). This same procedure code would be assigned a record type of Ancillary (A) when the provider type is non-clinician (see column "Non-Clinician Record Type" where procedureCode=99025).

Examples of record type assignment include:

- An office visit record provided by an internist will be assigned a "Clinician" provider type and a record type of "Management (M)"
- A cholecystectomy provided by a general surgeon will be assigned a "Clinician" provider type and a record type of "Surgery (S)"
- A pharmacy prescription will be assigned a record type of "Pharmaceutical (P)"
- An injection for chemotherapy (e.g., HCHPS J-code) will also be assigned a record type of "Pharmaceutical (P)"
- An imaging service provided by a radiologist, orthopedic surgeon, facility or any provider will be assigned a record type of "Ancillary (A)".

The worksheet "ExRecordType" in the attachment S5_Pneumonia_DataDictionary includes further examples.

The assigned record type provides information to the pneumonia episode-building methodology about the nature of the service and whether the diagnostic and other information on the service provides confirmatory information for a clinician service (versus potentially rule-out information from imaging, lab or other diagnostic services). Record type plays an important role in how services can trigger episodes of care and join and/or modify existing episodes.

Step 1B: Identify Anchor Records. The record type assigned in Step 1A is used to identify anchor records. An anchor record indicates that a clinician has evaluated the patient, assigned a diagnosis and has initiated the treatment and care of the patient for the condition. If the record type assigned to the service is M, S, or F (Management, Surgery or Facility), the service is an anchor record. All other services are considered non-anchor records.

Steps 1C through 1F: Before episodes can be built from anchor records and non-anchor services can be assigned to episodes, the relationship of diagnoses and procedures to each condition, including pneumonia, need to be assigned. Steps 1C through 1F describe how these relationships are defined. These initial steps categorize diagnoses and procedures relative to each condition, saving this information for use in the subsequent steps described in Step 2 and Step 3.

Note that in some instances a service may have a potential clinical relationship to more than one condition. This concept has importance to episode building, in general, and for episodes of pneumonia. While each service can inform grouping decisions across multiple episodes, the ETG methodology assigns each service uniquely to a single episode. Such an approach ensures that double-counting does not occur when considering service cost and utilization in the creation of resource use measures. As a result, accurate decisions on assigning a service to an episode of pneumonia or to another condition require the assessment of both the relationship of a service to pneumonia and to all other conditions for a patient. The methodology described in this section classifies diagnoses and procedures based on their relationship to pneumonia and also the strength of that relationship relative to other conditions. Using ETG, accurate episode grouping for pneumonia and other conditions must occur in the context of all of a patient's conditions.

Step 1C: Assign Diagnoses to Diagnosis Class

Assign each ICD-9 diagnosis code to a "diagnosis class". There are three diagnosis classes applied across all diagnosis codes, including diagnosis codes eligible for pneumonia:

- **Specific:** These diagnosis codes indicate a specific disease as opposed to a sign or symptom. These codes are specific enough to be linked to a single ETG. ICD-9 Diagnosis code 481 (pneumococcal pneumonia) is an example of a specific diagnosis code. It is primary to an episode of pneumonia. Specific diagnosis codes are usually primary to and eligible for a single ETG.

- **Non-Specific:** Like specific diagnoses, these diagnosis codes represent a disease or condition, but are not specific enough to support linkage to a single condition. ICD-9 Diagnosis code 079.89 (Other specified viral infection, in conditions classified elsewhere and of unspecified site) is an example of a non-specific ICD-9 code. Although it represents disease as opposed to a sign or symptoms, it is not specific as to representing a single disease. Services with this diagnosis will be assigned to an episode based on both information related to a pneumonia episode as well as information related to other potential conditions.

- **Signs and Symptom:** These diagnosis codes represent signs and symptoms of disease as opposed to a disease or condition. For example, ICD-9 Diagnosis code 780.61 (fever presenting with conditions classified elsewhere) represents a sign and symptom rather than a disease. This code could be related to multiple diseases. ETG assigns sign and symptoms diagnoses to the lowest specificity. Services with signs and symptoms diagnosis codes may be eligible for many ETGs due to their generic nature. These services will be gathered to episodes as a later step in the grouping process, after other, more specific, information has been considered.

Diagnosis class assignments determine how a service is grouped to an episode and the order in which it is considered. The ETG methodology considers one person at a time and an individual's medical and pharmacy service records are grouped in several distinct passes. The methodology first processes the specific and non-specific diagnosis codes on anchor records so that concrete conditions/diseases are created. It then processes services with sign and symptom diagnosis codes in reverse chronological order (based on dates of service) to determine the best episode these services can group to.

Step 1D: Identify the Clinical Relationship Between Diagnosis Codes and Conditions, Including Pneumonia

Match each diagnosis code with one or more conditions (ETGs) through a diagnosis eligibility table. In addition to mapping diagnosis codes to conditions, each diagnosis code is further ranked, based on its strength of association with a condition. A rank of "primary" or "incidental" is assigned to each diagnosis and condition combination, with a further ranking assigned to incidental relationships:

- **Primary:** A "primary" diagnosis/condition relationship is assigned where the diagnosis defines that condition. The diagnosis codes that are classified as primary to pneumonia are listed on the "PrimaryDxCodes" worksheet within the attachment "S5_Pneumonia_DataDictionary" (Note: the word "primary" here is used to describe the relationship between a diagnosis and an episode, it is not used to indicate the position of the diagnosis code on the claim line. The diagnosis in any position on the claim line can have a primary relationship with pneumonia). This map is used to identify primary diagnoses for pneumonia. Examples of diagnoses ranked as primary for pneumonia are 481 (pneumococcal pneumonia), 482 (Other bacterial pneumonia) and 483 (pneumonia due to other specified organism). Primary diagnosis codes can only be ranked as primary for a single ETG condition.

- **Incidental:** These diagnosis codes are eligible for a condition but are not classified as primary. These diagnosis codes can be incidental to other conditions. To support the linkage of these diagnosis codes to a final episode, a further ranking is assigned for each condition based on the relative strength of association between the diagnosis and condition. Values of low, medium, or high are assigned for each diagnosis/condition. The Diagnosis codes that are incidental to pneumonia are listed on the "IncidentalDxCodes" worksheet within the attachment "S5_Pneumonia_DataDictionary". The column "diagnosisEligibilityType" in the worksheet describes the relative strength ranking where 3 represents a high association, 2 represents a medium association and 1 represents a low association. A value of 5 means that the diagnosis code can shift the pneumonia ETG to a different ETG.

Step 1E: Identify Relationships between Procedure Codes and Conditions, Including Pneumonia

Match each procedure code with one or more conditions, including pneumonia, through a procedure eligibility table. All procedure codes that are eligible for pneumonia are listed on the “ProcedureCodes” worksheet within attachment “S5_Pneumonia_DataDictionary”. In the same way diagnoses can relate multiple conditions, a procedure can relate to more than one episode. The ProcedureCodes worksheet also includes a ranking of the strength of the clinical relationship of each CPT and HCPCS code with pneumonia, ranked from 1 to 4 based on the relative strength of the clinical relationship between the procedure and pneumonia. This relationship is included in the “ProcedureRank” column in the worksheet. A rank of 4 represents the strongest association and a rank of 1 the lowest. In this way, ETG considers not only the diagnostic information on a service when making grouping decisions around pneumonia, but also the service procedure and the strength of the relationship between the procedure and pneumonia relative to other potential conditions.

Step 1F: Identify Relationships Between Pharmacy Services and Conditions, Including Pneumonia

The relationship between pharmacy services and pneumonia and other conditions is based on the pharmacy code assigned to the service. To support this assessment, the ETG methodology assigns each pharmacy service to a Drug Category Code (DCC). The DCC describes the drug’s active ingredients and route of administration. DCCs are then mapped to ETGs and define the relationships between a drug and a condition. Most pharmacy services are defined using NDC procedure codes, however selected pharmacy services with a CPT or HCPCS code are also mapped to a DCC by ETG (e.g., J-codes describing injections).

The “Pharmacy” worksheet in the attachment “S5_Pneumonia_DataDictionary” describes the DCCs assigned to pneumonia. Similar to diagnoses and procedures, there are some instances a DCC code may be eligible for more than one ETG. In these cases, the ETG methodology uses strength of the clinical relationship between the DCC code and the episode condition. The “Rank” in the worksheet describes this strength of association for each DCC and pneumonia. The lower the value is for Rank, the stronger the association between the DCC and the episode. If multiple episodes are competing for a pharmacy service, this rank is used to support decisions on assignment.

Given the clinical relationships described in Steps 1A through 1F, the following steps are used to build episodes from anchor records.

Step 2- Build Episodes from Anchor Records.

Building pneumonia episodes from anchor records is a multi-step process that utilizes diagnostic and procedural information and the clinical relationships defined in Step 1. Anchor records are grouped in two passes through the patient’s data. The first pass groups the anchor records with specific and non-specific diagnoses. The second pass groups anchor records with sign and symptoms diagnoses. All anchor records are grouped before all non-anchor records.

Step 2A: Use Anchor Records to Start an Episode of pneumonia Using Specific and Non-Specific Diagnoses

A service must be an anchor record to start an episode of pneumonia. The service must also have a procedure code that is eligible for pneumonia and an ICD-9 diagnosis code that is primary for pneumonia. See worksheets “PrimaryDxCodes” and “ProcedureCodes” within attachment S5_Pneumonia_DataDictionary for a complete list of diagnosis codes and procedure codes that are primary for pneumonia. All codes within the “PrimaryDxCodes” worksheet are considered primary to pneumonia. If an anchor record meeting these requirements is observed, start an episode for pneumonia.

As an example of an anchor record that starts an episode of pneumonia, a pulmonologist sees a patient and submits a claim record using the CPT procedure code 99212 (Office visit, established patient) with and ICD-9 diagnosis code 481 (pneumococcal pneumonia).

Note that a single anchor record can start more than one episode. For example, an anchor record with a diagnosis and procedure code combination that is eligible for pneumonia will start a pneumonia episode. If that record also has a diagnosis and procedure code combination that is eligible for Hypertension, it will also start a Hypertension episode. (See Section I of the Attachment for S2 above for a discussion of the concept of “phantom episode clusters”).

Step 2B: Group Anchor Records to an Episode of pneumonia Using Specific and Non-Specific Diagnoses

Once an episode of pneumonia is started, group further anchor records to that episode. Consider specific and non-specific diagnoses on anchor records first.

First identify whether the anchor record is eligible for pneumonia. Eligible anchor records for pneumonia have a procedure code eligible for pneumonia and a diagnosis code that has either a primary or incidental relationship to pneumonia. See the “ProcedureCodes” worksheet within S5_Pneumonia_DataDictionary for the procedure codes eligible for pneumonia. See the “PrimaryDxCodes” and “IncidentalDxCodes” worksheets within S5_Pneumonia_DataDictionary for a list of the diagnosis codes primary and incidental to pneumonia.

For anchor records with eligibility to a pneumonia episode, apply the following steps to assign the anchor record to an episode.

Step 2B1 - If the anchor record is only eligible for the open pneumonia episode, group the anchor record to the pneumonia episode.

In some cases, an anchor record can be eligible to join more than one episode. This is true because a service may have more than one diagnosis code. Further, diagnosis codes that are incidental for pneumonia may also be eligible for another ETG condition.

Step 2B2 - If the anchor record is eligible for the pneumonia episode and another episode for the patient, apply the following tie-breaking steps to determine the episode an anchor record groups to:

- Assess the specificity of the diagnoses on the anchor record. Diagnosis class describes this specificity and was assigned to each diagnosis code in Step 1C (specific or non-specific).

- Assign the anchor record to an episode based on the diagnosis class. Episodes related to specific diagnoses take precedence over episodes related to non-specific diagnoses.

Specific diagnoses:

- If a diagnosis on the anchor record is specific and has a relationship with a single episode, assign the anchor record to that episode.

- If the anchor record has more than one specific diagnosis and those diagnoses have clinical relationships with more than one episode, then use the strength of association of the procedure code for the anchor record to determine the episode that the anchor groups to.

- If the strength of relationship between the procedure code and the different episode conditions is the same for the specific diagnoses, then the strength of the association of the diagnosis codes themselves are used to determine the grouping of the anchor record. As discussed in Step 1D, strength of association between diagnosis codes and conditions are described as primary or incidental. Primary relationships between diagnosis codes and episode conditions have precedence over incidental relationships.

- If the strength of the relationship between the specific diagnosis codes and the episode conditions is the same, the time between the anchor record and the closest anchor for the open episode is used.

Non-specific diagnoses:

- If no specific diagnoses are observed on the anchor record, consider non-specific diagnoses in assigning the anchor record to an episode. Apply the same order of logic described directly above for specific diagnoses to the assignment of anchor records based on non-specific diagnoses.

At the completion of Step 2B, each anchor record with a specific or non-specific diagnosis has been assigned to an episode, including episodes of pneumonia.

Note that in the same way a single anchor record can start more than one episode (Step 2A), a single anchor record can also extend more than one episode, however the anchor record itself can only be assigned to one episode, as described above. For example, an anchor record with a diagnosis and procedure code combination that is eligible for pneumonia can extend a pneumonia episode. If that record also has a diagnosis and procedure code combination that is eligible for Hypertension, it can also extend a Hypertension episode. (See Section I of the Attachment for S2 above for a discussion of the concept of “phantom episode clusters” and the concept of extending episodes.)

Step 2C: Group Anchor Records to an Episode of pneumonia Using Sign and Symptom Diagnoses

The last step in grouping Anchor records to pneumonia and other episodes involves processing anchor records with only sign and symptom diagnosis codes. All sign and symptom diagnosis codes for pneumonia are listed within the S5_Pneumonia_DataDictionary on worksheet “IncidentalDxCodes” where column “specificity”=“Sign and Symptom”.

An example is Chest Pain (ICD-9 786.5).

For these anchor records with eligibility to a pneumonia episode, apply the following steps to assign the anchor record to an episode.

Step 2C1 - If the anchor record is only eligible for the open pneumonia episode, group the anchor record to the pneumonia episode.

Step 2C2 - If the anchor record is eligible for the pneumonia episode and another episode for the patient, apply the following tie-breaking steps to determine the episode an anchor record groups to:

- If the anchor record has more than one sign and symptom diagnosis and those diagnoses have clinical relationships with more than one episode, then use the strength of association of the procedure code for the anchor record to determine the episode that the anchor groups to.

- If the strength of relationship between the procedure code and the different episode conditions is the same for the sign and symptom diagnoses, then the strength of the association of the diagnosis codes themselves are used to determine the grouping of the anchor record. As discussed in Step 1D, strength of association between diagnosis codes and conditions are described as primary or incidental. For sign and symptom diagnoses, incidental relationships between diagnosis codes and episode conditions have precedence over primary relationships.

-If the strength of the relationship between the sign and symptom diagnosis codes and the episode conditions is the same, the time between the anchor record and the closest anchor for the open episode is used.
At the completion of Step 2C, each anchor record with a sign and symptom diagnosis has been assigned to an episode, including episodes of pneumonia.

After completing these steps, anchor records have been used to open episodes of pneumonia, as well as episodes for other conditions. Anchor records have been assigned uniquely to individual episodes based on the clinical logic described above and in the attachment "S5_Pneumonia_DataDictionary".

Step 3. Group Non-Anchor Records to Episodes.

Non-anchor records (record type "Ancillary" and "Pharmacy") can not open episodes on their own, but can join episodes. For example, a service for 89399 (Pathology lab procedure) and an ICD-9 code of 481 (pneumococcal pneumonia) can group to an open episode of pneumonia but can not open the episode itself.

Step 3A: Group Non-Anchor Records other than Pharmacy to an Episode of pneumonia Using Specific and Non-Specific Diagnoses

Once an episode of pneumonia is started and anchor records have been grouped, non-anchor records can group to that episode. Consider specific and non-specific diagnoses on non-anchor records first.

First identify whether the non-anchor record is eligible for pneumonia. Eligible non-anchor records for pneumonia have a procedure code eligible for pneumonia and a diagnosis code that has either a primary or incidental relationship to pneumonia. See the "ProcedureCodes" worksheet within S5_Pneumonia_DataDictionary for the procedure codes eligible for pneumonia. See the "Pharmacy" worksheet within S5_Pneumonia_DataDictionary for the pharmacy codes eligible for pneumonia. See the "PrimaryDxCodes" and "IncidentalDxCodes" worksheets within S5_Pneumonia_DataDictionary for a list of the diagnosis codes primary and incidental to pneumonia.

For non-anchor records with eligibility to a pneumonia episode, apply the following steps to assign the record to an episode.

Step 3A1 - If the non-anchor record is only eligible for the open pneumonia episode, group the record to the pneumonia episode.

In some cases, a non-anchor record can be eligible to join more than one episode. This is true because a service may have more than one diagnosis code. Further, diagnosis codes that are incidental for pneumonia may also be eligible for another ETG condition.

Step 3A2 - If the non-anchor record is eligible for the pneumonia episode and another episode for the patient, apply the following tie-breaking steps to determine the episode the record groups to:

- Assess the specificity of the diagnoses on the non-anchor record. Diagnosis class describes this specificity and was assigned to each diagnosis code in Step 1C (specific or non-specific).

- Assign the non-anchor record to an episode based on the diagnosis class. Episodes related to specific diagnoses take precedence over episodes related to non-specific diagnoses.

Specific diagnoses:

- If a diagnosis on the non-anchor record is specific and has a relationship with a single episode, assign the record to that episode.

- If the non-anchor record has more than one specific diagnosis and those diagnoses have clinical relationships with more than one episode, then use the strength of association of the procedure code for the record to determine the episode that the anchor groups to.

- If the strength of relationship between the procedure code and the different episode conditions is the same for the specific diagnoses, then the strength of the association of the diagnosis codes themselves are used to determine the grouping of the non-anchor record. As discussed in Step 1D, strength of association between diagnosis codes and conditions are described as primary or incidental. Primary relationships between diagnosis codes and episode conditions have precedence over incidental relationships.

- If the strength of the relationship between the specific diagnosis codes and the episode conditions is the same, the time between the non-anchor record and the closest anchor for the open episode is used.

Non-specific diagnoses:

- If no specific diagnoses are observed on the non-anchor record, consider non-specific diagnoses in assigning the record to an episode. Apply the same order of logic described directly above for specific diagnoses to the assignment of non-anchor records based on non-specific diagnoses.

At the completion of Step 3A, each non-anchor record with a specific or non-specific diagnosis has been assigned to an episode, including episodes of pneumonia.

Step 3B: Group Non-Anchor Records other than Pharmacy to an Episode of pneumonia Using Sign and Symptom Diagnoses

The last step in grouping non-anchor records to pneumonia and other episodes involves processing non-anchor records

with only sign and symptom diagnosis codes. All sign and symptom diagnosis codes for pneumonia are listed within the S5_Pneumonia_DataDictionary on worksheet “IncidentalDxCodes” where column “specificity”=“Sign and Symptom”. An example is Chest Pain (ICD-9 786.5).

For these non-anchor records with eligibility to a pneumonia episode, apply the following steps to assign the record to an episode.

Step 3B1 -If the non-anchor record is only eligible for the open pneumonia episode, group the record to the pneumonia episode.

Step 3B2 - If the anchor record is eligible for the pneumonia episode and another episode for the patient, apply the following tie-breaking steps to determine the episode the record groups to:

-If the non-anchor record has more than one sign and symptom diagnosis and those diagnoses have clinical relationships with more than one episode, then use the strength of association of the procedure code for the record to determine the episode that the record groups to.

-If the strength of relationship between the procedure code and the different episode conditions is the same for the sign and symptom diagnoses, then the strength of the association of the diagnosis codes themselves are used to determine the grouping of the non-anchor record. As discussed in Step 1D, strength of association between diagnosis codes and conditions are described as primary or incidental. For sign and symptom diagnoses, incidental relationships between diagnosis codes and episode conditions have precedence over primary relationships.

-If the strength of the relationship between the sign and symptom diagnosis codes and the episode conditions is the same, the time between the non-anchor record and the closest anchor for the open episode is used.

Step 3C: Group Pharmacy Records to an Episode of Pneumonia

Pharmacy services group differently than other non-anchor records because they usually do not have ICD-9 diagnosis codes associated with them to use in grouping. Instead, pharmacy records are assigned to pneumonia and other episodes using a table that maps NDC to a DCC code (Drug Category Code) based on the drug’s active ingredients and route of administration. A DCC to ETG map is then used to inform the grouping for the service. The relationship between DCC codes and pneumonia are described in the “Pharmacy” worksheet in the attachment “S5_Pneumonia_DataDictionary”. In some instances a DCC code may be eligible for pneumonia and another open episode for a patient. In these cases, where multiple episodes are observed for a patient where the DCC code has eligibility, use the strength of the clinical relationship between the DCC code and the episode to determine final assignment. The column “Rank” in the “Pharmacy” worksheet within attachment “S5_Pneumonia_DataDictionary” describes that strength of association. The lower the value is for Rank, the stronger the association between the DCC and the episode.

Due to the size of the attachment the full list of NDC to DCC mappings has not been provided within this submission. This file is available upon request. The DCC mappings included in the S5 attachment provide a summary of the key clinical relationships between drugs and the conditions described by the relevant ETGs. The NDC to DCC map would include the individual NDCs within a DCC that map to those relationships.

At the completion of Step 3C, all relevant records for pneumonia episodes have been assigned.

Step 4: Finalize the Episodes

Finalizing an episode of pneumonia involves determining whether or not the episode is complete, assigning co-morbidities and condition status factors and calculating a severity score and associated severity level. Co-morbidities and condition status factors will be discussed in section 8.3 and severity score calculation and level assignment is addressed in section 8.5.

In terms of episode completeness, pneumonia is an acute condition. Therefore the general clean period logic described in the attachment for question S2 above is applicable. All clinically consistent treatments for the care of a pneumonia patient will group to the episode of pneumonia for as long as data are available.

S8.3. Comorbid and interactions

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

Please note that this specification will reference different attachments included with the submission for these measures, including:

- S2_ETG_Construction_Logic_Pneumonia. This attachment provides an overview of ETGs and a summary of the

methodology used for pneumonia episodes.

- S8_Pneumonia_ClinicalLogic (Excel workbook attachment). This attachment includes Worksheets that describe the details around the components of pneumonia methodologies that relate to co-morbidities, condition status factors, and severity adjustment.

Co-morbidities and condition status factors are identified for each pneumonia episode. These factors provide specificity of the episode's clinical condition and also play a key role in assigning a severity score and level to the episode.

Steps to Assign Co-morbidities and Condition Status Factors to Pneumonia Episodes:

Step 1 – Condition Status Factors for Pneumonia Episodes.

Each pneumonia episode is evaluated to determine whether any Condition Status Factors for pneumonia are observed. To do this, the anchor records for the episode are evaluated using a comparison of their ICD-9 diagnoses with the diagnoses for the conditions status factors for pneumonia. The condition status factors used for pneumonia and the matching diagnoses for each are included in the "ConditionStatusToDxCODEMap" Worksheet in the attachment "S8_Pneumonia_ClinicalLogic".

The following condition status factors are defined for pneumonia:

- H. influenzae pneumonia
- K. pneumoniae pneumonia
- Pseudomonal pneumonia
- Staphylococcal pneumonia
- Pneumonia, with abscess
- Chlamydia pneumonia
- Legionnaires pneumonia
- Other gram negative pneumonia
- Anaerobic pneumonia
- Streptococcal pneumonia
- Pneumococcal pneumonia

If these Condition Status Factor diagnosis codes are present on the anchor records for a pneumonia episode, that condition status factor is recorded for the episode.

Each condition status factor belongs to a condition status group. If several condition status factors eligible for a given ETG occur in different condition status groups then all are assigned to the ETG. If there are 2 or more distinct condition status factors in a single condition status group then each condition status factor within the group has a ranking and only a single condition status factor is assigned to the ETG based on its ranking within the condition status group. If two or more condition status factors have the same ranking within a given condition status group, then they may or may not compete to be assigned to the ETG. If the compete flag is 1 or "on" then only one of the condition status factors is assigned to the ETG depending on which condition status factors occurs with the largest number of anchor claims. If the compete flag is 0 or "off" then all of the conditionStatuses with the same ranking are assigned to the ETG.

For the case of pneumonia, all condition status factors are in the same condition status group, all have the same priority and the compete flag is 0 (off). All condition status factors that occur for a given episode or pneumonia are assigned to the episode.

Step 2 –Comorbidity Factors for Pneumonia Episodes.

Each pneumonia episode is evaluated to determine whether any Comorbidity Factors for pneumonia are observed. To do this, the anchor records outside the pneumonia episode are evaluated using a comparison of their ICD-9 diagnoses with the diagnoses for the comorbidity factors for pneumonia. The comorbidity used for pneumonia and the matching diagnoses for each are included in the "ComorbtoDxCODEMap" Worksheet in the attachment "S8_Pneumonia_ClinicalLogic".

Examples of the comorbidity groups for pneumonia include AIDS, Septicemia and Other Infectious Diseases. In the example included in the S8_Pneumonia_ClinicalLogic attachment (see worksheet "ExSevScore&Level"), the comorbidities 80071 (Other Drug Dependence) are assigned to the pneumonia episode based upon the diagnosis information on anchor records that occur outside of the pneumonia episode.

As is the case for condition status factors, each comorbidity belongs to a comorbidity group and when several comorbidities are identified for an episode, the rankings of the comorbidities within the specific comorbidity groups and the status of the complete flag will determine whether or not the comorbidity is assigned to the given episode.

In addition to co-morbidities and condition status factors, interactions between two co-morbidities or two condition status factors are also assessed. These interactions are used in assigning severity to a pneumonia episode, as described in section 8.5.

S8.4. Clinical hierarchies

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

As noted in S8.2 and S8.3, ETG uses different clinical relationships between diagnosis and procedure codes and conditions to support the creation of pneumonia episodes. Many of these relationships involve clinical hierarchies, including how specific and non-specific and signs and symptoms diagnosis codes are used. The relationship between primary and incidental diagnoses and the strength of association of incidental diagnoses to pneumonia and other episode concepts is a further example. A third example is the procedure hierarchies that apply across all concepts for pneumonia. Please see the discussion for sections S8.2 and S8.3 and the attachment for S2 for a summary of the role of rankings, strength of association and hierarchies are used in the ETG methodology for pneumonia. Further, as described below in the discussion of severity adjustment, ETG also uses hierarchies to identify the most important co-morbidities within a related set of co-morbidities for use in measuring severity.

S8.5. Clinical severity levels

Detail the method used for assigning severity level and provide rationale for this methodology.

Please note that this specification will reference different attachments included with the submission for these measures, including:

- S2_ETG_Construction_Logic_Pneumonia. This attachment provides an overview of ETGs and a summary of the methodology used for pneumonia episodes.
- S8_Pneumonia_ClinicalLogic (Excel workbook attachment). This attachment includes Worksheets that describe the details around the components of pneumonia methodologies that relate to co-morbidities, condition status factors, and severity adjustment.

More specifically, apply the following steps:

Step 1 – Identify Condition Status Factors and Comorbidities in an Episode

Assignment of severity occurs after the identification of condition status factors and comorbidities as detailed in specification S8.3. Interactions between various co-morbidities also play a role in severity assignment as well as demographic factors. The combination of all of these factors are used to describe a “severity” score and level for an episode, where a higher level of severity indicates an expectation of a higher level of resources required to diagnose, manage and treat an episode of pneumonia.

The steps required to identify condition status and comorbidity factors for pneumonia are described in S8.3.

Step 2 – Map Episode Comorbidities to the Final Comorbidities used to Calculate Episode Severity

The individual comorbidities identified in S8.3 are further grouped to the final comorbidity factors used in calculating episode severity. This step is performed to combine the effects of related comorbidities on severity. Further, in some cases, hierarchies are used to limit final factors to those comorbidities within a related group that have the greatest impact on episode severity. For example, for pneumonia, Other inflammatory lung diseases, Occupational & environmental pulmonary diseases, Chronic bronchitis, and Asthma are all qualified as comorbidities and are all conditions categorized as Bronchial Inflammation. Given the related nature of these comorbidities, only one factor is used as the final comorbidity factor for computing severity. Steps 2.1 through 2.4 describe how this final comorbidity is selected.

Worksheet “Comorbidities” – includes the ComorbidityCodes and Comorbidity Groups used to determine severity for pneumonia. Co-MorbidityGroup2 is the final comorbidity factor used to compute episode severity. To determine this factor:

Step 2.1 – Assign ComorbidityGroup1 and ComorbidityGroup2 to each ComorbidityCode. Using Bronchial Inflammation as an example, Other inflammatory lung diseases, Occupational & environmental pulmonary diseases, Chronic bronchitis, and Asthma would all be assigned to Bronchial Inflammation for ComorbidityGroup1. Other inflammatory lung diseases, Occupational & environmental pulmonary diseases, and Chronic bronchitis would be assigned to "Bronchial Inflammation 2" for ComorbidityGroup2 and Asthma would be assigned to "Bronchial Inflammation 1" for ComorbidityGroup2.

Step 2.2 – Assign Priority to each ComorbidityCode. Other inflammatory lung diseases, Occupational & environmental pulmonary diseases, Chronic bronchitis, and Asthma would be assigned a Priority value of 1, 2, 3, and 4, respectively.

Step 2.3 – Across all of the values for ComorbidityCode within each ComorbidityGroup1, select the ComorbidityCode with the lowest value for Priority. As an example, if Chronic bronchitis and Asthma were both observed, Chronic Bronchitis would be selected due to its lower value for Priority (a Priority value of 3 take precedence over a Priority value of 4)

The remaining values for ComorbidityCode and ComorbidityGroup2 define the final comorbidity factors used in determining pneumonia severity. In the above example (where Chronic bronchitis and Asthma were both observed), Bronchial Inflammation 2 (Chronic Bronchitis) would be selected as the final comorbidity within Bronchial Inflammation.

Step 2.4 – Assign a risk weight to each remaining factor. Each risk weight reflects the incremental contribution of having a specific comorbidity factor on pneumonia severity. If the patient’s age is less than 65, assign a risk weight using the column “Weight”. If the patient’s age is 65 or higher, use the risk weight using the value in the column ElderlyWeight. Use patient age as of the ending date for the measurement period to determine the appropriate weight. For Bronchial Inflammation 2, a risk weight of 0.78 would be assigned for a non-elderly patient. A risk weight of 0.78 would also be assigned for an elderly patient.

Step 3 – Identify Comorbidity Interactions

The interaction between two observed comorbidities can contribute to episode severity. Worksheet “ComorbidityInteractions” includes the interactions between Comorbidity Groups used to determine severity for pneumonia. The table describes pairings of the final comorbidity factors produced by Step 2 (identified by the values for ComorbidityGroup2).

Step 3.1 – Identify pairings of ComorbidityGroup2 for the episode that are also observed in the Worksheet “ComorbidityInteractions”

Step 3.1 – Assign a risk weight to each qualified interaction. Each risk weight reflects the incremental contribution of having a specific comorbidity interaction on pneumonia severity. If the patient’s age is less than 65, assign a risk weight using the column “Weight”. If the patient’s age is 65 or higher, use the risk weight using the value in the column ElderlyWeight. Use patient age as of the ending date for the measurement period to determine the appropriate weight.

Step 4 – Identify Comorbidity Counts

For some ETG conditions the number of final comorbidity factors will impact episode severity – for example, where 3 or more co-morbidity factors are observed. For these episodes, a separate Worksheet “ComorbidityCounts” includes these additional severity factors and their assigned risk weights added for those episodes. pneumonia does not include any Comorbidity Count factors; this step does not apply to pneumonia.

Step 5 – Condition Status Factors

The Worksheet “ConditionStatuses” – includes the Condition Status factors used to determine severity for pneumonia. The rightmost columns include risk weights for the non-elderly and elderly models. Each risk weight reflects the incremental contribution of having a specific Condition Status factor on pneumonia severity.

For each condition status factor observed, assign a risk weight. If the patient’s age is less than 65, assign a risk weight using the column “Weight”. If the patient’s age is 65 or higher, use the risk weight using the value in the column

ElderlyWeight. Use patient age as of the ending date for the measurement period to determine the appropriate weight.

Step 6 – Identify Condition Status Interactions

For some ETG conditions, the interaction between two observed condition status factors can contribute to episode severity. A separate tab, Worksheet “ConditionStatusInteractions” would be used to identify qualified pairings and their weight in calculating severity. Pneumonia episodes do not use condition status interactions in calculating severity. Step 6 does not apply to pneumonia.

Step 7 – Identify Condition Status Counts

For some ETG conditions the number of final condition status factors will impact episode severity – for example, where 3 or more condition status factors are observed. For these episodes, a separate Worksheet “ConditionStatusCounts” includes these additional severity factors and their assigned risk weights added for those episodes. Pneumonia does not include any condition status count factors; this step does not apply to pneumonia.

Step 8 – Assign Demographic Factors

The Worksheet “Demographics” includes the additional severity factors added based on age and gender. Each risk weight reflects the incremental contribution of having a specific Demographic factor on pneumonia severity. Based on patient age, assign the patient to an age range group. Using gender and age group, assign a demographic factor weight. Use patient age as of the ending date for the measurement period to determine the appropriate age range group.

Step 9 – Compute Severity Score

Sum the risk weights assigned for each of the relevant factors identified above. The sum of these weights is the overall severity score for the episode. As noted above, the higher the severity score for an episode, the more resources are expected relative to other pneumonia episodes.

As a note, the estimation of the risk weights used in computing severity for pneumonia episodes is based on empirical analyses of healthcare data for a benchmark population of over 25 million individuals. In particular, multivariate regression analyses were used where cost per episode for individual pneumonia episodes was the dependent variable and the defined array of co-morbidity and condition status factors and patient age and gender were the independent variables. The model was run separately for individuals 65 and over and those under 65 years of age. The resulting estimated parameters were used to assign weights to each factor described in the above tables. These weights and the presence of a particular set of factors for an episode are used to determine a pneumonia severity score for the episode.

Step 10 – Compute Severity Level

Based on the severity score, the severity “level” indicates a categorical ranking of where the specific episode is relative to the population of all pneumonia episodes. There are four potential severity levels for pneumonia, where the value 1 indicates a less severe episode and the value 4 indicates the most severe episode. The “Thresholds” Worksheet in attachment “S8_Pneumonia_ClinicalLogic” describe the three cut-off points that define the four levels of severity for pneumonia episodes.

Assign severity level to the episode depending on the episode severity score calculated in Steps 1-9 and where that score falls within the ranges defined in the “Threshold” Worksheet.

Example: Assigning Severity Score and Level to Pneumonia Episodes

The example included within the S8_Pneumonia_ClinicalLogic attachment (see worksheet “ExSevScore&Level”) shows the calculation of severity score and level for a pneumonia episode.

The example describes a Female patient, age 32, observed to have a number of anchor records with a diagnosis that maps to the pneumonia ETG. The patient is also observed to have one condition status factor and one comorbidity that are also eligible for pneumonia. The condition status factor (70142: Mycoplasma pneumonia) was identified through one or more anchor records observed within the episode where the diagnosis on the records mapped to that condition status factor. The comorbidity (80071: Other drug dependence) was identified on one or more anchor records observed outside of the pneumonia episode.

Assign severity markers and weights: The patient receives a severity marker for each of the condition status and comorbidity factors and a risk weight is assigned to each. The patient also receives severity weight related to her age and gender which fall into the "Female 19-34" range.

Calculate severity score: A severity score of 0.4807 is calculated based upon the sum of:

-The Demographic weight of 0.2498 (see worksheet "Demographics" within S8_Pneumonia_ClinicalLogic where column "gender"=F and column "ageRange"=19-34);

-The condition status weight for Mycoplasma pneumonia of 0 (see worksheet "ConditionStatuses" within S8_Pneumonia_ClinicalLogic where column "conditionStatusCode"=70142),

-The comorbidity weight for Other drug dependence of 0.2309 (see worksheet "Comorbidities" within S8_Pneumonia_ClinicalLogic where column "comorbiditycode"=80071. The Other drug dependence comorbidity belongs to the Comorbiditygroup2 of Drug Dependence.);

-The final severity score is calculated as $0.2498 + 0 + 0.2309 = 0.4807$

Calculate severity level: The severity score of 0.4807 falls with the range of <0.7 and the episode is assigned to Severity Level 1.

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.

ETG does provide methodology to deal with cases where a code will shift an episode from one ETG to another. For example, a concurrent renal transplant procedure will shift an episode of ETG Chronic renal failure to an episode of ETG Kidney transplant. For pneumonia there are 21 diagnosis codes which would cause an episode of pneumonia to shift to an episode of Viral pneumonia. An example of a diagnosis code that will cause pneumonia to shift to Viral pneumonia is 480.2 (pneumonia due to parainfluenza virus). Please refer to worksheet "IncidentalDxCodes" within attachment S5_Pneumonia_DataDictionary for a list of diagnosis codes that will result in an episode shift (see where column "diagnosisEligibilityType"=5). Within the same spreadsheet the columns "shiftedEtg" and "shiftedEtgDescription" describe the ETG the episode will shift to.

In the case where a diagnosis and procedure code on a claim are eligible for multiple episodes, a specific hierarchy of rules determines the most appropriate episode to group to, based on the rankings of the diagnosis and procedure code for the ETG of each episode. All of the eligibility and ranking information is contained in a set of tables within the ETG application.

For more information about episode building construction/logic, please refer to the attachment for S.2 .

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:

Please supply the username and password:

Attachment: ETG Construction Logic Pneumonia-634413056088309687.doc

S9.1. Brief Description of Construction Logic

Briefly describe the measure's construction logic.

Please refer to information provided in S2 and S8 for construction logic

S9.2. Construction Logic

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

Please refer to information provided in S2 and S8 for construction logic

S9.3. Measure Trigger and End mechanisms

Detail the measure's trigger and end mechanisms and provide rationale for this methodology.

As described in detail in S8, an episode is triggered by an anchor record. This is a claim record with a procedure indicating a face to face physician encounter, a surgical procedure by a physician or a facility charge indicating a confinement. The rationale for this is that the diagnosis and procedure codes on these record types are most likely to specify a valid clinical condition related to the individual. The length of the episode will depend on the subsequent records that occur within the ETGs clean period. When there is an interval longer than the clean period of the episode without any records eligible to group to the episode, it is considered complete.

Pneumonia is an acute condition. Therefore the clean periods described in Section I as part of the general ETG methodology are applicable. All clinically consistent treatments for the care of a Pneumonia patient will group to the episode of Pneumonia based upon the clean period.

For more information about episode building construction/logic, please refer to S8 and the attachment we provided in s.2 .

S9.4.Measure redundancy or overlap

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

The ETG application is able to keep related conditions separate. For example, suppose that there are concurrent episodes of pneumonia and Diabetes and there is record eligible for both ETGs. A specific hierarchy of rules coupled with a set of eligibility tables with strengths of association of each diagnosis and procedure code for each ETG will uniquely determine which episode the record will group to. There are no ambiguous assignments and episode assignment of each claim record will be unique. For more information about episode building construction/logic, please refer to S8 and the attachment we provided in s.2 .

S9.5.Complementary services

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

ETG does not group based on complimentary services. All claims group to the appropriate episode on their own merits.) For more information about episode building construction/logic, please refer to the attachment we provided in s.2 .

S9.6.Resource Use Service Categories

Inpatient services: Inpatient facility 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

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.

The following resource-use categories are included as measures for this submission.

Cost of Care per Episode

1. Total
2. Primary Care Core Services, Total
3. Primary Care Core Services, Visits
4. Primary Care Core Services, Other (Non-Visits)
5. ER Services
6. Hospital Services, Total
7. Inpatient Acute
8. Inpatient Non-Acute
9. Other Outpatient
10. Laboratory Services
11. Radiology Services, Diagnostic, Total
12. Radiology, MRI, CT Scan Services
13. Radiology, Other Diagnostic Services
14. Specialty Care Services, Total
15. Specialty Care, Other Diagnostic Testing Services
16. Specialty Care, Evaluation & Management Services
17. Specialty Care, Medicine Services
18. Specialty Care, Surgery Services
19. Specialty Care, Other Services
20. Pharmacy Prescription Services

Utilization per 1,000 Episodes

1. PCP Visits
2. Specialist Visits
3. Specialist Referrals
4. Total Evaluation & Management Visits
5. ER Visits
6. Hospital Inpatient Admits, Acute
7. Hospital Inpatient Days, Acute
8. Laboratory Services
9. Radiology Services, Diagnostic, Total
10. Radiology Services, MRI/CT Scan Services
11. Radiology Services, Other Diagnostic Services
12. Pharmacy Prescriptions Services

Each resource use category measure is described below, including reference to the specific codes and logic used to identify the services involved.

I. General Methods

The following notes on General Methods apply to all resource measures described here and provide guidelines on service costs, the treatment of incomplete and outlier episodes, and the selection of time periods. The logic described for type of service plays a specific role in each measure. These general methods are employed across all submitted measures:

-- Service cost – as a guideline, the service cost used in resource use measurement should reflect the actual payments or costs associated with the service or a standard-priced resource cost amount. As a further guideline, the financial amount used in resource measurement should reflect all payments for a service, including those made to the provider by payer, patient and other entities. The allowed or equivalent payment is an example.

-- Complete episodes – Only complete episodes should be included in resource measurement. See the attachment for s.2 for a discussion of how ETG assigns completion status to an episode.

-- Outlier episodes – as a guideline, low outlier cost episodes should be excluded from resource use measurement. High outlier cost episodes should be included, but all costs truncated at the high outlier cost threshold used for the episode (a technique called “winsorization”). Where costs by type of service are used in measurement, individual service costs can be pro-rated to reflect the truncated total cost for a high cost outlier episode.

-- Episode Time periods – as a guideline, the episodes included in resource use measure should focus on a specific 12 month period, for example, all episodes ending in calendar year 2010.

-- Selecting Clinical Episodes – For pneumonia, select all remaining episodes with pneumonia Base ETG

-- Type of Service. The type of service logic for each measure is described in the sections below. Each type of service definition includes an overview of the key steps used in identifying the relevant services used in measuring cost and utilization. As an initial step, prescription pharmacy services and hospital inpatient confinements are identified (more detail below). For the remaining services:

- a. Providers are categorized into facility, anesthesiology specialties and other professional (not anesthesiology);
- b. The attached document S9.5_RU_Categories then describes two levels of specifications used in assigning services to a type of service category;
- c. The first table in the attachment IMAP_TOS_PROC includes one row per procedure code (CPT, HCPCS, Revenue). For each row, the table includes the procedure code, a short description and the columns PROFTOS, ANESTOS, OPTOS, and PCC_TYPE. PROFTOS, ANESTOS, OPTOS include standard TOS_I codes that are assigned to each procedure code based on whether the provider is a facility, anesthesiologist or other professional, using OPTOS, ANESTOS and PROFTOS, respectively;
- d. Some services are also assigned a value for PCC_TYPE (described below);
- e. The second table, IMAP_TOS, includes one row for each of the standard TOS codes included in PROFTOS, ANESTOS and PROFTOS and columns for the TOS_I codes, ENC_TOS, and ENC_TOP and a brief description of the TOS_I. ENC_TOS and ENC_TOP are used in defining encounters below.
- f. These two tables are used in creating the measures described below.

-- Encounters. An Encounter is contact between an individual and the health care system for a related set of services. It is based on the type of service and the type of provider for a member on a specific day. Providing the ability to view data by encounters helps convey the scope and influence of all services associated with patient-health care system meetings. The concept of an encounter is used for the utilization measures described below. The following steps are used to assign an encounter value to each service record:

- a. Hospital inpatient admissions. A hospital inpatient confinement is considered a single encounter (ENCOUNTER=1).
- b. Prescription pharmacy. A pharmacy service record (claim record) is considered a single encounter (ENCOUNTER=1).
- c. Ancillary Drug Administered Services. All Ancillary, Drugs Administered (TOS_I values 201 thru 211), are considered an encounter (ENCOUNTER=1).
- d. For all other services, the number of encounters is dependent on the Type of Service and the Type of Provider assigned to the claims. In particular, the values included in the table IMAP_TOS for Encounter Type of Service (ENC_TOS) and Encounter Type of Provider (ENC_TOP) are used. As shown in IMAP_TOS, both the Encounter TOS and Encounter TOP are based on Type of Service (TOS_I) and can be assigned using table IMAP_TOS, and joining on TOS_I from the service record.
- e. For these other services, medical service records are sorted by Member, Date of Service, ENC_TOS and ENC_TOP.
- f. The calculation of encounters for services other than emergency room, laboratory and radiology services is 1 divided by the total number of records in the combination of Member, Date of Service, Encounter TOS, and Encounter TOP.
- g. Additional logic. Emergency room, laboratory and radiology services need to have a different logic because these services often are billed using both a technical and professional component – where both a professional provider and facility provider are involved.
- h. Any service with the following Encounter TOS values will use the additional logic when calculating encounters.
 1. ER professional and facility services (ENC_TOS=24)
 2. Lab and pathology professional and facility services (ENC_TOS=29, 31)
 3. Diagnostic and therapeutic radiology professional and facility services (ENC_TOS=47, 49)
- i. For the services using the additional logic, for each Member, Date of Service, and ENC_TOS distinct

combination, sum the number of records for each of the Encounter TOP values of 1 and 2.

1. Two cases can exist for these services: there are both facility and professional records in the combination; or there are only facility records or only professional records.
 2. Where at least one facility record and one professional record, the encounter is divided up equally between the professional and technical components. Therefore, the calculations for Encounters for these situations are: 0.5 divided by {number of records with Encounter TOP = 1 (Facility)} and 0.5 divided by {number of records with Encounter TOP = 2 (Professional)}
 3. Where all records have the same ENC_TOP value, the encounters calculation will be the generic calculation: 1 divided by {number of records in the combination of Member, Date of Service, Encounter TOS, Encounter TOP}
- Cost and Utilization Measures. The actual resource use for an episode is the sum of the costs or encounters for those services grouped to the episode. Measures of actual cost or use per episode across episodes, is the sum of cost or use divided by the total number of episodes included in the measurement.

II. Cost of Care per Episode

Total Service Costs. Total services costs include the total costs for all services included in the selected clinical episodes.

Primary Care Core Services Costs. Primary Care Core (PCC) services include a select group of services traditionally performed by an individual's primary care physician. The PCC concept is similar to the idea of the group of services typically included in a primary care capitation definition. In particular, these services include non-inpatient evaluation and management services and selected imaging, diagnostic and minor procedure services. PCC Services are identified as follows:

-- First select services rendered by a primary care provider. The identification of primary care providers can be made configurable. At a minimum, these providers include the individual's assigned PCP. Further, to include covering providers, other primary care providers in the network are included, defined using either a list of provider ids or all physicians with a specialty of internal medicine, family practice, geriatric medicine, adolescent medicine and pediatrics, or both (e.g., using a list to include specific OB/GYN providers in addition to all providers with primary care specialties).

- i. The CPT procedure code on the selected services is then used to identify:
 1. PCC Services Total
 2. PCC Services, Visits and
 3. PCC Services Other.
- ii. The CPT procedure codes assigned to these categories are included in the column PCC_TYPE in the attachment table IMAP_TOS_PROC. Values of "Visit" and "Other" are used. Blank entries for a procedure code indicate that they are not included as a PCC service.

-- ER Service Costs. These services include professional and facility emergency room services.

- i. Professional ER Services are identified as having values of 1803 thru 1805 in IMAP_TOS
- ii. Facility ER Services are identified as having values of 801 and 802 in IMAP_TOS

-- Hospital Costs. Includes the facility cost of an inpatient stay and services provided by an outpatient facility other than those defined elsewhere (e.g., ER, Lab, Radiology, Other). These services include professional and facility emergency room services.

- i. Inpatient Acute Services are identified as having a value of 601 in IMAP_TOS
- ii. Non-Inpatient Acute Services are identified as having a value of 703 in IMAP_TOS
- iii. Other Outpatient Hospital Services are identified as having values of 901 thru 1399 in IMAP_TOS

-- Laboratory Services. These services include professional and facility laboratory services, other than those professional services assigned to Primary Care Core.

- i. Professional Lab Services are identified as having values of 2101-2118 (Professional, Lab) or 2501-2511 (Professional, Pathology) in IMAP_TOS
- ii. Facility LAB Services are identified as having values of 1001 thru 1005 in IMAP_TOS

-- Radiology Services, Diagnostic. These services include diagnostic professional and facility radiology services, other than those professional services assigned to Primary Care Core:

- i. Professional Radiology, MRI, CT Scan Services are identified as having values of 2901 thru 2903 in

IMAP_TOS

- ii. Facility Radiology, MRI, CT Scan Services are identified as having values of 1201, 1203, 1204 in IMAP_TOS
- iii. Professional Radiology, Other Diagnostic Services are identified as having values of 2905, 2906, 2907, 2908 in IMAP_TOS
- iv. Facility Radiology, Other Diagnostic Services are identified as having values of 1202, 1206, 1207, 1208 in IMAP_TOS
- v. Note that Therapeutic Radiology is included in Specialty Care Services, Medicine

-- Specialty Care Services. These services include those services not identified above and are categorized as follows (including TOS_I values in IMAP_TOS):

- i. Specialty Care, Other Diagnostic Testing
 - 1. 1701-1733 (Professional, Diagnostic)
- ii. Specialty Care, Evaluation & Management
 - 1. 1601-1609 (Professional, Consult)
 - 2. 2001-2013 (Professional, Inpatient Visit)
 - 3. 2401-2411 (Professional, Office Visit)
 - 4. 2717-2719 (Professional, Home Visit)
 - 5. 2729-2731 (Professional, Domiciliary/Rest Home Visit)
 - 6. 2801-2807 (Professional, Preventive Medicine)
 - 7. Excludes any services assigned to Primary Care Core
- iii. Specialty Care, Medicine
 - 1. 1401-1405 (Professional, Allergy Tests)
 - 2. 1901-1901 (Professional, Immunizations / Injection)
 - 3. 2909-2915 (Professional, Therapeutic Radiology)
- iv. Specialty Care, Surgery
 - 1. 3001-3214 (Professional, Surgery)
- v. Specialty Care, Other
 - 1. 101-131 (Ancillary, DME)
 - 2. 201-211 (Ancillary, Drug Admin)
 - 3. 301-307 (Ancillary, Home Health)
 - 4. 401-403, 431 (Ancillary, Services and Supplies)
 - 5. 405-414 (Ancillary, Med and Surg Supplies)
 - 6. 416-424 (Ancillary, Orthotics)
 - 7. 425-429, 432 (Ancillary, Supplies)
 - 8. 433-436 (Ancillary, Oxygen/Resp)
 - 9. 437-446 (Ancillary, Prosthetics)
 - 10. 448-449 (Ancillary, Vision)
 - 11. 450-459 (Ancillary, Rpt/Trking)
 - 12. 501-503 (Ancillary, Transportation)
 - 13. 1501-1599 (Professional, Anesthesia)
 - 14. 2203-2212 (Professional, Mental Health)
 - 15. 2302-2317 (Professional, Obstetrics)
 - 16. 2601-2625 (Professional, Phys Medicine/Rehab)
 - 17. 2701-2715, 2721-2728 (Professional, Professional Other)

III. Utilization per 1,000 Episodes

Encounters are used for all utilization counts for the utilization measures described below.

Evaluation and Management Visits. E&M Visit services by all professional providers and include the following TOS_I values from IMAP_TOS:

- i. 1601-1609 (Professional, Consult)
- ii. 1803-1805 (Professional, ER)
- iii. 2001-2013 (Professional, Inpatient Visit)
- iv. 2401-2411 (Professional, Office Visit)
- v. 2717-2719 (Professional, Home Visit)
- vi. 2729-2731 (Professional, Domiciliary/Rest Home Visit)
- vii. 2801-2807 (Professional, Preventive Medicine)

PCP Visits. PCP Visits include E&M visits rendered by a PCP or a PCP covering provider (see discussion above for PCC services).

Specialist Visits. Specialist Visits include E&M visits rendered by a provider other than a PCP or a PCP covering provider (see discussion above for PCC services).

Specialist Referrals. A Specialist Referral is indicated using E&M visits and indicates the first instance of the Provider for an E&M service for that member. A specialist is a provider other than a PCP or a PCP covering provider (see discussion above for PCC services).

ER Visits. Indicates an ER service encounter. ER services are defined by a TOS_I value of Facility Outpatient, ER (801, 802) or Professional, ER (1803, 1805).

Radiology Services, Diagnostic. Radiology utilization is defined as an encounter for the following Types of Service:
 -MRI/Cat Scans – Facility Outpatient (1201, 1203, 1204), Professional (2901, 2902, 2903)
 -Other Diagnostic Radiology – Facility Outpatient, Diag. Radiology (1202, 1206, 1207, 1208), Professional, Diagnostic Radiology, Nuclear Medicine (2905 thru 2908)

Laboratory Services. Laboratory utilization is defined as an encounter for the following Types of Service:

- i. Facility Outpatient, Lab (1001, 1003, 1005)
- ii. Professional, Lab, (2101 thru 2118)
- iii. Professional, Pathology (2501 thru 2511)

Pharmacy Services. A pharmacy service prescription record.

Inpatient Admits and Days. Number of unique inpatient stays. An inpatient stay describes the entire stay by a patient in a facility at the same level of care. Transfers to a different level of care at the same facility results in a new admission. Acute inpatient stays describe inpatient confinements in an acute care facility. Non-acute inpatient stays describe inpatient confinements in a skilled nursing facility, transitional care unit/rehab, or other longer term/sub-acute facility. Inpatient days describe the difference between inpatient admission and discharge dates. Inpatient stays where the admission and discharge dates are equal are assigned one inpatient day.

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

URL:

Please supply the username and password:

Attachment: S9.7_RU_Categories-634413225228606337.xls

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

Emergency Medical Services/Ambulance

Home Health

Hospice

Hospital/Acute Care Facility

Imaging Facility

Laboratory

Post Acute/Long Term Care Facility : Nursing Home/Skilled Nursing Facility

Post Acute/Long Term Care Facility : Rehabilitation

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.

The attachment for S2 and responses to S8 above provided a description of the approach used by ETG to assign a severity score and level to each pneumonia episode. To do this, ETG first assesses the observed co-morbidities and condition status factors for an episode and the patient's age and gender. ETG then assigns a weight to each factor found to influence the relative risk of an episode of pneumonia. These weights and factors are condition-specific and were estimated using pneumonia episode results for a large population. The overall severity score for an episode is the sum of these weights for all factors observed. Using the severity score, a severity level is created, with each pneumonia episode assigned to one of four severity levels.

The approach used by ETG to assign episode severity has several advantages. First, the approach uses broad clinical profile of an episode, describing its clinical status and that of the patient. Second, the weightings assigned describe the incremental contribution of each factor to overall episode severity. Further, the approach used for severity is condition-specific – a separate model and weightings are constructed for each condition, including pneumonia. These severity results provide the key information required to support risk adjusted comparisons using pneumonia episodes.

Risk adjustment is an important step in resource use measurement. Measures of the cost of care for an organization or provider can be impacted by the underlying risk and severity of the patients they enroll or manage. Case-mix or risk adjustment addresses these differences and supports more consistent and equitable comparisons. These approaches allow a focus on differences in resource use deriving from differences in the practice of medicine rather than differences in the mix of episodes or patients.

The level of severity assigned by ETG to an episode is used to support risk adjustment. The risk adjustment approach includes three important steps:

- Compute the observed experience for the provider being measured, across all episodes to be included in the comparison;
- Compute the experience for peers or a best practice benchmark. Compute this experience at the level of the risk adjustment, in this case ETG base condition and severity level. For a peers benchmark, average cost per episode across all peers for the ETG base condition and episode level can be computed;
- Compare the observed experience with the risk adjusted peers or benchmark experience – often called the “expected” result. This expected result is adjusted to reflect both the peers/benchmark levels of performance and also the provider's own case mix of episodes by condition and level of severity. The ratio of observed to expected results can be termed the relative cost ratio and is a risk adjusted measure.

The table in S10.1 provides an example comparing the cost of care performance of two cardiologists using episodes of care and the condition of CHF. The analysis used only complete, non-outlier CHF episodes. The upper section of the table summarizes results at the condition and severity level. A higher severity level for a condition indicates the presence of one or more condition status factors and/or co-morbidities that impact the resources required for treatment. The table also summarizes results for CHF, across all severity levels.

The table shows the number of episodes attributed to the cardiologist, the observed cost per episode, peers cost per episode (the “expected” amount), and the ratio of the cost per episode of the cardiologist to his peers. By condition and severity level, the peers cost per episode is the average experience of all cardiologists included in the measurement for those episodes. The peer's experience is risk adjusted and assumes the same mix of episodes (by condition and severity) as the physician being measured. Notice that for the overall CHF summary, the peers cost per episode for Dr. Jones is \$2,081, while that amount for Dr. Smith is \$1,841. The higher amount for Dr. Jones indicates a higher case-mix and greater expected costs relative to Dr. Smith. These peer amounts, adjusted for the specific mix of episodes observed for the physician being measured, capture the risk adjustment appropriate for the analysis.

In the last column, a relative cost ratio less than 1.00 indicates that the observed cost per episode for a provider is less than his peers. As shown, Dr. Jones cost is lower than peers and Dr. Smith is higher cost than peers. An additional

report using the same measure information could summarize results by type of service, or specific utilization such as the use of a specific diagnostic test or treatment, providing greater insights into the factors behind differences in resource use. The risk adjustment for these measures would use the same approach as described here for total cost per episode.

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: S10_Risk Adjustment Method Example-634413225817356337.xls

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

ETG stratifies episodes by the intensity of service, or total cost. For a given episode, a severity score is assigned based on demographic factors (gender and age) and the presence of comorbidities and complications. The determination of this severity score is described in sections 8.3, 8.4 and 8.5. Once a severity score is determined, a severity level, a number between 1 and 4 is assigned based on a table that relates severity levels to severity scores for each ETG. The method for determining the severity levels is described in section 8.5. The severity level can then be used to stratify episodes by severity, measured as resource consumption.)

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.

The financial amounts used should be complete and valid, reflecting the total payments related to the service. The financial amount used in resource measurement should reflect all payments for a service, including those made to the provider by payer, patient and other entities. The allowed or equivalent payment is an example. The use of allowed payments provides the best estimate of the actual costs involved in delivering the medical and pharmacy services included in the measure. Allowed payments will reflect both the quantity of different services provided as well as the actual unit price of those same services. Allowed amounts are used extensively in the industry as a measure of cost of care, including comparison of physicians and delivery systems.

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.

Attributing patients and episodes to appropriate physicians and groups is a challenging step in cost measurement. Over some period of time a patient can have multiple conditions and, in many cases, multiple providers caring for the same condition. For example, for an episode of hypertension, a patient can be managed by their primary care physician, an internist, and also receive services from a cardiologist. For a patient with coronary artery disease, an internist, a cardiologist, and a surgeon can all play a key role in providing the patient's care. A methodology is required to identify these episodes for a patient and the providers responsible for the services performed within those episodes. As a guideline, some principles are involved in determining a valid approach to be used in assigning episodes:

- The approach must be valid conceptually. It must be defensible, understandable and accepted by providers, health plans, and other users of the measurement results;
- The approach must be supported by readily available information, including the outputs from an episode grouping;
- The approach should be robust across applications – working well for different sources of health plan data, patient populations and over time;
- The approach should be flexible and consider the characteristics of the specialists being compared and the nature and severity of their patients and episodes;
- Both activity-based and population-based approaches should be supported. An activity-based approach, describes attribution where an episode is assigned to the providers responsible for the greatest amount of activity during the course of the episode. Activity can be measured using different concepts including service costs, episode clusters, or patient visits.

A population, or panel-based approach is sometimes used when measuring performance for primary care physicians (PCPs), in particular where providers are performing a gatekeeper function for a population of members. In this case, responsibility for a member's qualified episodes of care may be attributed to the member's PCP — whether or not the PCP provided any of the services for that member during those episodes.

- “Sufficient” evidence of the provider's responsibility for the episode should exist. Thresholds should be considered that prevent providers from “winning” episodes where they have a small amount of involvement – relative to their physician peers or relative to all physicians involved in the episode.
- Attributing the same episode to multiple providers in different specialties should be considered, when appropriate.

Care during an episode can include two types of services: services where important clinical decisions are made regarding the course of care and services that are a response to those decisions. Office visits, consultations and other evaluation and management services are examples of the first type of services. As part of these services, decisions to perform tests, prescribe drugs or order other ancillary services are made. The second type of service includes diagnostic lab, imaging, other tests, DME, drug therapies and treatments. These services are typically responses to decisions made regarding the course of care. Some services, such as surgery, may describe a closely linked bundle of care and relate to both categories – where the surgeon has some role in the decision to perform the procedure and also performs the surgery itself.

The dichotomy above suggests two important concepts for assessing approaches to attribution. First, the measure of “activity” to be used in identifying a responsible provider should focus on those types of service where decisions regarding the course of care and management of the episode take place. Second, the decision on the approach to be used for attribution may differ by specialty. In the case of a group of providers such as surgeons, where the majority of their services may be of the second type – after the decision to undergo surgery has been made – using cost as the activity measure for attribution may make sense. However in the case of PCPs or medical specialists, non-acute E&M visits or the number of episode clusters (qualified services), may be a superior service activity measure for determining episode responsibility.

As a guideline, four different general options for physician episode attribution can be considered to attribute episodes to individual providers – three activity-based and one population-based approach. Each of these options can be supported using standard outputs from ETG and the measures described in this submission. For each option, the description below assumes the following steps have been performed prior to attribution:

- ETG episode grouping – producing the detail and summary output files to be used in attribution and measurement;
- Identification of the comparison peer group and the individual physicians to be included;
- The selection of qualified episodes for the peer group. Qualified episodes include those episodes with an ETG that matches the pre-defined list to be used for that peer group. Qualified episodes are further limited to complete, non-outlier episodes that fall within the time period defined for measurement.

For this discussion, it is assumed that the objective is to assign a single winner, if possible, for each peer group in which the episode is relevant, but allow providers in different peer groups to be assigned the same episode. To support this, the following logic would be applied separately, peer group by peer group. The activity-based options are described first.

Although these approaches are described for attribution at the individual physician level, they could also be applied using physician groups as the unit for attribution.

Approach 1 - Physician Episode Attribution using Professional Service Costs. This attribution approach identifies the responsible physician for an episode as that provider rendering the greatest amount of professional service costs during the episode.

Professional services are those performed by a clinician in managing and treating the patient during an episode of care, including visits and consultations, surgery and therapies. Professional services exclude inpatient and outpatient services billed by a facility and also typically exclude ancillary services, such as laboratory, imaging, DME, injectibles, medical and surgical supplies, transportation, pharmaceuticals, etc. One modification of the “professional services” to be used in this attribution approach that has been proposed by some is the use of information on the “ordering” provider, for a pharmacy prescription or diagnostic test. If available, this information could be used to extend the concept of services “rendered” by a professional provider. Some ETG users have assigned total costs for a cluster to the cluster provider as a way to extend this type of concept for attribution – the argument being that cluster ownership may suggest that the physician played an important role in the decisions to perform the ancillary services grouped to the cluster.

Using professional service costs for attribution involves the following steps:

-- For each qualified episode, sum the costs of all professional services grouped to that episode, by physician.

-- Identify those physicians with episode costs (if any) that are also included in the peer group being measured. Disregard any episodes without one or more physicians for that peer group;

-- Identify the peer group physician with the greatest amount of total costs. If two or more peers are found to have the most costs, apply an appropriate “tie-breaker” to determine the winning physician (discussed below).

-- For each physician, compute their professional costs, as a percentage of costs for all clinicians for the episode and also as a percentage of all costs for all physicians in the peer group. These amounts can be used to compare against percentage thresholds to determine the degree to which a provider is “dominant” within an episode (discussed below).

After application of appropriate tie-breaker and threshold comparisons, the peer group physician with greatest amount of professional costs, is the responsible provider for that episode for that peer group.

Approach 2 - Physician Episode Attribution using Episode Clusters. This attribution approach identifies the responsible physician for an episode as that provider in the peer group owning the greatest number of “clusters” within the episode. As described in the attachment for S.2, other than the individual service, the cluster is the basic unit of an ETG episode. Episode clusters are created using anchor records. Anchor records represent services provided by a clinician engaging in the direct evaluation, management or treatment of a patient. Office visits, therapies, and surgical procedures are examples. An anchor record indicates that a clinician has evaluated a patient’s illness and has decided on the types of services required to further identify and treat the patient’s condition. ETG links an anchor record with related services to form a cluster. Clinically homogeneous clusters are then combined to create episodes of care.

The clinical nature of an episode cluster makes it a natural candidate as an activity measure for episode attribution. In particular, the anchor records that define a cluster represent those types of service where decisions regarding the course of care and management of an episode take place. An additional benefit of episode clusters is that an anchor record service for a cluster can reside in another episode of care, but the cluster and cluster provider can still be identified for the episode of interest.

Using episode clusters for attribution involves the following steps:

-- For each qualified episode, sum the number of clusters “owned” by each clinician. The detail output file from ETG can be used for this purpose. For each service that can be assigned to an episode, the detail file identifies a unique cluster number and a cluster provider ID (same as the servicing provider ID for the cluster anchor record). Using this file, the unique cluster providers for an episode and the number of clusters each provider owns can be identified.

-- Identify those physicians with episode clusters (if any) that are also included in the peer group being measured. Disregard any episodes without one or more cluster providers from that peer group;

-- Identify the peer group physician with the greatest number of episode clusters. If two or more providers are found to have the most clusters, apply an appropriate “tie-breaker” to determine the winning provider (discussed below).

-- For each peer group physician, compute their number of clusters, as a percentage of clusters for all clinicians for the episode and also as a percentage of all clusters for all physicians in that peer group. These amounts can be used to compare against percentage thresholds to determine the degree to which a provider is “dominant” within an episode

(discussed below).

After application of appropriate tie-breaker and threshold comparisons, the peer group physician with greatest number of clusters is the responsible provider for that peer group.

Approach 3 - Physician Episode Attribution using Non-Acute Evaluation and Management (E/M) Visits. This attribution approach identifies the responsible physician for an episode as that physician providing the greatest number of non-acute E/M visits within the episode.

Non-Acute E/M services include office visits and consultations and other E/M services that occur outside of an acute setting where a provider is managing patients and their care. For example, these services exclude initial and subsequent inpatient visits, inpatient consultations, ER visits and critical care visits. It includes office visits and consults, home visits, SNF visits, psychiatric evaluations and therapy and preventive services.

The clinical nature of these services makes them a logical candidate as an activity measure for episode attribution. In particular, these services represent encounters where decisions regarding the course of care and management of an episode take place. This subset of services will be narrower than that described by episode clusters.

Using non-acute E/M visits for attribution involves the following steps:

-- For each qualified episode, sum the number of non-acute E/M visits (visits) rendered by each clinician during the episode.

-- Identify those physicians with these visits (if any) that are also included in the peer group being measured. Disregard any episodes without one or more visit providers from that peer group;

-- Identify the peer group physician with the greatest number of visits. If two or more providers are found to have the most visits, apply an appropriate "tie-breaker" to determine the winning provider (discussed below).

-- For each peer group physician, compute their number of visits, as a percentage of visits for all clinicians for the episode and also as a percentage of all visits for peer group physicians. These amounts can then be used to compare against percentage thresholds to determine the degree to which a provider is "dominant" within an episode (discussed below).

After application of appropriate tie-breaker and threshold comparisons, the peer group physician with greatest number of visits is the responsible provider for that episode for that peer group.

Approach 4 - Physician Episode Attribution using a Primary Care, Population-based Approach. As noted above, a "population" or "panel" based approach is sometimes used when measuring performance for peer groups comprised of primary care physicians. In particular, this approach is often considered where the PCPs are performing a gatekeeper function for a population of members. In this case, responsibility for a member's qualified episodes of care may be attributed to the member's PCP — whether or not the PCP provided any of the services for that member during those episodes.

This approach requires two important steps:

-- Identification of a PCP for each member. This identification can often be obtained from the member's eligibility record which can include a notation of their assigned PCP for a period of time. Alternatively, a PCP can be "imputed" for a member based on that primary care specialist providing the greatest number of services or service costs for selected primary care. When imputing, the list of eligible providers is typically limited to those physicians involved in primary care. Using either approach, a member is linked to a PCP for a defined period of time.

-- For each qualified episode, identify the patient's assigned PCP during the episode period. Most users of this approach will select the member's assigned PCP at the beginning or ending date of the episode (episode begin and end date is available as part of the standard ETG output).

Using this approach, the peer group physician would be assigned all qualified episodes where they were determined to be the patient's PCP during the defined time period.

Physician Episode Attribution – Other Issues. Some general issues around episode attribution remain. The first involves tie-breakers. When using activity-based attribution for some episodes, two or more providers may have the same amount of costs, clusters or visits. In this case, a tie-breaker is often applied to determine the responsible physician for the episode. Useful candidates for this purpose are the alternative activity measures described here. For example, if two physicians own the same number of clusters within an episode, the physician with the greatest amount of professional services costs could be selected. If a tie still remains, the physician with the greatest number of visits could be chosen, and so on.

A second issue involves setting appropriate thresholds to determine sufficient activity. As noted above, most activity-based attribution approaches involve some screening of the winning provider to ensure that they owned sufficient

activity relative to their peers and to other providers during the course of the episode. This is typically done using two threshold comparisons – a provider’s percentage of the total activity of peers and a provider’s percentage of the total activity described by all clinicians for the episode. This percentage is then compared to a predefined threshold(s). For the physician with the greatest activity, if their percentages exceed both of these thresholds, they are determined to be responsible for the episode.

As an example, for an episode with 10 clusters, Dr. Jones is responsible for 2 of the 10 clusters and 8 other physicians are responsible for 1 cluster each. Even though Dr. Jones has the most clusters, he still may not be assigned the episode because his involvement was very small.

Most users set these thresholds at 25 or 30 percent. For example, the winning provider must own 25% or more of all of the episode clusters owned by peers and 25% or more of all episode clusters owned by all clinicians.

As a final point, it is useful to summarize the issues around allowing an episode to be attributed to multiple providers.

As noted above, many ETG users who employ episode results to support physician measurement perform attribution separately for each specialty peer group of interest, including primary care. In doing this, they select a single winner, if possible, for each peer group in which the episode is relevant, but allow providers in different peer groups to be assigned the same episode, if attribution requirements are met.

In this way, it is theoretically possible to assign more than one physician to an episode if each peer group is considered separately. Users typically do not assign two physicians from the same peer group to the same episode.

To support multiple attribution across peer groups, users would repeat the attribution step selected from above separately for each peer group. Those physicians both meeting the dominant provider status for their peer group and also exceeding the threshold requirements could be responsible for the episode.

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 groups define the group of physicians being compared. For example, a common practice in physician episode measurement is to assess the actual costs for those episodes attributed to an individual physician or practice and compare actual costs to peer results, risk adjusted to support more valid comparisons. The peer values use in these comparisons will be influenced by the selection of providers included in the peer group.

In defining a peer group for cost of care measurement, most organizations will include physicians from the same specialty or area of expertise. For organizations with a network covering broad geographic area, some distinction by provider geography can also be used. Internal medicine, cardiology, or general surgery within a certain geographic area are examples of a peer group. Although not directly related to defining a group of providers as peers, many organizations provide separate measurements by line of business, separating results and peer comparisons by commercial, Medicare and Medicaid products.

S11.3. Level of Analysis:

Clinician : Group/Practice

Clinician : Individual

Clinician : Team

Facility

Health Plan

Integrated Delivery System

Population : Community

Population : County or City

Population : National

Population : Regional

Population : State

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 : Outlier episodes – as a guideline, low outlier cost episodes should be excluded from resource use measurement. High outlier cost episodes should be included, but all costs truncated at the high outlier cost threshold used for the episode (a technique called “winsorization”). Where costs by type of service are used in measurement,

individual service costs can be pro-rated to reflect the truncated total cost for a high cost outlier episode.

S11.5.Detail sample size requirements

Detail the sample size requirement including rules associated with the type of measure

Guidelines : The choice of sample size is less important using techniques that include statistical methods that find only statistically significant difference. If your choice of sample size is low, you will not find many cases that are statistically significantly different. A sample size of 30 is chosen because this is when the normal distribution is a good approximation of the student's t distribution. However, the choice of sample size is less critical when using tests of statistical significance.

S11.6.Define benchmarking or comparative estimates

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

Guidelines : The response to section S10.1 includes examples on how to compare the results for a physician with that of their peers or with external best practice benchmarks. As a guideline, in making comparative estimates, the following considerations should be made:

- As described in S10.1, comparative results should be risk adjusted to support more valid comparisons;
- Differences in fee schedules and contracts – for some comparisons using cost of care, differences between actual practice and the benchmark can be influenced by different unit pricing assumptions. In these cases standard pricing or general adjustments to cost levels can be made; and
- Practice styles and service utilization can differ between geographic areas and also between physicians in different specialties. Although comparisons across areas and specialties can provide insights, proper care should be taken in interpreting and communicating results.

S12.Type of Score:

Continuous variable
Count
Rate/Proportion
Ratio

If available, please provide a sample report:

[S12_sample_score_report_EPI-634413291076418837.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 measures described in this submission include continuous cost measures, counts of utilization, rates and proportions (per episode), and the ratio of observed to expected results, based on risk adjusted comparisons.

For the continuous cost per episode measures (also a rate), an increase in costs can be interpreted as an increase in the resources used to diagnose, manage and treat the episodes in question. This score provides a representation of the weighted utilization expended, where the weights are based on the cost assigned to each individual service.

For the counts of utilization measures per 1,000 episodes (also a rate), an increase in utilization can be interpreted as an increase in the resources used to diagnose, manage and treat the episodes in question. This score provides a representation of un-weighted utilization. Counts of utilization measures are most useful when the services being aggregated are similar (e.g., inpatient admits, E&M visits, MRI services).

The risk adjusted observed to expected cost or utilization ratio (O/E ratio) includes three important steps:

- Compute the observed experience for the provider being measured, across all episodes to be included in the comparison;
- Compute the experience for peers or a best practice benchmark. Compute this experience at the level of the risk adjustment, in this case ETG base condition and severity level. For a peers benchmark, average cost per episode across all peers for the ETG base condition and episode level can be computed;
- Compare the observed experience with the risk adjusted peers or benchmark experience – often called the “expected” result. This expected result is adjusted to reflect both the peers/benchmark levels of performance and also the provider’s own case mix of episodes by condition and level of severity. The ratio of observed to expected results can be termed the relative cost ratio and is a risk adjusted measure.

The O/E ratio (relative resource use ratio) can be interpreted based on its magnitude and relationship to a peer average or other guidelines. A relative cost ratio less than 1.00 indicates that the observed resource use per episode for a provider is less than his peers. A relative cost ratio greater than 1.00 indicates that the observed resource use per episode for a provider is greater than his risk adjusted peers.

S12.2. Detail Score Estimation

Detail steps to estimate measure score.

The measures described in this submission include continuous cost measures, counts of utilization, rates and proportions (per episode), and the ratio of observed to expected results, based on risk adjusted comparisons. The continuous cost measures, counts of utilization, and rates per episode are described in detail in S9.5. The details involved in computing the O/E ratio measure is provided in S10.1.

S12.3. Describe discriminating results approach

Detail methods for discriminating differences (reporting with descriptive statistics--e.g., distribution, confidence intervals)

In all of these measures we end up with an O/E ratio for a provider. In order to determine the statistical accuracy of this measure we start by measuring the variance of this metric:

$\text{Var}(\text{O/E})$

The Variance of this metric has been estimated by the following expression in a number of journal articles :

$\text{Var}(\text{O/E}) = (\text{Sum}(\text{Var}(\text{O}_i)) / [\text{Sum}(\text{E}_i)]^2$

Where $\text{Var}(\text{O}_i)$ is the variance for each of the physician’s episodes across all episodes in it’s statistical unit for the peer group.

Then the standard error (SE) for this measurement is $\text{Sqrt}(\text{Var}(\text{O/E}))$.

Finally, a 95% confidence interval could be calculated by:

$(\text{O/E} - 1.96 * \text{SE}, \text{O/E} + 1.96 * \text{SE})$

Alternatively, a 90% confidence interval could be calculated by: $(\text{O/E} - 1.64 * \text{SE}, \text{O/E} + 1.64 * \text{SE})$

Adams et al. BMC Health Services Research 2010, 10:57 <http://www.biomedcentral.com/1472-6963/10/57>

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.

TESTING ATTACHMENT (5MB or less) or URL:

If needed, attach supplemental documentation (Save file as: SA_Reliability_VValidity 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_VValidity Testing_Pneumonia.xls

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)

Data used to support validity testing is based upon a National Commercial member health care services benchmark database representing more than 25 million covered lives for calendar year 2009. Various permutations of the 25 million unique members are pulled to support testing initiatives, for example:

- 4 million member sample used for face validity evaluation of ETG processing
- 7 million member sample used for reliability evaluation of ETG processing and associated Resource Utilization measures
- 75,000 member sample, with manipulated data for content validation testing of ETG processing and associated Resource Utilization measures

SA1.2. Analytic Methods

(Describe method of reliability testing and rationale)

Reliability refers to the consistency of a measure. A measure is considered reliable when the same result is produced repeatedly. Reliability of ETGs and Resource Utilization Measures are judged based upon an internal consistency reliability approach. The first level of internal consistency reliability focuses on high-level parallel processing tests and regressions performed by internal Quality Assurance (QA) teams. This level focuses on assessment of results compared to a baseline set of expected results developed based upon the experience of a benchmark database of member and health care services covering more than 25 million lives as described in SA1.1.

The second level of internal consistency reliability involves detailed parallel processing comparisons between ETG and Resource Utilization Measure software and SAS-based software prototypes. Software prototypes are developed and maintained by analysts familiar with the detailed methodology of the measures for the purpose of Content Validation (CV). This form of parallel reliability testing requires that the results of both the software and prototype match exactly and are executing the logic in accordance with methodological specifications. Observed differences in the output are researched and resolved prior to releasing the software for use. Multiple parallel processing comparisons are performed to assure that the software is producing reliable results using a variety of processing configuration options and data input scenarios.

SA1.3. Testing Results

(reliability statistics, assessment of adequacy in the context of norms for the test conducted)

The extensive testing of ETG produces volumes of results across the test cases and other concepts described above. In terms of validity and assessing the reliability of the implementation, testing of the measurement software with the parallel SAS prototype involves iterations until a high degree of matching of results is observed (over 99.9%). The statistic used in this testing is the exact match of the grouping of records and assignment of resource measures. The difference in the result for each measure between the methodology and prototype is calculated and differences equal to zero are considered an exact match.

In terms of testing of measures across organizations, the following results provide examples of consistency for the submitted measures. These data were not standard priced, so some observed variation is the result of differences in fee schedules and contracts between the organizations. A table, "Reliability Across HCOs" is included in the attachment for SA (SA_Reliability_VValidity Testing). The table shows measures of resource use for nine healthcare organizations (HCOs) (columns) with a separate comparison provided for selected resource use measures included with this submission. (The 7 million member sample from 9 health care organizations used for reliability assessment described in SA1.1). The results include combined findings across all severity levels for the base condition, with results risk adjusted to reflect the same mix of episodes by severity level across each organization. Separate results are shown for

2a2

H ☐
M ☐
L ☐
I ☐

relevant peer groupings (e.g., internal medicine, cardiology). These peer group results are based on episode attributed to each provider, with the estimates describing the peer level findings across all physicians and episodes included in the measurement. As shown, the results suggest a level of consistency across health plans implying reliability in both the measure specification and how it can be applied to different organizations.

A further assessment of reliability and face validity can be made using measure results attributed to physicians in different specialties. The tables, “Results Across PeerGrps, Cost” and “Results Across PeerGrps, Utils” included in the attachment for SA (SA_Reliability_VValidity Testing). Provide a comparison of the cost and use per episode for episodes attributed to different specialties. The tables also show results by episode severity level, supporting an assessment of how cost and use measure results vary as severity level increases. The results also show a strong relationship between episode severity and resource use.

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

As noted in SA1.3, the findings on reliability and validity suggest the measures could be applied in a consistent way, the results matched well to clinical expectations, and the results from the measurement software were consistent with those produced by a parallel process using prototype implementation of the methodologies.

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)

Different samples of data are used in testing ETG and the Resource Use Measures described in this submission. The general source of information is the Ingenix National health care services benchmark database. This database describes enrollment, medical and pharmacy services, and providers for a population of more than 25 million covered lives. The data used in the testing described in this submission was primarily for commercial non-elderly individuals and covered the years 2006 thru 2010, depending on the test. The primary test databases used to support the tests described in the SA section are as follows:

- 4 million member sample used for validity and reliability of the ETG methodology and the software used for ETG processing;
- 250,000 member sample, with manipulated data for content validation testing of the post-ETG processing associated with Resource Utilization measures (measures described in S9.5);
- 7 million member sample from 9 health care organizations used for reliability assessment (consistency across data sources). This sample was also used to support the empirical estimates for the Importance section of this submission (IM1)

SA2.2.Analytic Method

(Describe method of validity testing and rationale; if face validity, describe systematic assessment)

Also, please see our responses to SA1 which relate to both reliability and validity.

Validity determines if the output of the measure is accurate. The measure must be valid in order for the results to be accurately applied and interpreted. Validity of a measure is not determined by a single statistic, but by evaluating the complete result of the measures and demonstrating the relationship between the result and the intended purpose of the measure. Validity of ETGs and Resource Use Utilization Measures are judged based upon both content validity and face validity.

Content validation testing involves detailed parallel processing comparisons between ETG and Resource Use Utilization Measure software and SAS-based software prototypes. Software prototypes are developed and maintained by analysts familiar with the detailed methodology of the measures for the purpose of Content Validation (CV). This form of parallel testing requires that the results of both the software and prototype match exactly and are executing the logic in accordance with methodological specifications. Observed differences in the output are researched and resolved prior to releasing the software for use. Multiple parallel processing comparisons are performed to assure that the software is producing valid results using a variety of processing configuration options and data input scenarios.

The face validity approach assesses if the measure result is reasonable and functioning according to expectations. This

2b2

H ☐
M ☐
L ☐
I ☐

form of validation is most typically performed when modifications to the methodology intentionally change the result of the measure. When this occurs a pre- and post-modification parallel run is created and changes in the measure output are validated for accuracy at face value. Episodes are evaluated for validity in terms of distribution of ETGs, Episode Types, Record Types, Outlier Status and Type of Service. Resource Utilization Measures are evaluated for validity in terms of measure Cost per Episode by Peer Group as well as overall evaluation of the utilization measures by Peer Group.

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)

Please see our responses to SA1 which relate to both reliability and validity.

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

Please see our responses to SA1 which relate to both reliability and validity.

SA3. Testing for Measure Exclusions

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

ETG includes logic to identify high or low cost outliers at the episode level. In particular, ETG has a comprehensive method for identifying outlier episodes where the resource cost is high or low enough relative to norms for the clinical condition to distort the results. A table of thresholds, or “trim points”, is used to describe levels of costs considered extremely high or low relative to the norm. Specific trim points are defined for each base condition (e.g., pneumonia) and also for each level of severity and the presence of surgical treatments. These values have been determined using a benchmark database describing the experience of more than 25 million covered lives. Note that severity of illness and treatment indicators are assigned as described in the general methodology paper on ETG included in the response to S2. Low and high outlier episodes are noted by ETG.

As described in the general methodology paper on ETG (included in the response to S2), ETG considers an episode incomplete if the clean period of the episode overlaps with the boundaries of the overall time period being used for measurement (e.g., calendar years 2009 and 2010) or the member’s eligibility start and end dates. Incomplete episodes may have either an unknown start or an unknown finish. ETG clean periods are described in detail in the general methodology paper on ETG (see S2 response). To summarize, clean periods describe the amount time before and after an episode where clinical activity related to the episode is assessed to determine episode completeness. If no relevant clinical activity is observed and the clean period does not overlap with the overall analytic time period begin and end dates or the member’s eligibility begin and end dates, the episode can be considered complete. Different rules are applied to acute and chronic episode conditions to do this. Complete and incomplete episode status and type are noted by ETG.

It is recommended that incomplete episodes be excluded from resource use measurement and comparisons. It is recommended that low outlier cost episodes be excluded from resource use measurement. It is recommended that high outlier cost episodes be included in resource use measurement, but truncated at the high outlier trim point.

In terms of resource use measure construction following ETG grouping, no additional data inclusion or exclusion are applied. Only condition episodes are included in the measurement of episode-based resource use for that condition, including the individual services that ETG groups to those episodes. As noted, it is recommended that incomplete episodes be excluded from resource measurement and outlier episodes be treated as described above.

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)

Different samples of data are used in testing ETG and the Resource Use Measures described in this submission. The general source of information is the Ingenix National health care services benchmark database. This database describes enrollment, medical and pharmacy services, and providers for a population of more than 25 million covered lives. The

2b3

H ☐
M ☐
L ☐
I ☐

data used in the testing described in this submission was primarily for commercial non-elderly individuals and covered the years 2006 thru 2010, depending on the test. The primary test databases used to support the tests described in the SA section are as follows:

- 4 million member sample used for validity and reliability of the ETG methodology and the software used for ETG processing;
- 250,000 member sample, with manipulated data for content validation testing of the post-ETG processing associated with Resource Utilization measures (measures described in S9.5);
- 7 million member sample from 9 health care organizations used for reliability assessment (consistency across data sources). This sample was also used to support the empirical estimates for the Importance section of this submission (IM1)

SA3.3. Analytic Method

(Describe type of analysis and rationale for examining exclusions, including exclusion related to patient preference)

Reliability and testing of exclusions for ETGs and Resource Utilization Measures are judged based upon an internal consistency reliability approach. The first level of internal consistency reliability focuses on high-level parallel processing tests and regressions performed by internal Quality Assurance (QA) teams. This level focuses on assessment of results compared to a baseline set of expected results developed based upon the experience of the benchmark described above in SA2.1.

The second level of internal consistency reliability for exclusions involves detailed parallel processing comparisons between ETG and Resource Use Measure software and SAS-based software prototypes. Software prototypes are developed and maintained by analysts familiar with the detailed methodology of the measures for the purpose of Content Validation (CV). This form of parallel reliability testing requires that the results of both the software and prototype match exactly and are executing the logic in accordance with methodological specifications. Observed differences in the output are researched and resolved prior to releasing the software for use. Multiple parallel processing comparisons are performed to assure that the software is producing reliable results using a variety of processing configuration options and data input scenarios.

As an example, the text below provides the Table of Contents for an ETG testing plan for ETG Version 7.0. The plan includes processes around data used, test cases created, comparison of software results with those produced by a SAS prototype (to determine matching across parallel implementations of the methodology), and a review by clinical analysts to assess face validity. A similar testing approach is used for the resource use measures that are processed following ETG grouping. Note that steps 2.4 and 2.5 relate to exclusions around episode completeness and outlier status.

ETG TEST PLAN DOCUMENT – EXAMPLE TABLE OF CONTENTS

SECTION 1—OVERVIEW

- 1.1 PURPOSE OF TEST PLAN DOCUMENT
- 1.2 TESTING APPROACH AND DELIVERABLES
- 1.3 SCOPE OF TESTING
- 1.4 DATA
- 1.5 ETG GROUPER

SECTION 2—BENCHMARK TEST CASES

- 2.1 ACCOUNTING OF GROUPED VS. UNGROUPED RECORDS
- 2.2 DISTRIBUTION BY ETG
- 2.3 DISTRIBUTION BY MPC
- 2.4 DISTRIBUTION BY EPISODE COMPLETENESS
- 2.5 DISTRIBUTION BY OUTLIERS
- 2.6 EPISODE AGE/GENDER PROFILE

SECTION 3—FEATURE-RELATED TEST CASES

- 3.1 COMPARISON OF SOFTWARE TO PROTOTYPE
- 3.2 SEVERITY ADJUSTMENT
- 3.3 COMPLICATIONS
- 3.4 COMORBIDITIES
- 3.5 TREATMENT INDICATORS
- 3.6 EPISODE INDICATORS

SECTION 4—REVISION HISTORY

Finally, the results are applied to the healthcare data of different organizations to assess both the ability of the organization's data to support the measurements and also the consistency of results across the organizations. This assessment of reliability also provides evidence that the measures are being applied in a consistent and valid way.

SA3.4. Results

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

See Attachment SA_Reliability_Velocity Testing for a comparison of episode outlier and completion results across sources of data from ETG processing.

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

As noted in SA1.3, the findings on reliability and validity suggest the measures could be applied in a consistent way, the results matched well to clinical expectations, and the results from the measurement software were consistent with those produced by a parallel process using prototype implementation of the methodologies. This statement applies to all methodologies involved, including exclusions.

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, <i>Scientific Acceptability of Measure Properties</i> , met? Rationale:	Y <input type="checkbox"/> N <input type="checkbox"/>
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
<p>Meaningful, Understandable, and Useful Information</p> <p>U1. Current Use:</p> <p>Internal quality improvement Payment Public reporting (disclosure to performance results to the public at large) Quality improvement with external benchmarking</p> <p>U1.1. Use in Public Reporting Initiative Use in Public Reporting. <i>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)</i></p> <p>Several users of ETGs and Resource Use Measures rely on the analysis to support Public Reporting initiatives. Examples include:</p> <ul style="list-style-type: none"> -- Health Care Organization #1: Measuring Provider Efficiency -- HCO #1 ranks providers based on efficiency by ETG using a single provider ETG overview. Using COGNOS reporting capabilities the organization is able to drill down into procedure and drug level comparisons. -- Health Care Organization #2: Corporate Wellness Programs -- HCO #2 uses ETG output to analyze utilization patterns and identify potential diseases and populations to target for intervention. ERGs are used to adjust the average and comparison population expenditures and Specialty profiles are created using both ETG and ERG results. ERG scores are used to identify patients who could be potential high utilizers. -- Health Care Organization #3: Physician Profiling and Clinical Benchmarking -- HCO #3 has embarked upon an initiative to use ETG information for clinical reporting and benchmarking. ERG output complements the ETG information for underwriting and physician profiling programs as well. -- Health Care Organization #4: Provider Specialty Profiling and Predictive Modeling -- HCO #4 utilizes Resource Use Measures and ETG to identify variations in practice patterns, measure performance and examine utilization and disease management. The primary focus is on high cost specialties and ETGs are used to identify the top 5 conditions to support specialty profiles and cost comparisons and drill downs. ERG scores are used to risk adjust PCP profiles to adjust for patient severity. <p>Please note that Health Care Organization names were not provided to protect the confidentiality of our users. HCO names for reference purposes are available upon request.</p> <p>U1.2. Use in QI <i>(If used in improvement programs, provide name of program(s), locations, Web page URL(s)).</i></p> <p>Examples of ETGs and Resource Use Measures in action within health care industry quality improvement initiatives include:</p> <ul style="list-style-type: none"> -- Health Care Organization #5: Internal Quality Improvement – Disease Management -- HCO #5 utilizes 30 months of medical and pharmacy data totaling more than 17 million claim lines to support identification of member risk and stratification of members for care management teams. ETG and ERG groupers are embedded within their claims datamart with other sources of data and support the identification of clinical care gaps and impactable dollars for quality improvement. -- Health Care Organization #6: Employer Group Utilization Reports to Identify Provider Variance -- HCO #6 generates Employer or Account Group Utilization Reports which includes a global view of ETGs for the 	<p>3a</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/></p>

<p>population. These reports are used to identify the top 5 ETGs where variance is the greatest to target specific procedures for a particular ETG in order to improve quality for the Employer group.</p> <p>-- Health Care Organization #7: Cesarean Section Study</p> <p>-- HCO #7 conducted a study on Cesarean Section, Infertility and multiple births using ETGs. Providers with high rates of Cesarean Section were identified and compared based upon severity indices. The study determined that multiple births were a significant contributor to a market's cost and procedure variances. The study further identified infertility treatment specialists who need improvement based upon the comparison to their peers of best practices and procedures.</p> <p>Please note that Health Care Organization names were not provided to protect the confidentiality of our users. HCO names for reference purposes are available upon request.</p> <p>U1.3. Use for other Accountability Functions (payment, certification, accreditation) <i>(If used in a public accountability program, provide name of program(s), locations, Web page URL(s)).</i></p> <p>Other examples of industry use of ETGs and Resource Use Measures include Provider Pay for Excellence programs and Member Cost Analysis Tools. Specific examples include:</p> <p>-- Health Care Organization #8: Provider Analytics Team</p> <p>-- HCO #8 leverages the power of ETGs and Resource Use Measures to support their internal Provider Analytics team. This team manages the Provider Profiling program to support the Medical Directors' high-level physician review and network physician meetings as well as bi-annual provider profiling reports. In addition to provider profiling the Provider Analytics team uses ETG and Resource Use Measures to Impute PCP information to identify gaps in care, support physician group award programs and Patient Centered Medical Home projects.</p> <p>-- Health Care Organization #9: Member Cost Analysis Tools</p> <p>-- HCO #9 has created a patient website with cost calculation tools to provide detailed treatment costs for the patient based upon ETG analysis. The website includes tips on how to reduce costs as well as a pharmacy co-pay calculator. Users may access median cost reports for an ETG as well as cost ranges for procedures based upon CPT codes, pharmaceuticals and office visits. The website also provides comparison data for providers based upon performance indices.</p> <p>Please note that Health Care Organization names were not provided to protect the confidentiality of our users. HCO names for reference purposes are available upon request</p>	
<p>U2. Testing of Interpretability <i>(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).</i></p> <p>U2.1. If understanding or usefulness was demonstrated <i>(e.g., through systematic feedback from users, focus group, cognitive testing, analysis of quality improvement initiatives) describe the data, methods, and results.</i></p> <p>The assessment of the usability of the results from ETG-based measures of resource use is primarily from two entities: the ETG Medical Advisory Board and the Ingenix User Forums around these measures. The Medical Advisory Board is comprised of medical directors from healthcare organizations that employ episode based measures to assess resource use. Input and feedback from these clinicians inform both the ETG methodology itself and also how it is used in creating and sharing provider measurement results. The Ingenix User Forums include technical experts from organizations that use ETG. Similar to the Medical Advisory Board, input and feedback from this group informs the ETG methodology, but primarily is focused on how ETG results are used to create and share provider measurement results.</p>	<p>3b</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> NA <input type="checkbox"/></p>
<p>U2.2. Resource use data and result can be decomposed for transparency and understanding.</p> <p><i>Refer to items S11 -S12.3.</i></p>	<p>3c</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/></p>
<p>U3. If there are similar or related measures (either same measure focus or target population)</p>	

<p>measures (both the same measure focus and same target population), list the NQF # and title of all related and/or similar measures.</p> <p>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?</p> <p>U3.2. If the measure specifications are not completely harmonized identify the differences, rationale, and impact on interpretability and data collection burden. <i>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.)</i></p>	<p>3d</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/> NA <input type="checkbox"/></p>
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Usability</i>?</p>	
<p>Steering Committee: Overall, to what extent was the criterion, <i>Usability</i>, met? Rationale:</p>	<p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/></p>
<p>FEASIBILITY</p>	
<p>Extent to which the required data are readily available, retrievable without undue burden, and can be implemented for performance measurement.</p>	<p>Eval Rating</p>
<p>F1. Data Elements Generated as Byproduct of Care Processes <i>How are the data elements needed to compute measure scores generated? Data used in the measure are:</i></p> <p>Generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)</p>	<p>4a</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/></p>
<p>F2. Electronic Sources <i>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)</i></p> <p>ALL data elements in electronic claims</p> <p>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.</p>	<p>4b</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/></p>
<p>F3. Susceptibility to Inaccuracies, Errors, or Unintended Consequences <i>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.</i></p> <p>The main source of inaccuracies relate to small sample size. There are lower limits on the number of episodes for a given provider or specialty that are allowed for inclusion in the analysis. Sample sizes that are determined to be too</p>	<p>4c</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/></p>

<p>small are eliminated from the analysis.</p> <p>These situations will occur infrequently, as the sample sizes that are customarily dealt with are very large. A methodology for applying statistical techniques to determine confidence intervals of the results has been created and can be applied to gauge the accuracy of the analysis. In addition, sample size is less of an issue when multiple episode types are combined for a single metric.</p> <p>In some cases, there are physicians that are "ultra" specialized that may not have a reasonably sized peer group for comparison. Sub-specialties like hepatology, or muscular dystrophy specialists may fall into this category.) A second source of potential inaccuracies relate to the validity and completeness of the administrative data available to support the measurement. As described in S6.1, a careful evaluation of the data to be used to support the measurement is required and actions taken to address identified issues.</p>	
<p>F4. Data Collection Strategy <i>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).</i></p> <p>The measure is in use beyond internal QI. Please see the section on Usability.</p>	<p>4d</p> <p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/> I <input type="checkbox"/></p>
<p>TAP/Workgroup: What are the strengths and weaknesses in relation to the subcriteria for <i>Feasibility</i>?</p>	
<p>Steering Committee: Overall, to what extent was the criterion, <i>Feasibility</i>, met? Rationale:</p>	<p>H <input type="checkbox"/> M <input type="checkbox"/> L <input type="checkbox"/></p>
<p>RECOMMENDATION</p>	
<p>Steering Committee: Do you recommend for endorsement? Comments:</p>	<p>Y <input type="checkbox"/> N <input type="checkbox"/> A <input type="checkbox"/></p>
<p>CONTACT INFORMATION</p>	
<p>Co.1 Measure Steward (Intellectual Property Owner)</p> <p>Co.1 Organization</p> <p>Ingenix, 950 Winter Street, Suite 3800, Waltham, Massachusetts, 02154</p> <p>Co.2 Point of Contact</p> <p>Jen, Pearse, Jennifer_J_Pearse@ingenix.com, 781-419-8628-</p>	
<p>Measure Developer If different from Measure Steward</p> <p>Co.3 Organization</p> <p>Ingenix, 950 Winter Street, Suite 3800, Waltham, Massachusetts, 02154</p> <p>Co.4 Point of Contact</p> <p>Dan, Dunn, daniel.dunn@ingenixconsulting.com, 781-419-8425-</p>	
<p>Co.5 Submitter If different from Measure Steward POC</p>	

Jen, Pearse, Jennifer_J_Pearse@ingenix.com, 781-419-8628-, Ingenix

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

ADDITIONAL INFORMATION

Workgroup/Expert Panel involved in measure development

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

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released:

Ad.3 Month and Year of most recent revision:

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

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

Ad.6 Copyright statement:

Information submitted is confidential/Proprietary to Ingenix, copyright 2011

Ad.7 Disclaimers:

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

03/31/2011

ETG METHODS DOCUMENT

Building Episodes with Episode Treatment Groups (ETG): General Methodology and Application for Pneumonia

This document provides an overview of the Ingenix Episode Treatment Groups (ETG) methodology and its application for creating pneumonia episodes of care. ETG groups individual medical and pharmacy services to unique episodes of care defining a condition for a patient and is used extensively to support episode-based measurement of cost of care. The first section of this document describes the general approach used by ETG. The second section beginning on page 12 summarizes methods for pneumonia.

I. Episode Treatment Groups (ETG) Construction Logic

ETG is an episode grouping methodology that identifies a unique clinical condition for a patient and the services involved in diagnosing, managing and treating that condition. ETG organizes routinely-collected professional, inpatient, outpatient and ancillary services, including pharmaceutical services, into episodes of care. ETG evaluates each claim service record with respect to provider type, procedure and diagnoses codes and other information to assign the record to an appropriate episode. In doing this, all conditions and episodes are considered for a patient, including concurrently occurring conditions.

ETG covers the breadth of clinical medicine. Examples of ETG based conditions include diabetes, asthma and chronic sinusitis. Each episode is further assigned a condition-specific severity level, supporting case-mix adjusted comparisons within and across conditions.

ETG uses as input data information from administrative medical and pharmacy claim service records and encounters describing the individual services provided to a patient. ETG also uses information describing each patient, including age and gender and time enrolled with a health plan or other organization.

The Episode Building Process

The ETG episode building process has four important steps:

1. Identify Records; Assign Record Type and Anchor Records, Classify Diagnoses and Procedures
2. Build Episodes from Anchor Records
3. Group Non-Anchor Records to Episodes
4. Finalize the Episodes (identify comorbidities and complicating factors, and assign episode severity)

Step 1: Identify Records; Assign Record Type and Anchor Records, Classify Diagnoses and Procedures

Assign record type to each Service

In building an episode the first step involves assigning a Record Type to each service record. The Record Type assigned to a record is determined by the Provider Type, Procedure Code and/or Revenue Code Service, and National Drug Code (NDC) (if any), on the record. Provider Type values are based on the mapping of individual provider specialties to one of three values recognized by ETG: Clinician, Facility and Other. The Provider Type values and their definitions are as follows:

Provider Type	Definition
Clinician	Providers who make diagnoses and recommend treatment
Facility	Acute and long term care providers such as short-term hospitals, skilled nursing facilities, and psychiatric or chemical dependency facilities
Other/Non-Clinician	All other healthcare providers

Identify Anchor Records

Service records containing a NDC code are assigned a Pharmacy Record Type. For other services, ETG assigns one of the following Record Types to the service record using Provider Type and the procedure/revenue code and also determines if that Record Type can anchor (begin or continue) an episode. The following table describes the Record Type and Anchor relationship:

Record Type	Record Type Value	Anchor or Non-Anchor
Management	A record submitted by a clinician for services related to the evaluation of a patient's condition.	Anchor
Surgery	A record submitted by a clinician for surgical or related procedures.	Anchor
Ancillary	A record submitted by any provider for laboratory, radiological or similar services.	Non-Anchor
Facility	A record submitted by a treatment facility for room & board services.	Anchor
Pharmacy	A record for a prescription drug service.	Non-Anchor

Most management records contain evaluation and management CPT-4 codes. Surgery records are primarily procedural CPT-4 codes. Facility records are room and board revenue codes billed by a facility (also referred to as a confinement). Pharmacy records are claims containing a NDC or certain HCPCS codes related to the administration of a drug. Record Types of management, surgery and facility are considered *anchor* records. The identification of an anchor record is significant because it indicates that a clinician has evaluated a patient, and has decided on the types of services required to further identify and treat the patient's condition. Non-anchor records describe *ancillary* services that aid in evaluating and treating the patient, such as x-rays and laboratory services.

Assign Diagnoses to Diagnosis Class

The way in which records are grouped to an episode is governed mainly by the diagnosis, revenue, and procedure codes on the service record. Each ICD-9-CM, CPT-4/HCPCS, and revenue code has been mapped to ETG concepts through extensively vetted and continually updated clinical tables. (ICD-9 procedure codes are not used in grouping.)

Diagnosis Codes

The software relies heavily on the diagnosis codes to help identify discrete episodes. The diagnosis identifies the condition being treated, which broadly translates to an ETG. Each diagnosis code is identified with a given diagnosis class. There are three diagnosis classes:

- **Specific:** These are ICD-9 diagnosis codes that indicate a specific disease. This code represents a disease or condition (as opposed to a sign or symptom) and is specific enough to be linked to a single ETG.
- **Non-Specific:** These ICD-9 diagnosis codes represent a disease or condition (as opposed to a sign or symptom), but may not be specific enough to identify a single ETG.
- **Sign and Symptom:** These ICD-9 diagnosis codes represent signs and symptoms of disease as opposed to disease or condition.

The software runs one member at a time and processes the anchor records with a 365-day moving window. The diagnosis codes are grouped in several distinct passes. This is done so that the grouper processes the more specific codes first, leaving the sign & symptom codes until later, when it is more likely that there is a more specific episode for these claims to join.

Each diagnosis code is matched with one or more ETGs through a diagnosis eligibility table. The exception is 'E' codes which are not grouped. Each diagnosis code is further ranked, based on its strength of association with the ETG. The rank values are as follows: low, medium, high and primary. Low, medium, and high represent the strength of the match association. A primary rank describes conditions that define a disease and are the main codes that impact grouping decisions. The grouper first processes the specific and non-specific diagnosis codes so that concrete conditions/diseases are created. It then processes the sign and symptom diagnosis codes in reverse chronological order based on service dates to determine the best episode each of them can group to.

Identify the Clinical Relationship Between Diagnosis Codes and Conditions

Match each diagnosis code with one or more conditions (ETGs) through a diagnosis eligibility table. In addition to mapping diagnosis codes to conditions, each diagnosis code is further ranked, based on its strength of association with a condition. A rank of "primary" or "incidental" is assigned to each diagnosis and condition combination, with a further ranking assigned to incidental relationships:

- Primary: A "primary" diagnosis/condition relationship is assigned in a map where the diagnosis defines that condition. (Note: the word "primary" here is used to describe the relationship between a diagnosis and an episode, it is not used to indicate the position of the diagnosis code on the claim line. The diagnosis in any position on the claim line can have a primary relationship with an ETG). This map is used to identify primary diagnoses for the ETG. Primary diagnosis codes can only be ranked as primary for a single ETG condition.

- Incidental: These diagnosis codes are eligible for a condition but are not classified as primary. These diagnosis codes can be incidental to other conditions. To support the linkage of these diagnosis codes to a final episode, a further ranking is assigned for each condition based on the relative strength of association between the diagnosis and condition. Values of low, medium, or high are assigned for each diagnosis/condition.

Identify Relationships between Procedure Codes and Conditions

In building episodes, the procedure or revenue code can help to identify the ETG to which a particular claim record can be assigned. A given procedure may be valid for several ETGs, though not equally so. A procedure eligibility table therefore ranks the valid ETGs for each procedure to give a better sense of how closely related the service is to each ETG. The ranking options are: Very Low, Low, Medium, and High, with High being the strongest rank.

The following table provides an example of a rhinoplasty surgical procedure and selected ETGs it is eligible for and the rank for each ETG.

ETG	Rank
Trauma to ear/nose/throat	High
Other inflammatory conditions of ear/nose/throat	High
Allergic rhinitis	Medium
Chronic sinusitis	Medium
Trauma of oral cavity	Medium
Open fracture or dislocation - head & face	Medium
Congenital & acquired anomalies of ear/nose/throat	Medium
Closed fracture or dislocation – head & face	Low
Cocaine or amphetamine dependence	Very Low
Other disorders of ear/nose/throat	Very Low

For a record to be eligible to start or join an episode, the diagnosis code and the procedure/revenue code must both be eligible for an ETG. Where an anchor record can be assigned to more than one observed episode for a patient, the record is assigned to an episode according to the best combination of the procedure/revenue code and the diagnosis code.

- ❖ The ETG Online Clinical Knowledge Base application on the Ingenix website (www.ingenix.com/transparency) provides more information about the diagnosis and procedure associations to an ETG.

Identify Relationships Between Pharmacy Services and Conditions

The relationship between pharmacy services and episodes is based on the pharmacy code assigned to the service in a mapping. To support this assessment, the ETG methodology assigns each pharmacy service to a Drug Category Code (DCC). The DCC describes the drug's active ingredients and route of administration. DCCs are then mapped to ETGs and define the relationships between a drug and a condition. Most pharmacy services are defined using NDC procedure codes, however selected pharmacy services with a CPT or HCPCS code are also mapped to a DCC by ETG (e.g., J-codes describing injections).

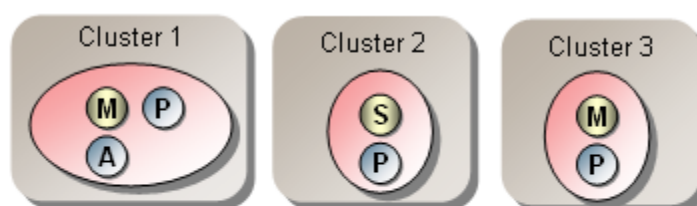
Step 2: Build Episodes from Anchor Records

Only anchor records can start or continue an episode.

Anchor records can do the following:

- Begin a *cluster* that can open a new episode or join an existing episode
 - Extend an episode (time-wise) – providing evidence that the episode has not yet completed
 - Create one or more or *phantom* clusters – when there are multiple diagnosis codes on the same anchor record
 - Determine if episodes incur complications, comorbidities and significant surgery/treatment
-

Each anchor record forms a cluster. A cluster is the basic unit of an episode. Each cluster is comprised of an anchor record and zero, one, or more ancillary and pharmacy records. Each episode consists of one or more clusters. The illustration below demonstrates this concept, showing management (M), ancillary (A) and pharmacy (P) records within clusters.



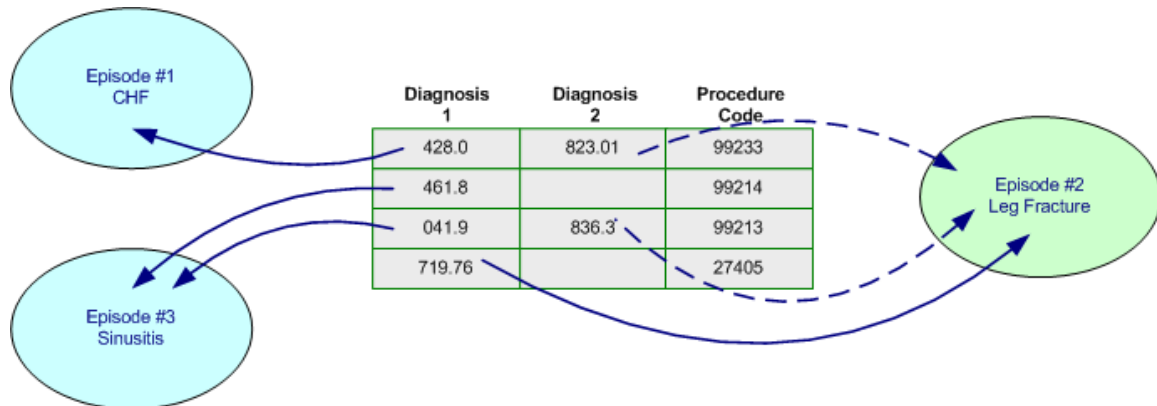
Each cluster has only one anchor record

All records in a cluster have the same cluster number

Clusters: Real and Phantom

Once the anchor record has been assigned to an episode using a diagnosis, the remaining diagnosis codes on the record, if any, are examined. If a remaining diagnosis would more appropriately belong to a different episode than the episode the anchor record is assigned to, the software starts a phantom cluster for a new episode. At this point, phantom clusters are episodes

created that will not have any costs assigned to them. Subsequent service records for a patient will now have available additional episodes for potential grouping, so the software will be able to assign these subsequent services more accurately than it would without using phantoms. This allows the diagnostic information to be utilized fully to identify and track all of the conditions for which the member is being treated, yet still assign records to only one episode. The diagram below provides an illustration. The dotted line indicates a phantom episode was started, a straight line indicates a real episode was started. In the case of diagnosis code 719.76, it joined episode #2 which originated as a phantom episode, thereby converting it to a real episode.



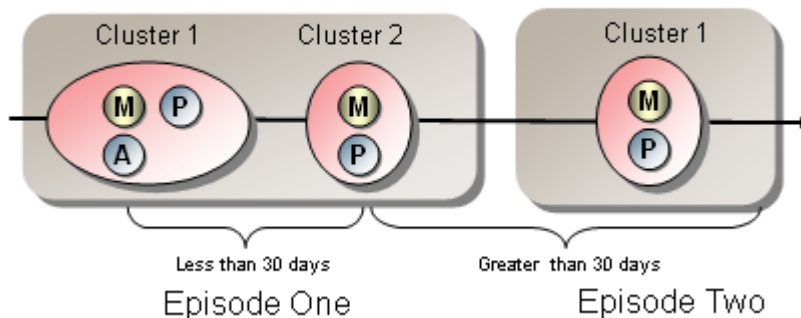
Time Windows: Clean Periods and Member Eligibility

Along with the clinical aspects of starting and grouping records to an episode, the method of episode completion is a crucial feature of ETG. The approach taken for the identification of a complete episode relies on a flexible, rather than a fixed length of time. There are no standard definitions of an episode's chronological length. The episode grouper continues to identify and track all clinical activity for an episode for as long as a condition is actively treated – a concept described as discrete dynamic clean periods. A clean period is defined as the absence of treatment for a specified period of time. Each ETG has its own unique clean period. For an acute condition the concept of a clean period is of most importance. For example, the clean period for Acute Bronchitis is 30 days. Once an episode has started for this ETG, anchor records clinically consistent for acute bronchitis group to this episode until such time as 30 days passes without any corresponding clinically consistent treatment. For Chronic Bronchitis, the clean period is 180 days, consistent with a more chronic illness. In some obvious instances, e.g. benign hypertension or diabetes, there is no clean period. The condition is basically life-long (chronic) and all clinically consistent treatments group to an episode of benign hypertension for as long as data are available.

The clean period window is dynamic in that each new anchor record that joins an episode moves the clean period window by extending the episode's dates. In this way, as long as a condition is consistently treated such that the date of each successive anchor record is less than or equal to the clean period date for the ETG, the episode can last forever.

The following diagram provides an illustration of this concept for an acute condition.

A member has been identified as having Acute Bronchitis.
The Clean Period for this ETG is 30 days.



In this example, two episodes of **Acute Bronchitis** are created.

- Three office visits occurred for the treatment of acute bronchitis (record type M)
- The time frame between the second office visit and the third office visit was greater than 30 days, the clean period of this ETG. Therefore, a second episode was created for this condition

If the example above had been for a chronic condition, such as benign hypertension, all services would be grouped into a single episode since chronic conditions do not necessarily have an end to their clean period. To allow for analysis on chronic conditions, we offer 5 options for users to parse the episode into annual increments:

1. User chooses any month to begin year long episodes
2. Year long episodes will start from the beginning of the grouped data
3. Year long episodes will start from the member's eligible start date
4. Year long episodes will end at the end of the grouped data
5. Year long episodes will end at the member's eligible end date

Step 3: Group Non-Anchor Records

Non-anchor records represent services that are incidental to the direct evaluation, management and treatment of a patient. There are two types of non-anchor records: pharmacy records and ancillary records (such as laboratory tests, x-rays, and the facility component of ambulatory surgery centers services). Each non-anchor record links to only one cluster and eventually becomes part of the episode that the cluster is finally grouped to.

Ancillary records can do the following:

- Join an episode
- Convert a phantom episode into a real episode

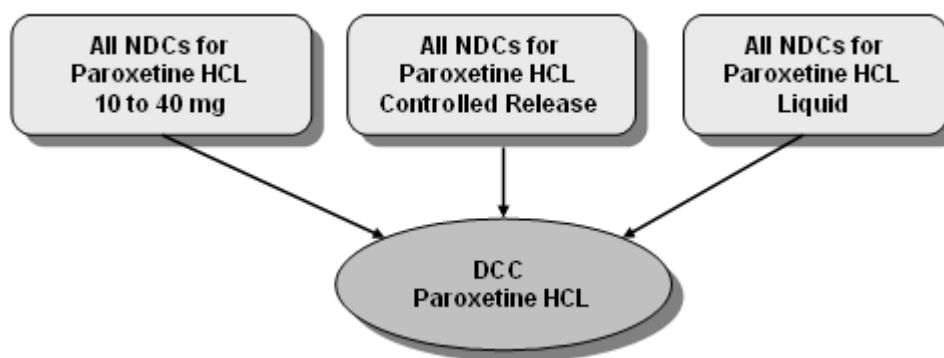
When the grouper assigns an ancillary record to an episode, it uses the ancillary record's diagnosis and procedure/revenue codes. It first evaluates diagnosis codes classified as *specific and nonspecific* to determine if these records can join an episode and then evaluates diagnosis codes classified as *sign and symptoms*. The ancillary record must occur within the clean period time window around an existing episode in order to be eligible to group to an existing episode. An ancillary record cannot extend an episode's length. It can only join an episode.

It is possible for an ancillary claim record to be medically inappropriate for any episode or condition for a member. If an ancillary record is not eligible to join an open episode it is then evaluated to determine if it can be assigned to a preventive ETG (screening and immunizations). If an ancillary record cannot be

assigned to a valid ETG or a preventive ETG, it is identified as an orphan record. An example of this would be when a provider calls in a prescription for the patient rather than seeing the patient in his/her office. The pharmacy claim would not have an anchor record to group to, so it would be considered an orphan.

For drug records, the methodology evaluates each pharmacy record against the episodes for which the patient is being treated. The NDC code assigned to the pharmacy record provides the clinical information to support this evaluation. Just as with the procedure and diagnosis codes, a drug eligibility table identifies ETGs to which an NDC can be associated and the strength of that association (low, medium, high), allowing the grouper to assign the drug claim record to the most clinically appropriate episode. HCPCS Level II procedure codes which represent a drug and its administration (e.g., injectables) are also considered to be pharmacy records, and are grouped in the same way. Due to the large number of NDCs defined for pharmacy services, the ETG methodology uses a drug classification hierarchy to support grouping. Each drug is associated with a Drug Classification Code (DCC) which represents a drug, or a specific dosage form of a drug. For example, the NDCs for all strengths of the antidepressant Paroxetine maps to the DCC of Paroxetine. The DCC concept assigned to the pharmacy services then supports grouping, not the NDC.

The following diagram illustrates this drug hierarchy.



Like ancillary records, drug records cannot extend an episode's length; they can only join an episode. A drug record must occur within an episode's clean period (pre and post) in order to be eligible to group to it.

Step 4: Finalize the Episodes (identify comorbidities and complicating factors, and assign episode severity)

After all claim records have grouped to an episode, the grouper then has all of the information it needs to finalize the episode.

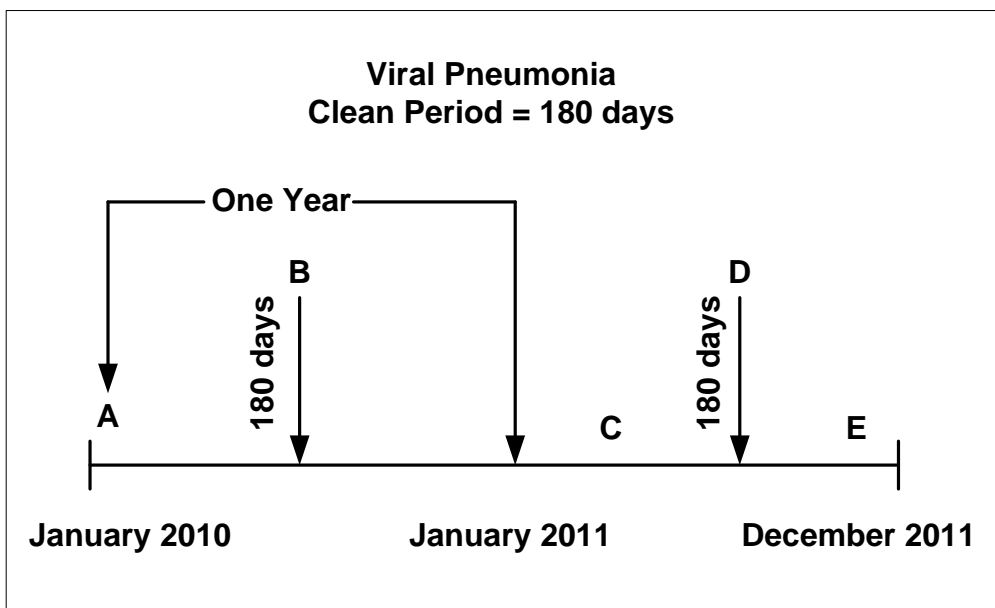
Episode Completeness

The notion of a complete episode is complex in the reality of service data. For example, assume the grouping start date is January 1, 2010. Does an episode for an acute condition with its first anchor record on January 3, 2010 begin with this claim or is the episode in progress? The episode of the acute condition might have begun sometime earlier (prior to January 1, 2010) but the data to identify the exact begin date are not available. The opposite is also true. With data available from January 1, 2009 through December 31, 2010, can it be known if a record incurred on December 21st for an existing episode is the end of the episode? The answer to both questions is that under certain circumstances it cannot be known whether a claim service record is actually the true beginning or the true end of an episode. A distinction must be

made between episodes which are to be considered complete from those whose completeness cannot be determined.

A clean start is defined as a situation where the true beginning date for an episode is known. The ETG methodology identifies a clean start by comparing the incurred date of the first anchor record of an episode with the beginning date of the overall service data range used in the grouping (or a member's beginning eligibility date, if later), with the episode's ETG clean period. If that anchor record date starts after the number of pre-episode clean period days, the episode is considered to have a clean start. If it occurs within the clean period days, it is considered to have an unknown start. The same methodology is true for a clean finish. A clean finish uses the same number of clean period days to determine a known finish. If the last anchor record occurs prior to the clean period days, the episode is determined to have a clean finish. If the last anchor record occurs within the clean period days, it has an unknown finish.

The following diagram illustrates this concept. In this example, anchor records for this episode occur at dates A, B, C, D and E. Note that treatment for this episode spans well over one year.



Assume that the time frame from each anchor record to the next is less than 180 days.

- The anchor record at date A is an unknown start.
- The anchor records at dates B and C (if either were the first anchor records in this episode) represent a clean start.
- The anchor records at dates D and E (if either were the last anchor records in this episode) represent an unknown finish.

The Episode Type identifies the completeness of an episode. Each acute episode is assessed for its status as a full year episode, and if it has a clean start and/or a clean finish. The episode's start and end dates are compared against the clean period days. From this information, the Episode Type can be determined.

The following table identifies the episode type values and whether they are considered complete or incomplete.

Episode Type	Description	Completeness Status
0	Clean start, clean finish	Complete
1	Clean start, unknown finish (full year)	Complete
2	Unknown start, clean finish (full year)	Complete
3	Unknown start, unknown finish (full year)	Complete
4	Clean start, unknown finish	Incomplete
5	Unknown start, clean finish	Incomplete
6	Unknown start, unknown finish	Incomplete
7	Incomplete annual episode	Incomplete

To account for chronic conditions, the ETG methodology utilizes different logic than the clean/unknown starts and finishes approach described above. ETG does this since chronic conditions are life-long going

forward. Further, to support proper episode-to-episode comparisons, the grouper limits the length of each episode for a chronic condition to one year. Such episodes which extend beyond one year and are subsequently limited to one year for analytical purposes are referred to as chronic annual episodes. As mentioned above, the grouper provides different configurable options on how to decide the starting point for chronic episodes: start month (a static month), grouping start date, grouping end date, eligibility start date and eligibility end date.

The grouper uses that selection and looks forward or back 365 days, collects all anchor records within that timeframe and assigns them to an episode. It does this in segments of 365 days. It then collects the non-anchor records and assigns them to the appropriate annual episode. To determine, within an annual year, if a chronic annual episode is considered complete, the grouper determines the member's enrollment during that time span: if the member is eligible for the entire year, that episode is considered complete (episode type 0); if not, the episode is considered incomplete (episode type 7).

The start date and end date for chronic annual episodes is based on the configurable selection made and is a full year date span. It does not reflect the date of the first and last anchor records within the episode, as acute episodes do.

Assign Complications/Condition Status, Comorbidities and Treatments to Episodes

The ETG methodology also identifies complication, comorbidity and treatment factors observed for each episode. After core grouping, episodes are evaluated to determine if they have any complicating factors, if there are any comorbidities associated with the episode's condition, and if the activity within the episode contains any treatment indicators. This information is reflected in the ETG number, allowing one to see specific characteristics of each episode. The first 6 digits are the base class, a unique number identifying the ETG; the 7th, 8th and 9th digits are the flags (with "0" indicating the factor was not observed, and "1" indicating it was) for with or without complication, with or without comorbidity and with or without treatments. The following table provides an illustration of the ETG numbers for Diabetes.

Base ETG	ETG Number	ETG Long Description
163000	163000000	Diabetes, w/o complication, w/o comorbidity, w/o surgery
163000	163000001	Diabetes, w/o complication, w/o comorbidity, with surgery
163000	163000010	Diabetes, w/o complication, with comorbidity, w/o surgery
163000	163000011	Diabetes, w/o complication, with comorbidity, with surgery
163000	163000100	Diabetes, with complication, w/o comorbidity, w/o surgery
163000	163000101	Diabetes, with complication, w/o comorbidity, with surgery
163000	163000110	Diabetes, with complication, with comorbidity, w/o surgery
163000	163000111	Diabetes, with complication, with comorbidity, with surgery

Identifying the condition status/complications for an episode provides specificity of the episode's clinical condition, any complications associated with the episode, and the disease progression, when applicable. The ETG methodology categorizes some diagnosis codes into groupings of similar diagnoses, referred to as condition status codes. For example, condition statuses for Diabetes include Diabetes Type 1 and Diabetes Type 2. Examples of condition statuses that specify complications of Diabetes are Diabetic Coma and Diabetic Ketoacidosis.

Condition status codes are identified by diagnosis codes on anchor records, are ETG-specific and must occur within an episode in order for the episode to be designated as *with complication*. For example, the diagnosis of diabetic coma would not be a condition status code for an episode of chronic bronchitis. It would, however, be a condition status code for an episode of diabetes. In addition to flagging the ETG as *with complication*, the grouper provides an optional output that lists each condition status that was identified within an episode.

A comorbidity is defined as the presence of more than one disease or health condition in a member at a given time. The ETG methodology categorizes some diagnosis codes into groupings of similar diagnoses, referred to as comorbidity codes. For example, the comorbidity *Chronic bronchitis* is a

compilation of the various diagnosis codes designated as such (e.g. Bronchiectasis, Chronic bronchitis NOS, etc.). The grouper identifies comorbidities by evaluating diagnosis codes on the records designated as anchor records. It keeps track of all of a member's comorbidities, gives each comorbidity an active period (approximately two years) and uses that information to determine what episodes can be labeled as *with comorbidity*.

Comorbidities are ETG-specific. For example, the comorbidity of Chronic Bronchitis would not be a comorbidity for an episode of Lymphoma. It would, however, be a comorbidity for an episode of Congestive Heart Failure. Any comorbidity that has an active period that occurs during an eligible episode's time frame is considered a comorbidity for that episode.

Treatment indicators are categorizations of services such as defining surgeries and active management procedures for malignant neoplasms (chemotherapy and radiation therapy services). These categories are a grouping of similar procedures. For example, the treatment indicator for Chemotherapy is a compilation of the procedure codes and revenue codes that are classified as chemotherapy services.

When flagging the ETG as *with or without surgery*, the ETG methodology provides more specificity for certain conditions. For malignant neoplasms, the grouper will also designate if an episode incurred active management services. For cardiology conditions, the grouper will also designate if an episode incurred these specific defining surgeries: angioplasty, CABG and valve surgery. The exact nature of the treatment will be specified by the value of the treatment indicator digit. The procedure and/or revenue codes categorized as a treatment indicator must occur within an episode in order for the episode to be flagged as such.

Given the ETG numbering scheme, where the first six digits define the base condition and the remaining digits describe treatment and other clinical factors, users of the ETG outputs have flexibility in how the grouped results are applied. For example, if the desire is to measure at the condition level, episodes are combined for analysis using the first six digits of the ETG number (the first six digits identify the base ETG). If the combination of condition and the presence (or not) of a significant surgery are desired to support comparisons, users would combine episodes using the first six digits and the ninth digit of the ETG number. As described below, severity levels can also be used in addition to support comparisons.

Severity Adjusting Episodes

Complications, comorbidities and member demographics are used in determining the severity of the member's episode. The ETG methodology takes advantage of the relevant complication and comorbidity factors (indicating a sicker member who may require more extensive treatment for a related condition) when determining an episode's severity. The result is a severity score and severity level for episodes. The higher the severity score, the more resources are expected relative to episodes with a lower severity score for the same base ETG.

After condition statuses and comorbidities have been assigned to an episode, the grouper can determine the severity score and severity level for each episode. Each contributing factor to an episode is given a weight: a demographic weight (age & gender), condition status and comorbidities weight, additional weights if there are interactions between multiple complications and interactions between multiple comorbidities (interaction weight), and weights for multiple complications and/or multiple comorbidities (multiple count weights). These weights are then summarized to generate an overall severity score for the episode.

A separate set of weights is computed for each base ETG where severity is measured. There are separate age/gender weights for elderly (age 65 and older) and non-elderly weights for many conditions.

Based on the severity score, the severity level indicates a ranking of where the specific episode is relative to the population of all episodes within that base ETG. There are four potential severity levels, where the value 1 indicates a less severe episode and the value 4 indicates the most severe episode. Not all ETGs are severity adjusted and not all ETGs have 4 severity levels. All episodes for ETGs that are not severity adjusted have a severity score of 1.00 and a severity level of 1.

Outlier Status

Outlier status is the comparison of an episode's costs to a dollar amount specified for each ETG. An episode is considered a low outlier if its costs are below the ETG-specific low outlier amount; an episode is considered a high outlier if its costs are above the ETG-specific high outlier amount. The ETG Base Class in combination with the episode's severity level is used to determine the outlier status. All costs within the episode are evaluated (i.e., all record types).

II. ETG Construction Logic for Pneumonia Episodes of Care

Episodes for the submitted Pneumonia measures are defined using the Episode Treatment Group (ETG) methodology. Section I of this document describes the general approach used by ETG to create episodes of care. This section applies that general methodology to create Pneumonia episodes. Also, please note that this description will reference a number of attachments included with the submission for these measures, including:

- S5_Pneumonia_DataDictionary (Excel workbook attachment). This attachment describes the clinical relationships between diagnosis and procedure codes and the episode condition.
- S8_Pneumonia_ClinicalLogic (Excel workbook attachment). This attachment includes Worksheets that describe the details around the components of Pneumonia methodologies that relate to comorbidities, condition status factors, and severity adjustment.

The individual Worksheets in these attachments that relate to the specific components of the methodology are referenced in the following discussion.

As noted above, the ETG episode building process has four important steps:

1. Identify Records; Assign Record Type and Anchor Records, Classify Diagnoses and Procedures
2. Build Episodes from Anchor Records
3. Group Non-Anchor Records to Episodes
4. Finalize the Episodes (identify comorbidities and complicating factors, and assign episode severity)

In this section we discuss how these steps apply specifically to creating Pneumonia episodes.

Step 1 (PNEUMONIA). Assign a Record Type and Anchor Records, Classify Diagnoses and Procedures

Record Type Assignment

Each service record, or claim line, is assigned a record type. Assigning Record Type uses a combination of the procedure code and the provider type on the claim. As described in Section I, there are 5 record types used by ETG:

- Management Records (for example, an office visit or consultation)
- Surgery Records (for example, a surgical procedure)
- Ancillary Records (for example, a lab test or imaging service)
- Facility Records (room and board)
- Pharmacy

Anchor Record Assignment

Anchor Records are also identified as part of this step. Anchor records play an important role in building Pneumonia episodes. Anchor records have a record type of Management, Surgery, or Facility. An anchor record indicates that a clinician has evaluated the patient and has initiated the treatment and care of the patient for the condition.

Classify Diagnosis Codes

As described in Section I of this document, ETG relies heavily on the diagnosis codes to help identify discrete episodes. The diagnosis identifies the condition being treated, which broadly translates to an ETG. Each diagnosis code is identified with a given diagnosis class. There are three diagnosis classes applied across all episodes, including ETG:

- **Specific:** These are ICD-9 diagnosis codes that indicate a specific disease. This code represents a disease or condition (as opposed to a sign or symptom) and is specific enough to be linked to a single ETG. ICD-9 Diagnosis code 481 (pneumococcal pneumonia) is an example of a specific diagnosis code. It is primary to an episode of Pneumonia.
- **Non-Specific:** These ICD-9 diagnosis codes represent a disease or condition (as opposed to a sign or symptom), but may not be specific enough to identify a single ETG. ICD-9 Diagnosis code 079.89 (Other specified viral infection, in conditions classified elsewhere and of unspecified site) is an example of a non-specific ICD-9 code. Although this code represents disease as opposed to signs or symptoms of disease, it is not specific as to representing a single disease. This code is assigned a lower specificity—Non-specific.
- **Signs and Symptom:** These ICD-9 diagnosis codes represent signs and symptoms of disease as opposed to disease or condition. ICD-9 Diagnosis code 780.61 (fever presenting with conditions classified elsewhere) does not represent diseases, but only signs and symptoms that could be related to multiple diseases. These codes are assigned the lowest specificity— Signs and Symptoms. Signs and Symptoms codes may be eligible for many ETGs due to their generic nature.

The ETG methodology considers one member at a time. The service records and their diagnosis codes are grouped in several distinct passes for a member. The methodology first processes the specific and non-specific diagnosis codes on anchor records so that concrete conditions/diseases are created. It then processes the sign and symptom diagnosis codes in reverse chronological order based on service dates to determine the best episode each of them can group to. Using this approach, the logic described below that links service records to Pneumonia episodes is applied.

Each diagnosis code is matched with one or more ETGs through a diagnosis eligibility table, including codes that match to the ETG for Pneumonia. Each diagnosis code is further ranked, based on its strength of association with the Pneumonia ETG and other ETGs. The rank values are:

- **Primary Classification Ranking diagnoses:** A primary ranking classification for a diagnosis describes a condition that defines Pneumonia. These are the main diagnosis codes that impact grouping decisions for Pneumonia. The Diagnosis codes that are classified as primary to Pneumonia are listed on the “PrimaryDxCodes” worksheet within the attachment “S5_Pneumonia_DataDictionary”.
- **Incidental Classification Ranking diagnoses:** Incidental diagnosis codes are eligible for Pneumonia, but not classified as primary. Incidental diagnoses are further ranked as low, medium, and high, representing the strength of the match association with Pneumonia. The Diagnosis codes that are incidental to Pneumonia are listed on the “IncidentalDxCodes” worksheet within the attachment “S5_Pneumonia_DataDictionary”. The column “diagnosisEligibilityType” in the worksheet describes the ranking where 3 represents a high association, 2 represents a medium association and 1 represents a low association.

Also note that a diagnosisEligibilityType value of “5” in the attachment “S5_Pneumonia_DataDictionary” means that the diagnosis code can shift a Pneumonia ETG to a different ETG. Episodes in the Pneumonia ETG can shift to the Viral Pneumonia ETG. This will occur if a certain set of diagnosis codes referring to that condition appear on a claim record. This shift, however, would still keep the episode within the broader category of Pneumonia and within the submitted measures.

Classify Procedure Codes

Procedure codes are also matched to Pneumonia. All procedure codes that are eligible for Pneumonia are listed on the “ProcedureCodes” worksheet within attachment “S5_Pneumonia_DataDictionary”. In

some instances a procedure code may be eligible for more than one ETG. In these cases, where multiple episodes are observed for a member where the procedure code has eligibility, the ETG methodology uses strength of the clinical relationship between the procedure code and the episode. The CPT and HCPCS procedure codes on this worksheet are ranked from 1 to 4 to specify the strength of the clinical relationship between the procedure code and Pneumonia. The column “ProcedureRank” in the worksheet describes that strength of association, with 4 being the strongest association and 1 being the lowest. The grouping of services based on diagnosis and procedure codes is further described below.

Step 2 (PNEUMONIA). Build Episodes from Anchor Records.

Given the clinical relationships described above, the following steps are further used to build episodes from anchor records:

- a. Anchor records are grouped in two passes through the member’s data. The first pass groups the anchor records with specific and non-specific diagnoses. The second pass groups anchor records with sign and symptoms diagnoses.
- b. All anchor records are grouped before all non-anchor records. Non-anchor records have a record type of Ancillary or Pharmacy.
- c. An episode of Pneumonia requires an anchor record to start an episode. For an anchor record to start an episode of Pneumonia, it must have a procedure code that is eligible for Pneumonia and an ICD-9 diagnosis code that is primary for Pneumonia. As an example of an anchor record that starts an episode of Pneumonia, a pulmonologist sees a member and submits a claim record using the CPT procedure code 99212 (Office visit, established patient) with and ICD-9 diagnosis code 481 (pneumococcal pneumonia).

Note that a single anchor record can start more than one episode. For example, an anchor record with a diagnosis and procedure code combination that is eligible for Pneumonia will start a Pneumonia episode. If that record also has a diagnosis and procedure code combination that is eligible for Hypertension, it will also start a Hypertension episode. (See Section I above for a discussion of the concept of phantom episode clusters.)

- d. Once an episode of Pneumonia is started, further anchor records can group to that episode. For a record to be eligible to join an already open episode of Pneumonia the procedure code for the record must be eligible for Pneumonia and the diagnosis code must have either a primary or incidental relationship to Pneumonia.
- e. In some cases, an anchor record can be eligible to join more than one episode (because it may have more than one diagnosis code). When determining the episode an anchor record groups to, the specificity of the diagnoses determines the priority for grouping the record. For Pneumonia, a specific code (like 481 (pneumococcal pneumonia)) has priority over a non-specific code (like 079.89 (Other specified viral infection, in conditions classified elsewhere and of unspecified site)).
- f. As described above, diagnosis codes with specificity of sign and symptom have the lowest priority for grouping. An example of a sign and symptom code is 780.61 (fever presenting with conditions classified elsewhere). Anchor records with only sign and symptom diagnosis codes are not grouped until anchor records with more specific disease diagnosis codes are grouped. For example, an office visit record on Jan 15th with an ICD-9 code of 780.61 (fever presenting with conditions classified elsewhere) is followed by an office visit record on Feb 1st with an ICD-9 code of 481 (pneumococcal pneumonia). The grouper would skip the anchor record service on Jan 15th because it only had a sign and symptom diagnosis code. It would then open up an episode of Pneumonia based on the claim on Feb 1st. On the second pass, the grouper would use the incidental relationship between the sign and symptom ICD-9 code 780.61 to group this claim to the already open Pneumonia episode. Without this methodology, the claim on Jan 15th would not group to the Pneumonia episode on the first pass because at the time of the first pass evaluating the claim on Jan 15th, the Pneumonia episode did not exist.

- g. Following these steps, anchor records have been used to open episodes of Pneumonia, as well as episodes for other conditions and anchor records have been assigned uniquely to individual episodes based on the clinical logic described above and in the attachment “S5_Pneumonia_DataDictionary”.

Step 3 (PNEUMONIA). Group Non-Anchor Records to Episodes.

Non-anchor records (record type “Ancillary” and “Pharmacy”) can not open episodes on their own. For example, a service record with a procedure code of 89399 (Pathology lab procedure) and an ICD-9 code of 481 (pneumococcal pneumonia) can group to an open episode of Pneumonia but can not open the episode itself.

Ancillary service records group to Pneumonia based on a match of diagnosis and procedure code to Pneumonia. As described above, attachment S5_Pneumonia_DataDictionary includes the diagnosis and procedure mappings for Pneumonia that inform these assignments.

In some instances an Ancillary procedure code may be eligible for more than one ETG. In these cases, where multiple episodes are observed for a member where the procedure code has eligibility, the ETG methodology uses strength of the clinical relationship between the procedure code and the episode. The column “ProcedureRank” in the “ProcedureCodes” worksheet within attachment “S5_Pneumonia_DataDictionary” describes that strength of association, with 4 being the strongest association and 1 being the lowest.

Pharmacy services group differently because they usually do not have ICD-9 diagnosis codes associated with them. Pharmacy claims group by using a table that maps NDC to the DCC code (Drug Category Code) based on the drug’s active ingredients and route of administration. Selected pharmacy services with a CPT or HCPCS code are also mapped to a DCC (e.g., J-codes describing injections). For example, a service with an NDC code 68115006630 (Cefaclor extended release 500mg) will map to DCC 00408. The DCC 00408 has a relationship with Pneumonia as defined by the “Pharmacy” worksheet in the attachment “S5_Pneumonia_DataDictionary”. Therefore this claim could join an open episode of Pneumonia. It could not, however, start an episode of Pneumonia on its own.

In some instances a DCC code may be eligible for more than one ETG. In these cases, where multiple episodes are observed for a member where the DCC code has eligibility, the ETG methodology uses strength of the clinical relationship between the DCC code and the episode. The column “Rank” in the “Pharmacy” worksheet within attachment “S5_Pneumonia_DataDictionary” describes that strength of association. The lower the value is for Rank, the stronger the association between the DCC and the episode.

Due to the size of the attachment the full list of NDC to DCC mappings has not been provided within this submission. This file is available upon request.

Step 4 (PNEUMONIA): Finalize the Episodes (Identify Comorbidities and Complicating Factors, and Assign Episode Severity)

Episode Completeness

Episode completeness, the assignment of comorbidities and condition status, and the measurement of episode severity are the key steps in finalizing a Pneumonia episode.

In terms of episode completeness, Pneumonia is an acute condition. Therefore the clean periods described in Section I as part of the general ETG methodology are applicable. All clinically consistent treatments for the care of a Pneumonia patient will group to the episode of Pneumonia based upon the clean period.

Assigning Comorbidities and Condition Status Factors to Pneumonia Episodes

The ETG methodology identifies the comorbidities and condition status factors observed for each Pneumonia episode. These factors provide specificity of the episode's clinical condition and also play a key role in assigning a severity score and level to the episode. An example of the assignment of comorbidities and condition status factors and creation of a severity score and level is provided at the end of step 4 and references to this example are provided in the following text.

Condition status factors for Pneumonia episodes are identified by diagnosis codes on anchor records that occur within the Pneumonia episode. The "ConditionStatusToDxCodeMap" Worksheet in the attachment "S8_Pneumonia_ClinicalLogic" describes the mapping of diagnosis codes to condition status factors. In particular, the following condition status factors are defined for Pneumonia:

- H. influenzae pneumonia
- K. pneumoniae pneumonia
- Pseudomonal pneumonia
- Staphylococcal pneumonia
- Pneumonia, with abscess
- Chlamydia pneumonia
- Legionnaires pneumonia
- Other gram negative pneumonia
- Anaerobic pneumonia
- Streptococcal pneumonia
- Pneumococcal pneumonia

For all condition status factors that occur for a given episode of pneumonia, the severity weight is added to the total severity weight of the episode.

Each condition status factor belongs to a condition status group. If several condition status factors eligible for a given ETG occur in different condition status groups then all are assigned to the ETG. If there are 2 or more distinct condition status factors in a single condition status group then each condition status factor within the group has a ranking and only a single condition status factor is assigned to the ETG based on its ranking within the condition status group. If two or more condition status factors have the same ranking within a given condition status group, then they may or may not compete to be assigned to the ETG. If the compete flag is 1 or "on" then only one of the condition status factors is assigned to the ETG depending on which condition status factors occurs with the largest number of anchor claims. If the compete flag is 0 or "off" then all of the conditionStatuses with the same ranking are assigned to the ETG.

For the case of pneumonia, all condition status factors are in the same condition status group, all have the same priority and the compete flag is 0 (off). All condition status factors that occur for a given episode of pneumonia are assigned to the episode.

Comorbidity factors for Pneumonia episodes are identified by evaluating diagnosis codes on the records designated as anchor records from outside the Pneumonia episode. ETG tracks all of a member's comorbidities, gives each comorbidity an active period (approximately two years) and uses that information to determine what episodes can be labeled as with comorbidity. The comorbidities defined by ETG for Pneumonia are described in the "ComorbtoDxCodeMap" Worksheet in the attachment "S8_Pneumonia_ClinicalLogic", including the individual diagnosis codes that map to each. Examples of these comorbidities include AIDS, Septicemia and Other Infectious Diseases. In the example included below, the comorbidity 80071 (Other Drug Dependence) is assigned to the Pneumonia episode based upon the diagnosis information on anchor records that occur outside of the Pneumonia episode.

Each comorbidity belongs to comorbidity groups one and two. Comorbidities that belong to the same Group 2 are treated identically in the model and only counted once for each comorbidity in the same Group 2. Multiple Group 2's can belong to the same Group 1. If two comorbidities in two different Group 2's are in the same Group1 then only the comorbidity in the Group 2 with the higher severity is counted. When several comorbidities are identified for an episode, the rankings of the comorbidity's Group 2 within

Group 1 will determine whether or not the severity of a comorbidity is added to the severity score of the given episode.

As is the case for condition status factors, each comorbidity belongs to a comorbidity group and when several comorbidities are identified for an episode, the rankings of the comorbidities within the specific comorbidity groups and the status of the compete flag will determine whether or not the comorbidity is assigned to the given episode.

In addition to co-morbidities and condition status factors, interactions between two co-morbidities or two condition status factors are also assessed. These interactions are used in assigning severity to a pneumonia episode, as described in section 8.5.

Assigning Severity to Pneumonia Episodes

Condition status factors, comorbidities and patient demographics are used in determining the severity of the Pneumonia episode. The ETG methodology takes advantage of the relevant condition status and comorbidity factors when determining an episode's severity. In general, these factors indicate a higher risk patient who may require more extensive treatment for Pneumonia. The result is a severity score and severity level for each episode. The higher the severity score, the more resources are expected relative to other Pneumonia episodes.

The condition status and comorbidity factors found to have an impact on the required resources for Pneumonia episodes are included in the severity model. Each contributing factor to an episode is given a weight: a demographic weight (age & gender), condition status and comorbidities weight, additional weights if there are interactions between multiple complications and interactions between multiple comorbidities (interaction weight), and weights for multiple complications and/or multiple comorbidities (multiple count weights). These weights are then summarized to generate an overall severity score for the episode.

A separate set of weights is computed for the base ETG of Pneumonia. There are separate age/gender weights for elderly (age 65 and older) and non-elderly weights.

The following worksheets in the attachment "S8_Pneumonia_ClinicalLogic" describe the factors and weightings used in determining the level of severity for a Pneumonia episode (see the notes at the top of each worksheet for a further description of the comorbidity or condition status concept):

- Worksheet "Comorbidities" – includes the ComorbidityCodes and Comorbidity Groups used to determine severity for Pneumonia. The rightmost columns include a "Priority" hierarchy along with risk weights for the non-elderly and elderly models. The Priority column is applied where multiple ComorbidityCodes in the same Comorbidity group are identified, with the lowest number priority receiving precedence. Each risk weight reflects the incremental contribution of having a specific Comorbidity factor on Pneumonia severity. (Note that a number of the individual ComorbidityCodes that are clinically similar are combined and used as a group in measuring severity. Only one of these individual Codes is needed to trigger the aggregate Comorbidity Group2, after application of any relevant Priority.);
- Worksheet "ComorbidityInteractions" – includes the interactions between Comorbidity Groups used to determine severity for Pneumonia. The rightmost columns include risk weights for the non-elderly and elderly models. Each risk weight reflects the incremental contribution of having a specific Comorbidity interaction factor on Pneumonia severity;
- Worksheet "ComorbidityCounts" – includes the additional severity factors added for those episodes where 3 or more+ comorbidity factors were observed. The rightmost columns include risk weights for the non-elderly and elderly models. Each risk weight reflects the incremental contribution of having a specific Comorbidity Count factor on Pneumonia severity;
- Worksheet "ConditionStatuses" – includes the Condition Status factors used to determine severity for Pneumonia. The rightmost columns include risk weights for the non-elderly and elderly

models. Each risk weight reflects the incremental contribution of having a specific Condition Status factor on Pneumonia severity;

- Worksheet “ConditionStatusInteractions” – includes the interactions between Condition Status factors used to determine severity for Pneumonia. The rightmost columns include risk weights for the non-elderly and elderly models. Each risk weight reflects the incremental contribution of having a specific Condition Status interaction factor on Pneumonia severity;
- Worksheet “ConditionStatusCounts” – includes the additional severity factors added for those episodes where 3 or more condition status factors were observed. The rightmost columns include risk weights for the non-elderly and elderly models. Each risk weight reflects the incremental contribution of having a specific ConditionStatus Count factor on Pneumonia severity;
- Worksheet “Demographics” – includes the additional severity factors added based on age and gender. Each risk weight reflects the incremental contribution of having a specific Demographic factor on Pneumonia severity;

The severity score for a Pneumonia episode is the sum of the weights for each of the factors observed for the episode.

The following example shows the calculation of severity score and level for a Pneumonia episode. The example describes a Female patient, age 32, observed to have a number of anchor records with a diagnosis that maps to the Pneumonia ETG. The patient is also observed to have one condition status factor and one comorbidity that are also eligible for Pneumonia. The condition status factor (70142: Mycoplasma pneumonia) was identified through one or more anchor records observed within the episode where the diagnosis on the records mapped to that condition status factor. The comorbidity (80071: Other drug dependence) was identified on one or more anchor records observed outside of the Pneumonia episode.

The patient receives a severity marker for each of the condition status and comorbidity factors and a risk weight is assigned to each. The patient also receives severity weight related to her age and gender which fall into the “Female 19-34” range.

A severity score of 0.4807 is calculated based upon the sum of:

- The Demographic weight of 0.2498 (see worksheet “Demographics” within S8_Pneumonia_ClinicalLogic where column “gender”=F and column “ageRange”=19-34);
- The condition status weight for Mycoplasma pneumonia of 0 (see worksheet “ConditionStatuses” within S8_Pneumonia_ClinicalLogic where column “conditionStatusCode”=70142),
- The comorbidity weight for Other drug dependence of 0.2309 (see worksheet “Comorbidities” within S8_Pneumonia_ClinicalLogic where column “comorbiditycode”=80071. The Other drug dependence comorbidity belongs to the Comorbiditygroup2 of Drug Dependence.);
- The final severity score is calculated as $0.2498 + 0 + 0.2309 = 0.4807$

Based on the severity score, the severity “level” indicates a categorical ranking of where the specific episode is relative to the population of all Pneumonia episodes. There are four potential severity levels for Pneumonia, where the value 1 indicates a less severe episode and the value 4 indicates the most severe episode. The “Thresholds” Worksheet in attachment “S8_Pneumonia_ClinicalLogic” describe the three cut-off points that define the four levels of severity for Pneumonia episodes.

The following example shows the calculation of severity score and level for a Pneumonia episode.

Episode	ETG(Base Condition)	Complications			
0001	Bacterial lung infection	70142			
Comorbidities					
		80071			
Severity Level					
1	2	3	4		
< 0.7	0.7 - 1.5	1.5 - 2.5	> 2.5		
Calculation of Relative Episode Severity					
Indicator	Code	Description	Severity Weight		
Demographic	4	F19_34	0.2498		
Condition Status	70142	Mycoplasma pneumonia	0		
Co-morbidity	80071	Other drug dependence	0.2309		
Interaction					
Total			0.4807		

Example of Calculating ETG Episode Severity Score and Level.

The ETG methodology for Pneumonia uses medical and pharmacy service records/claims and member enrollment as input. Outputs for Pneumonia include the identification of the individual service records assigned to a Pneumonia episode, along with the details of the grouping, including ETG, episode ID, record type, cluster ID, and cluster provider. An episode summary record is also produced, describing the episode ID, the ETG assigned (Pneumonia), the severity score and severity level for the episode, episode completion status, and other episode-level characteristics.

Note that the episode grouping methodology for Pneumonia is applied in the context of the full-breadth of the ETG clinical methodology, where all clinical conditions and episodes can be considered and created for a member. In this way, decisions regarding the appropriate assignment of a service record to an individual Pneumonia episode can be made while considering all conditions and episodes for that member, including episodes other than Pneumonia.

The episode results can then be used to support episode-based measures of the resources involved in diagnosing, managing and treating Pneumonia as further discussed within the Pneumonia specifications provided in the submission form.



Physician		Number of Episodes:	93
Name:	Provider 6388502012	Case Mix, Episodes:	0.48
		Peer Group	
Primary ID:	6388502012	Peer Group Number of Episodes:	5,430
		Peer Group Name:	II Cardiology
Specialty:	Cardiology	Key Statistics	
		Overall Quality Index:	1.02
		Overall Cost Index, Episodes:	1.14
		Confidence Intervals for the Index	
		Overall Quality Index:	No data available
		Overall Cost Index, Episodes:	No data available

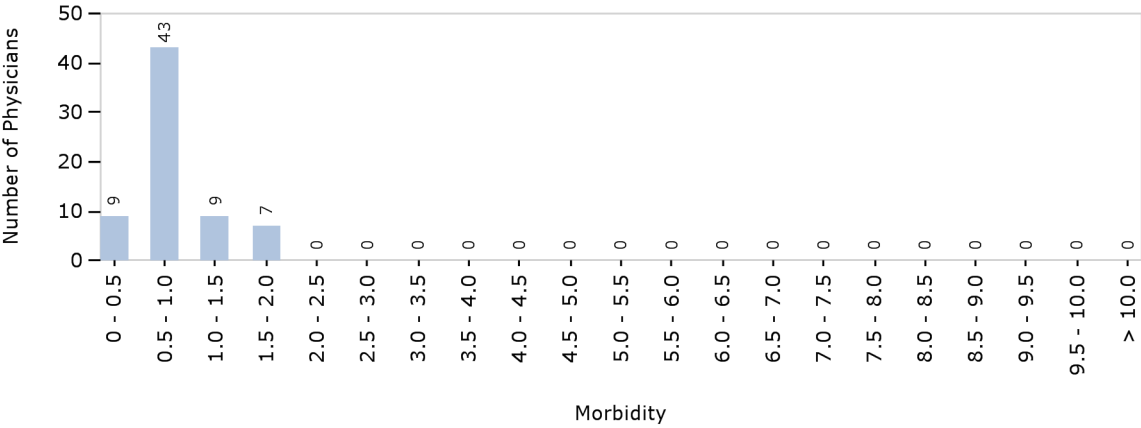
**Statistical significance of difference between
index and peer group average: * p<0.10; ** p <
0.05**

Episode Case Mix Summary

Top 10 ETGs, by Total Cost (Completed Episodes of Care)

ETG Family Description	Episodes			Encounters (Per 1000 Episodes)	
	Episodes	Specialist's Cost / Episode	Peers Cost / Episode	Specialist's Encounters / 1000 Episode	Peers Encounters / 1000 Episode
Hypertension	43	\$1,569.36	\$1,228.51	14,779	12,844
Hyperlipidemia, other	19	\$720.64	\$631.67	7,169	6,829
Ischemic heart disease	9	\$1,511.63	\$2,378.04	12,889	13,765
Valvular disorder	14	\$818.25	\$1,047.19	4,367	7,315
Cardiomyopathy	3	\$2,407.90	\$1,340.66	16,583	14,088
Pulmonary embolism	1	\$3,244.43	\$3,897.41	38,714	24,716
Congestive heart failure	1	\$2,817.56	\$1,496.61	6,600	14,084
Atherosclerosis	2	\$702.92	\$387.57	1,500	1,125
Atrial fibrillation & flutter	1	\$507.36	\$1,715.52	25,500	21,127
All Others	0	--	--	--	--
All Episodes	93	\$1,304.04	\$1,211.06	11,523	10,879

Relative Morbidity Histogram

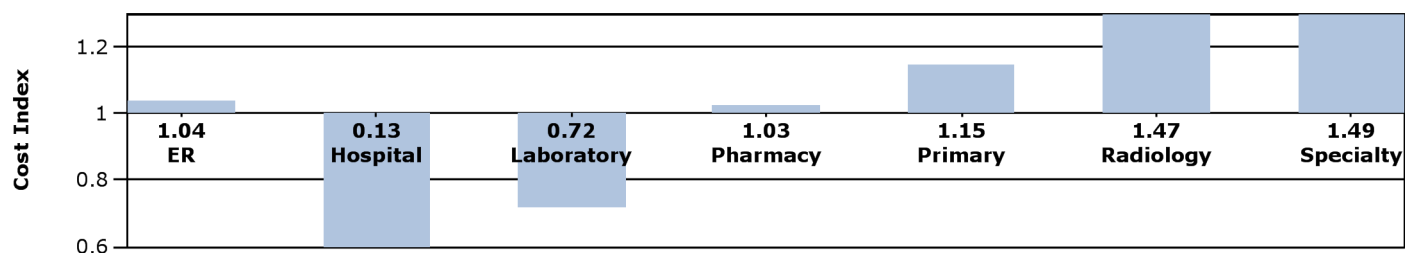


Quality Measures

**As of the End of the Report Period
(Members Must be Continuously Enrolled with Plan a Minimum of 12 Months)**

	Number of Quality Opportunities		Rates		Index
	With Compliance	Total	Provider Rate	Peer Rate	Quality Index
Cardiology					
HTN					
Pt(s) taking an ACE-inhibitor, angiotensin II receptor antagonist, diuretic, or aldosterone receptor blocker that had a serum K+ in last 12 rpt mos.	12	15	0.80	0.81	0.99
HTN					
Pt(s) taking an NSAID med.	21	23	0.91	0.92	0.99
HTN					
Pt(s) that had an annual physician visit.	23	23	1.00	0.97	1.03
HTN					
Pt(s) that had a serum creatinine in last 12 rpt mos.	19	23	0.83	0.80	1.03
CAD					
Pt(s) conon 2 meds (nitrate and phosphodiesterase type 5 inhibitor) w/ interacting properties.	6	6	1.00	1.00	1.00
CAD					
Pt(s) that had an OV for CAD care in last 12 rpt mos.	6	6	1.00	0.97	1.03
Endocrinology					
Hyperlipidemia					
Pt(s) taking a statin-containing med, nicotinic acid or fibric acid derivative that had an annual serum ALT or AST test.	10	10	1.00	0.92	1.09
Hyperlipidemia					
Pt(s) w/ a LDL cholesterol test in last 12 rpt mos.	16	16	1.00	0.90	1.12
Hyperlipidemia					
Pt(s) w/ the most recent LDL result <160mg/dL.	4	5	0.80	0.93	0.86
Hyperlipidemia					
Pt(s) w/ a HDL cholesterol test in last 12 rpt mos.	16	16	1.00	0.90	1.11
Hyperlipidemia					
Pt(s) w/ the most recent HDL result >= 40mg/dL.	1	5	0.20	0.68	0.29
Hyperlipidemia					
Pt(s) w/ a triglyceride test in last 12 rpt mos.	16	16	1.00	0.90	1.12
Total	150	164	0.91	0.89	1.03

Cost Index Summary, by Service Category



Cost and Utilization Summary Measures

Profiled Costs

	Actual Encounters	Peers Encounters	Actual Cost / Episode	Peers Cost / Episode	Cost / Episode Index	Actual Total Cost
ER	5	4	\$53.98	\$52.08	1.04	\$5,020
Facility	3	2	\$45.42	\$45.48		\$4,224
Professional	2	2	\$8.56	\$6.59		\$796
Hospital Services	8	17	\$29.49	\$227.90	0.13	\$2,743
Inpatient Facility	0	2	\$0.00	\$108.97		\$0
Outpatient Hospital Surgery	0	2	\$0.00	\$57.97		\$0
Laboratory	40	51	\$19.29	\$26.93	0.72	\$1,794
Facility	0	3	\$0.00	\$6.96		\$0
Professional	40	48	\$19.29	\$19.97		\$1,794
Pharmacy	492	499	\$271.71	\$264.70	1.03	\$25,269
Anti-Infective Agents	4	5	\$0.37	\$1.51		\$35
Cardiovascular agents	359	393	\$221.64	\$227.13		\$20,613
Primary Care Core	119	104	\$68.41	\$59.66	1.15	\$6,362
PCC Diagnostic	57	61	\$18.70	\$28.57		\$1,739
Radiology	34	27	\$210.93	\$143.69	1.47	\$19,617
Facility	1	3	\$2.51	\$25.55		\$234
Professional	34	24	\$208.42	\$118.14		\$19,383
Specialty Care	373	309	\$650.24	\$436.10	1.49	\$60,472
Medical Specialty	354	287	\$606.05	\$402.22		\$56,363
PCP Specialty	1	3	\$0.22	\$3.31		\$21
Surgical Specialty	3	4	\$1.38	\$4.23		\$129
Total	1,072	1,012	\$1,304.04	\$1,211.06	1.08	\$121,276

Overall Cost Index: 1.14

Utilization Rates Per 1,000 Episodes

Reporting Period : 1/1/2006 - 12/31/2007

Provider # : 6388502012

	Actual	Peers	Index
Specialist Visit Rate	1,387	1,407	0.99
Other Specialty Care Rate	839	616	1.36
Radiology Procedure Rate	391	365	1.07
MRI Procedure Rate	0	3	0.00
Laboratory Procedure Rate	908	887	1.02
Overall Prescribing Rate	5,290	5,360	0.99
Generic Prescribing %	0%	0%	--
ER Visit Rate	48	41	1.19
Admits per 1000 Episodes	0	25	0.00
Days per 1000 Episodes	0	63	0.00
Average Length of Stay	--	2.50	0.00

Episode Detail and Analysis

Atherosclerosis

Total Specialty Episode Costs: \$1,406

Cost per Episode	# of Episodes	Total	Primary Care Core	Specialty Care	Laboratory	Radiology	Hospital	Pharmacy	ER
Actual	2	\$702.92	\$0.00	\$702.92	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Peers		\$387.57	\$0.00	\$387.57	\$0.00	\$0.00	\$0.00	\$0.00	\$0.00
Index			--	--	--	--	--	--	--

Encounters per 1000 Episodes

Actual		0	1,500	0	0	0	0	0	0
Peers		0	1,125	0	0	0	0	0	0
Index		--	--	--	--	--	--	--	--

Atrial fibrillation & flutter

Total Specialty Episode Costs: \$507

Cost per Episode	# of Episodes	Total	Primary Care Core	Specialty Care	Laboratory	Radiology	Hospital	Pharmacy	ER
Actual	1	\$507.36	\$6.20	\$106.50	\$25.66	\$0.00	\$75.58	\$293.43	\$0.00
Peers		\$1,715.52	\$35.87	\$465.51	\$46.52	\$69.43	\$459.09	\$533.92	\$105.18
Index			--	--	--	--	--	--	--

Encounters per 1000 Episodes

Actual		1,000	9,000	3,500	0	1,000	11,000	0	0
Peers		1,435	6,459	2,597	208	319	9,968	141	141
Index		--	--	--	--	--	--	--	--

Cardiomyopathy

Total Specialty Episode Costs: \$7,224

Cost per Episode	# of Episodes	Total	Primary Care Core	Specialty Care	Laboratory	Radiology	Hospital	Pharmacy	ER
Actual	3	\$2,407.90	\$32.88	\$1,410.90	\$2.32	\$0.00	\$613.18	\$348.61	\$0.00
Peers		\$1,340.66	\$19.72	\$515.26	\$49.66	\$109.92	\$300.36	\$345.74	\$0.00
Index			--	--	--	--	--	--	--

Encounters per 1000 Episodes

Actual		1,333	3,750	167	0	1,000	10,333	0	0
Peers		511	3,479	736	205	379	8,779	0	0
Index		--	--	--	--	--	--	--	--

Congestive heart failure

Total Specialty Episode Costs: \$2,818

Cost per Episode	# of Episodes	Total	Primary Care Core	Specialty Care	Laboratory	Radiology	Hospital	Pharmacy	ER
Actual	1	\$2,817.56	\$0.00	\$655.48	\$28.58	\$682.19	\$384.57	\$0.00	\$1,066.73
Peers		\$1,496.61	\$27.44	\$714.02	\$20.78	\$106.20	\$314.81	\$286.36	\$26.99
Index			--	--	--	--	--	--	--

Encounters per 1000 Episodes

Actual		0	4,000	100	1,000	1,000	0	500	500
Peers		854	3,447	349	243	269	8,881	41	41
Index		--	--	--	--	--	--	--	--

Hyperlipidemia, other

Total Specialty Episode Costs: \$13,932

Cost per Episode	# of Episodes	Total	Primary Care Core	Specialty Care	Laboratory	Radiology	Hospital	Pharmacy	ER
Actual	19	\$720.64	\$38.46	\$188.41	\$20.36	\$35.22	\$0.00	\$421.22	\$16.97
Peers		\$631.67	\$28.58	\$106.52	\$34.61	\$37.56	\$9.55	\$409.05	\$5.80
Index			1.35	1.77	0.59	0.94	0.00	1.03	2.93

Encounters per 1000 Episodes

Actual			719	1,748	719	52	0	3,879	52
Peers			581	1,180	788	60	13	4,203	5
Index			1.24	1.48	0.91	0.86	0.00	0.92	11.35

Hypertension

Total Specialty Episode Costs: \$67,221

Cost per Episode	# of Episodes	Total	Primary Care Core	Specialty Care	Laboratory	Radiology	Hospital	Pharmacy	ER
Actual	43	\$1,569.36	\$88.65	\$760.21	\$27.68	\$311.39	\$7.03	\$324.61	\$49.79
Peers		\$1,228.51	\$75.29	\$468.78	\$19.68	\$188.49	\$148.75	\$266.33	\$61.20
Index			1.18	1.62	1.41	1.65	0.05	1.22	0.81

Encounters per 1000 Episodes

Actual			1,474	4,513	275	533	47	7,891	47
Peers			1,401	3,557	298	364	156	7,021	46
Index			1.05	1.27	0.92	1.46	0.30	1.12	1.02

Ischemic heart disease

Total Specialty Episode Costs: \$13,605

Cost per Episode	# of Episodes	Total	Primary Care Core	Specialty Care	Laboratory	Radiology	Hospital	Pharmacy	ER
Actual	9	\$1,511.63	\$160.14	\$759.84	\$7.31	\$381.47	\$0.00	\$202.87	\$0.00
Peers		\$2,378.04	\$45.89	\$672.60	\$29.37	\$278.61	\$978.17	\$288.30	\$85.11
Index			--	--	--	--	--	--	--

Encounters per 1000 Episodes

Actual			2,935	5,500	176	611	0	3,667	0
Peers			1,218	5,527	684	613	541	5,077	106
Index			--	--	--	--	--	--	--

Valvular disorder

Total Specialty Episode Costs: \$11,319

Cost per Episode	# of Episodes	Total	Primary Care Core	Specialty Care	Laboratory	Radiology	Hospital	Pharmacy	ER
Actual	14	\$818.25	\$17.60	\$679.04	\$0.90	\$106.43	\$10.24	\$4.04	\$0.00
Peers		\$1,047.19	\$32.37	\$590.16	\$14.37	\$108.66	\$179.66	\$61.34	\$60.62
Index			0.54	1.15	0.06	0.98	0.06	0.07	0.00

Encounters per 1000 Episodes

Actual			428	3,217	145	217	72	289	0
Peers			828	3,654	448	225	245	1,854	61
Index			0.52	0.88	0.32	0.96	0.29	0.16	0.00

Member Quality Non-Compliance List

Member ID	Member Name	Date of Birth	Gender	Age	Condition	Case	Rule
0231115813		3/25/1957	M	49	Cardiology	HTN	Pt(s) taking an NSAID med.
1576957219		9/21/1956	M	50	Cardiology	HTN	Pt(s) that had a serum creatinine in last 12 rpt mos.
1722584502		3/16/1959	F	47	Cardiology	HTN	Pt(s) taking an NSAID med.
3510814590		8/22/1968	M	38	Endocrinology	Hyperlipidemia	Pt(s) w/ the most recent HDL result \geq 40mg/dL.
5095625983		1/7/1951	F	55	Endocrinology	Hyperlipidemia	Pt(s) w/ the most recent LDL result $<$ 160mg/dL.
5095625983		1/7/1951	F	55	Endocrinology	Hyperlipidemia	Pt(s) w/ the most recent HDL result \geq 40mg/dL.
6189711566		7/4/1953	M	53	Cardiology	HTN	Pt(s) taking an ACE-inhibitor, angiotensin II receptor antagonist, diuretic, or aldosterone receptor blocker that had a serum K+ in last 12 rpt mos.
6189711566		7/4/1953	M	53	Cardiology	HTN	Pt(s) that had a serum creatinine in last 12 rpt mos.
7310373120		4/9/1960	M	46	Endocrinology	Hyperlipidemia	Pt(s) w/ the most recent HDL result \geq 40mg/dL.
8090910733		6/10/1963	F	43	Cardiology	HTN	Pt(s) taking an ACE-inhibitor, angiotensin II receptor antagonist, diuretic, or aldosterone receptor blocker that had a serum K+ in last 12 rpt mos.
8090910733		6/10/1963	F	43	Cardiology	HTN	Pt(s) that had a serum creatinine in last 12 rpt mos.
8365387487		11/5/1952	M	54	Endocrinology	Hyperlipidemia	Pt(s) w/ the most recent HDL result \geq 40mg/dL.
8577199106		6/16/1948	M	58	Cardiology	HTN	Pt(s) taking an ACE-inhibitor, angiotensin II receptor antagonist, diuretic, or aldosterone receptor blocker that had a serum K+ in last 12 rpt mos.
8577199106		6/16/1948	M	58	Cardiology	HTN	Pt(s) that had a serum creatinine in last 12 rpt mos.

Report Introduction and Interpretation

Patterns of Care

Episode Case Mix Summary

Panel Morbidity - Peer Distribution

Quality

Cost and Use

Episode Detail

Member Quality Non-Compliance

Information from Measure Evaluation

Measure Number and Name: ETG Based Chronic Pneumonia (#1611)

Description:

Measure Developer: Ingenix

Summary Assessment

The clinical and construction logic of the measure is described in detail. Measure users should be able to implement it using the description provided.

The reliability of both the data and the measure score has been established. Face validity is also established

The measure has only been tested in a commercial database → it can be endorsed for use in commercial populations.

The measure is submitted for implementation in:

Clinician : Group/Practice

Clinician : Individual

Clinician : Team

Facility

Health Plan

Integrated Delivery System

Population : Community

Population : County or City

Population : National

Population : Regional

Population : State

There are some items that merit attention:

- a) The statistical model and process used to derive weights from the model needs to be presented in more detail including all the appropriate measures of goodness of fit and calibration. The additional information provided does not assess the calibration of the model.
- b) Measure developers should have performed some sort of split-sample validation of statistical models.

Reliability (2a)

2a1. Is the measure well defined and precisely specified?

- a) Measure clinical logic described? Yes ☒ No ☐
- b) Measure construction logic described? Yes ☒ No ☐
- c) Risk-adjustment methodology described? Yes ☒ No ☐
- d) Is the data derivation process described in sufficient detail for users to implement the measure?

- i. Target population and data sources identified
 - ii. Measure specific target conditions and events identified
 - iii. Data elements and outcome variable(s) clearly defined
 - iv. Measurement windows, exclusions, risk adjustment methodology clearly defined and explained
- a) The description of the measure clinical logic is complete and exhaustive. The clinical care episode is defined using Episode Treatment Groups (ETG).
 - b) The measure construction logic is described in detail. The presentation is clear and organized including a description of how the data should be prepared and organized by potential measure users. A short list of pitfalls to avoid is also included.
 - c) The risk adjustment methodology is described but there is insufficient level of detail about the specific techniques used. It is mentioned that a multivariate regression was used but no statistics showing calibration and goodness of fit were presented. No information is given about how the thresholds to define severity groups were chosen.
 - d) The data derivation process is described in detail:
 - i. Data sources for measure users are administrative medical and pharmacy claims, member enrollment and demographic information. Tables are provided to implement to map codes to ETGs. The target population is identified as ETG 437400. This should represent all Pneumonia episodes of care (no age restrictions).
 - ii. The target condition for this measure is Pneumonia. The events associated with the target condition are also identified through the ETG methodology.
 - iii. A data dictionary is provided. The outcome variables are well defined as total cost per episode and measures of utilization per 1,000 episodes.
 - iv. The measurement window is 1 year worth of data. Exclusions are explained. The risk adjustment methodology needs more detail.

2a2 Reliability Testing

Data Reliability

- a) Was data reproducibility assessed?

Yes. The measure developers assessed data reproducibility by performing parallel development of the ETG and resource use calculations using two independent software approaches. The ETG and Resource Use Measure software results were compared to the results obtained from a SAS prototype developed using the exact same specifications. This analysis was performed on a sample of 4 million members.

The measure developers observed a matching rate of over 99.9% between the two approaches. A match is defined as exact agreement of the grouping of records and assignment of resource use.

Measure Score Reliability

- a) Measure score reliability tested (signal-to-noise ratio analysis by means of ANOVA, Intra-class Correlation Coefficient or other means)

Yes. The measure score was an integral part of the data reliability analysis described above.

Validity (2b)

2b1 Is there evidence presented that the measure specifications allow to demonstrate variations in resource use across providers and/ or population groups? Does the measure and risk-adjustment methodology address this variability allowing for fair comparisons?

2b2 Validity Testing

Data Elements

- a) Has the data been compared to other authoritative data sources? (Other databases, literature, etc.)

There is no comparison to similar independent claims databases. A comparison of the distribution of important variables to the literature could not be found.

- b) Data integrity checked? (e.g. Percent of missing values, missing diagnosis codes, inconsistent dates, range checks, etc.)

No evidence of checking for data integrity was found. There is no mention of any checks performed during measure development. The measure steward does recommend that users of the measure perform their own data integrity checks.

- c) Is the data representative of the target population?

Unclear. The main source of data is the Ingenix National health care services benchmark database. This is a large database with information on providers and medical and pharmacy services for a population of more than 25 million covered lives. There were no age restrictions in the definition of the measure. The database may not be representative of the over 65 population.

Measure Score

- a) Has the measure score validity been shown? (By correlating to another valid indicator, or showing that it produces different results when applied to subgroups known to have differences in resource use or by expert opinion or other methods)

Yes. A table is provided showing cost and utilization measures for different lines of service across severity levels for three different peer group definitions. The majority of cost is attributed to the specialty care followed by primary care in two of the peer groups (family practices and internal medicine); for the pediatrics peer group, the primary care costs are larger than the specialty care costs. Costs are positively correlated with severity level.

2b3 Are exclusions supported by clinical evidence?

- a) Has a sensitivity analysis been performed of the measure with and without the exclusions in terms of distribution of the outcome and number of patients affected?

No. There are exclusions related to data completeness and the presence of low outliers. In the latter case, an analysis is presented of their effect on the percent of episodes eligible for attribution for 9 HCOs, but not of their effect on the outcome. The distribution of the scores outlier status appears to be part of their testing plan, but the results are not shown.

- b) Are the reasons for exclusions properly addressed?

No. Patients considered low outliers are excluded from consideration. High outliers are included and their values are winsorized. The developers provide the thresholds for an observation to be considered a low or high, but do not explain how the cut-offs were selected. There is no explanation for the different treatment of high and low outliers.

- c) Are any of the exclusions based on patient preferences?

No

2b4 Is the measure risk-adjusted? If not, is there a rationale that supports no risk-adjustment/risk stratification?

- a) Is the risk-adjustment methodology described completely and accurately?

No. There are 3 stages in the development of the methodology. First, a multivariate regression model is fit where the dependent variable is the cost per episode. The list of predictors includes age, gender and comorbidities. Second, the coefficients of the model are used to derive weights which were then used to calculate a severity score. Third, the severity score is categorized into 4 groups which define four severity levels.

All steps are described, but more detail is needed about the fit and calibration of the model, the derivation of the weights and the categorization of the severity score into severity levels.

- b) If a statistical model was used, is it appropriate for the problem at hand?

Unclear. The developers did not provide sufficient detail about the modeling process.

- c) Candidate and final variable selection adequately described

It was not clear what the number of candidate variables was in the original regression model. No variable selection process is mentioned.

- d) Summary indicators of model fit, calibration and discrimination if appropriate provided

The information presented is insufficient to evaluate the fit and calibration of the model. R-squares should be calculated at the episode level and calibration curves should be provided.

e) Risk factors identified make clinical/practical sense

The risk factors are the sex-age groups, condition status, comorbidities and comorbidity interactions. Their appropriateness cannot be judged without more information about how they were derived.

f) Missing data/imputation methodology explained.

None used.

g) The model validates when applied to a new dataset (i.e., no overfitting)

Not tested

h) How are influential observations handled?

Low outliers are excluded. High outliers are winsorized.

2b5 Risk factors identified are associated with statistically significant and clinically meaningful differences

a) Are issues of statistical vs. practical significance addressed?

No

2b6 Demonstration that the method produces comparable results in different data sources

a) Does the method produce expected results when applied to different databases accounting for the differences in databases (e.g., an option to use administrative **or** medical record data)?

The method did not provide options for different data sources.

2c Are identified disparities in care being used as risk factors?

Factors that identify groups with differences/inequalities in care (race, socioeconomic status, gender, etc.) should not be part of the risk-adjustment methodology

Age and sex are part of the risk adjustment models.

Other comments:

Reviewer: Carlos Alzola