NATIONAL QUALITY FORUM

Measure Submission and Evaluation Worksheet 5.0

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

NQF #: 0334 NQF Project: Pulmonary Project

(for Endorsement Maintenance Review)
Original Endorsement Date: May 15, 2008 Most Recent Endorsement Date: May 15, 2008 Last Updated Date: Jan 20, 2012

BRIEF MEASURE INFORMATION

De.1 Measure Title: PICU Severity-adjusted Length of Stay

Co.1.1 Measure Steward: Virtual PICU Systems, LLC

De.2 Brief Description of Measure: The number of days between PICU admission and PICU discharge.

2a1.1 Numerator Statement: Number of PICU days, PICU days = Number of days between PICU admission and PICU discharge

2a1.4 Denominator Statement: Discharges from the PICU (including tranfers to other units) during the time period being reported

2a1.8 Denominator Exclusions: Patients => 18 years of age

1.1 Measure Type: Outcome
2a1. 25-26 Data Source: Administrative claims, Electronic Clinical Data : Registry, Paper Records
2a1.33 Level of Analysis: Facility

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

De.3 If included in a composite, please identify the composite measure (*title and NQF number if endorsed*):
Recommend use in conjunction with 0334 (PICU Severity Adjusted
Length of Stay) and 0335 (PICU Unplanned Readmission Rate) as balancing measures.

STAFF NOTES (*issues or questions regarding any criteria*)

Comments on Conditions for Consideration:

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

1a. Specific national health goal/priority identified by DHHS or NPP addressed by the measure (*check De.5*):
5. Similar/related endorsed or submitted measures (*check 5.1*):
Other Criteria:

Staff Reviewer Name(s):

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

Importance to Measure and Report is a threshold criterion that must be met in order to recommend a measure for endorsement. All three subcriteria must be met to pass this criterion. See guidance on evidence.
Measures must be judged to be important to measure and report in order to be evaluated against the remaining criteria. (evaluation criteria)

1a. High Impact: H☐ M☐ L☐ I☐
(*The measure directly addresses a specific national health goal/priority identified by DHHS or NPP, or some other high impact*

aspect of healthcare.)

De.4 Subject/Topic Areas *(Check all the areas that apply)*: Pulmonary/Critical Care, Pulmonary/Critical Care : Critical Care
De.5 Cross Cutting Areas (*Check all the areas that apply)*:

1a.1 Demonstrated High Impact Aspect of Healthcare: High resource use, Patient/societal consequences of poor quality, Severity of illness

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

1a.3 Summary of Evidence of High Impact *(Provide epidemiologic or resource use data)*:
ICUs are a source of signficant health care cost (1,2,5) PICUs have been shown to have varying degrees of efficiency with consumption of resources that could have been provided elsewhere (3,4). This, coupled with the potential for hospital acquired infections, supports caring only for those patients in the ICU who require ICU-level care.

1a.4 Citations for Evidence of High Impact cited in 1a.3: 1. Russell LB: The role of technology assessment in cost control, in McNeil BJ, Cravalho EG (eds): Critical Issues in Medical Technology. Boston, Auburn House, 1980, pp 129-138.
2. Knaus WA, Draper EA, Wagner DP: The use of intensive care: New research initiatives and their implications for national health policy. Milbank Q
1983;61:561-583
3. Pollack MM, Ruttimann UE, Glass NL, et al: Monitoring patients in pediatric intensive care. Pediatrics 1985;76:719-724.
4. Pollack MM. Getston PR, Ruttimann UE, et al. Efficiency of intensive care. A compartive analysis of eight pediatric intensive care units. JAMA 1987; 358:1481-1486
5. Brilli RJ, Spvetz A, Branson, RD, et al. Critical care delivery in the intensive care unit: Defining clinical roles and the best practice model. Critical Care Medicine; 2001: 29 (10), 2007-2019.

1b. Opportunity for Improvement: H☐ M☐ L☐ I ☐
(*There is a demonstrated performance gap - variability or overall less than optimal performance*)

1b.1 Briefly explain the benefits (improvements in quality) envisioned by use of this measure:
Use of this measure, coupled with 0335 and 0336 as balancing measures, allows for evaluation of appropriateness of resource utilization.

1b.2 Summary of Data Demonstrating Performance Gap *(Variation or overall less than optimal performance across providers)*:
[*For Maintenance – Descriptive statistics for performance results for this measure - distribution of scores for measured entities by quartile/decile, mean, median, SD, min, max, etc.*]
Adult data has found 24% of admissions to adult ICUs were for observation only in one study (1), while 77% of admissions to an adult ICU were for monitoring alone in another study (2)

Analysis of data PICUs submitting data in Q3 2011 to the VPS system revealed a of range of severity adjusted LOS from 1.71 to 4.02 days
Pediatric critical care studies found that 27% of admissions to one PICU received no benefit beyond what could be provided elsewhere (3), while an analysis of eight PICUs demonstrated varying efficiency with a potential saving of 5.1 to 17.2% of ICU days of care through earlier discharge (4).

1b.3 Citations for Data on Performance Gap: [*For Maintenance – Description of the data or sample for measure results reported in 1b.2 including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included*]
1.Knaus WA, Draper EA, Wagner DP: The use of intensive care: New research initiatives and their implications for national health policy. Milbank Q
1983;61:561-583
2. Thibault GE, Mulley AG, Barnett CO, et al: Medical intensive care: Indications, interventions, and outcome. N Engl J Med 1980;302:938-942.
3. Pollack MM, Ruttimann UE, Glass NL, et al: Monitoring patients in pediatric intensive care. Pediatrics 1985;76:719-724.
4. Pollack MM. Getston PR, Ruttimann UE, et al. Efficiency of intensive care. A compartive analysis of eight pediatric intensive care units. JAMA 1987; 358:1481-1486

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

Population differences have not been found to be variable in pediatric intensive care therapies. A study examined whether medical resources and outcomes for children admitted to pediatric intensive care units differed according to race, gender, or insurance status. After adjustment for differences in illness severity, standardized mortality ratios and overall resource use were similar with regard to race, gender, and insurance status, but uninsured children had significantly shorter lengths of stay in the pediatric intensive care unit. Uninsured children also had significantly greater physiologic derangement on admission (mortality probability, 8.1%; 95% confidence interval [CI], 6.2-10.0) than did publicly insured (3.6%; 95% CI, 3.2-4.0) and commercially insured patients (3.7%; 95% CI, 3.3-4.1). Consistent with greater physiologic derangement, hospital mortality was higher among uninsured children than insured children.

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

Lopez A, Tilford J, Anand K, et al. Variation in pediatric intensive care therapies and outcomes by race, gender and insurance status. Ped Crit Care Med 2006; 7(1). 2-6.

1c. Evidence (*Measure focus is a health outcome OR meets the criteria for quantity, quality, consistency of the body of evidence.*)
Is the measure focus a health outcome? Yes☐ No☐ If not a health outcome, rate the body of evidence.

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

Quantity	Quality	Consistency	Does the measure pass subcriterion1c?
M-H	M-H	M-H	Yes☐
L	M-H	M	Yes☐ IF additional research unlikely to change conclusion that benefits to patients outweigh harms: otherwise No☐
M-H	L	M-H	Yes☐ IF potential benefits to patients clearly outweigh potential harms: otherwise No☐
L-M-H	L-M-H	L	No ☐

Health outcome – rationale supports relationship to at least one healthcare structure, process, intervention, or service	Does the measure pass subcriterion1c? Yes☐ IF rationale supports relationship

1c.1 Structure-Process-Outcome Relationship *(Briefly state the measure focus, e.g., health outcome, intermediate clinical outcome, process, structure; then identify the appropriate links, e.g., structure-process-health outcome; process- health outcome; intermediate clinical outcome-health outcome)*:

This measure indirectly measures process (decision making related to PICU discharge) while directly measuring PICU resource utilization.

1c.2-3 Type of Evidence *(Check all that apply)*:

Selected individual studies (rather than entire body of evidence)

1c.4 Directness of Evidence to the Specified Measure *(State the central topic, population, and outcomes addressed in the body of evidence and identify any differences from the measure focus and measure target population)*:

Measurement of ICU and PICU length of stay (LOS) is was included as a measure or focus of study in over 9000 publications in a recent Pubmed search. Risk-adjustment of LOS is an established and accepted methodology.

1c.5 Quantity of Studies in the Body of Evidence *(Total number of studies, not articles)*:

1c.6 Quality of Body of Evidence (*Summarize the certainty or confidence in the estimates of benefits and harms to patients across studies in the body of evidence resulting from study factors. Please address: a) study design/flaws; b) directness/indirectness of the evidence to this measure (e.g., interventions, comparisons, outcomes assessed, population included in the evidence); and c) imprecision/wide confidence intervals due to few patients or events*): Length of stay may vary significantly based on the severity of illness of the patient at the time of admission. Failure to adjust for patient-level severity of illness may

result in inappropriate comparison.

1c.7 Consistency of Results across Studies *(Summarize the consistency of the magnitude and direction of the effect):* There are two categories of studies measuring LOS: those that use a risk-adjustment method and those that don´t. For those studies that compare the two approaches, risk adjustment is found to be an important factor. However, short of the studies that compare methodologies, there is no way to otherwise make comparisons

1c.8 Net Benefit *(Provide estimates of effect for benefit/outcome; identify harms addressed and estimates of effect; and net benefit - benefit over harms)*:
Risk-adjustment of LOS accounts for otherwise unexplained variation within and between centers which may result in flawed interpretations of performance.

1c.9 Grading of Strength/Quality of the Body of Evidence. Has the body of evidence been graded? No

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

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

1c.12 If other, identify and describe the grading scale with definitions: per 1c.9 above no grading has been done.

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

1c.14 Summary of Controversy/Contradictory Evidence: None

1c.15 Citations for Evidence other than Guidelines*(Guidelines addressed below)*:
Ruttimann UE, Pollack MM. Variability in duration of stay in pediatric intensive care units: A multiinstitutional study. The Journal of Pediatrics;1996:128(1), 35-43.

Straney L, Clements A, Slater A. Quantifying variation of paediatric length fo stay among intensive care units in Australia and New Zealand. Qual Saf Health Care. 2010. 19 1-5

Starney LD, Clement A, Alexander J, Slater A. Measuring efficiency in Australian and New Zealand paediatric intensive care units. Intensive Care Med. 2010; 36(8):1410-16

Niskanen M, Reinikainen M, Pettila V. Case-mix-adjusted length of stay and mortality in 23 Finnish ICUs. Intensive Care Med 2009; 35(6). 1060-7

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

1c.17 Clinical Practice Guideline Citation:

1c.18 National Guideline Clearinghouse or other URL: n/A

1c.19 Grading of Strength of Guideline Recommendation. Has the recommendation been graded? No

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

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

1c.22 If other, identify and describe the grading scale with definitions: per 1c.19 above no grading has been done.

1c.23 Grade Assigned to the Recommendation:

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

Based on the NQF descriptions for rating the evidence, what was the developer's assessment of the quantity, quality, and consistency of the body of evidence?
1c.25 Quantity: Moderate 1c.26 Quality: High1c.27 Consistency: High

Was the threshold criterion, *Importance to Measure and Report*, met?
(*1a & 1b must be rated moderate or high and 1c yes*) Yes☐ No☐
Provide rationale based on specific subcriteria:

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

2. RELIABILITY & VALIDITY - SCIENTIFIC ACCEPTABILITY OF MEASURE PROPERTIES

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. (evaluation criteria)
Measure testing must demonstrate adequate reliability and validity in order to be recommended for endorsement. Testing may be conducted for data elements and/or the computed measure score. Testing information and results should be entered in the appropriate field. Supplemental materials may be referenced or attached in item 2.1. See guidance on measure testing.

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

S.2 If yes, provide web page URL: https://portal.myvps.org/document/NQFMeasures.pdf

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

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

2a1.1 Numerator Statement *(Brief, narrative description of the measure focus or what is being measured about the target population, e.g., cases from the target population with the target process, condition, event, or outcome)*:
Number of PICU days, PICU days = Number of days between PICU admission and PICU discharge

2a1.2 Numerator Time Window *(The time period in which the target process, condition, event, or outcome is eligible for inclusion)*:
Submitted quarterly for all discharges during that time period

2a1.3 Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, codes with descriptors, and/or specific data collection items/responses*:
All patients < 18 years of age

2a1.4 Denominator Statement *(Brief, narrative description of the target population being measured)*:
Discharges from the PICU (including tranfers to other units) during the time period being reported

2a1.5 Target Population Category *(Check all the populations for which the measure is specified and tested if any)*: Children's Health

2a1.6 Denominator Time Window *(The time period in which cases are eligible for inclusion)*:
Submitted quarterly for all discharges during that time period

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

2a1.8 Denominator Exclusions *(Brief narrative description of exclusions from the target population)*:
Patients => 18 years of age

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

2a1.10 Stratification Details/Variables *(All information required to stratify the measure results including the stratification variables, codes with descriptors, definitions, and/or specific data collection items/responses)*:
Risk-adjustment using approved severity of illness tool.

2a1.11 Risk Adjustment Type *(Select type. Provide specifications for risk stratification in 2a1.10 and for statistical model in 2a1.13)*: Statistical risk model 2a1.12 If "Other," please describe:

2a1.13 Statistical Risk Model and Variables *(Name the statistical method - e.g., logistic regression and list all the risk factor variables. Note - risk model development should be addressed in 2b4.)*:
Selection criteria for risk adjustment tool for pediatric ICU's:
- Tool must allow quality assessment and comparison between intensive care units, and must be widely used
- Tool must be valid and reliable for severity adjustment and measurement of quality of care provided
- Computation of mortality risk must be in the public domain (i.e. free ofcharge)
- Algorithms must receive ongoing validation and recalibration

The PRISM 3 model meets these criteria.

1. Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated pediatric risk of mortality score. Crit Care Med 1996;24:743-52.

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

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

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

2a1.20 Calculation Algorithm/Measure Logic*(Describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; aggregating data; risk adjustment; etc.)*:
Numerator of number of days between PICU admission and PICU discharge is determined.
All discharges including transfer from PICU are counted for same time period to serve as denominator.

Risk stratification addressed using PRISM 3 methodology.

PRISM 3 is a valid, realiable and internationally accepted risk measurement tool. The methodology and measure specifications have been published(1) and are available at https://portal.myvps.org/document/NQFMeasures.pdf
1. Pollack MM, Patel KM, Ruttimann UE. PRISM III: an updated pediatric risk of mortality score. Crit Care Med 1996;24:743-52.

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

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

2a1.25 Data Source *(Check all the sources for which the measure is specified and tested*). If other, please describe:
Administrative claims, Electronic Clinical Data : Registry, Paper Records

2a1.26 Data Source/Data Collection Instrument (*Identify the specific data source/data collection instrument, e.g. name of database, clinical registry, collection instrument, etc.*): No mandatory data source or collection instrument for PICU community. Potential resources include PICU-specific databases or the VPS database (myvps.org).

Thus, 2a1.27 and 2a1.30 are not applicable

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

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

2a1.33 Level of Analysis (*Check the levels of analysis for which the measure is specified and tested)*: Facility

2a1.34-35 Care Setting *(Check all the settings for which the measure is specified and tested)*: Hospital/Acute Care Facility

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

2a2.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
Severity adjusting of LOS is an established method which has been demonstrated to be both valid and reliable, in addition to superior to unadjusted (raw) LOS. Further measure testing is not indicated.

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

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

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

2b1.1 Describe how the measure specifications *(measure focus, target population, and exclusions)* are consistent with the evidence cited in support of the measure focus (*criterion 1c)* and identify any differences from the evidence:
Exclusion criteria for the severity adjustment tool (PRISM 3) assures accuracy of severity adjusted LOS calculation.

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

2b2.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
Severity adjusting of LOS is an established method which has been demonstrated to be both valid and reliable, in addition to superior to unadjusted (raw) LOS. Further measure testing is not indicated.

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

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

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

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

2b3.1 Data/Sample for analysis of exclusions *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
The endorsed severity adjustment methodology used for calculating severity adjusted LOS excludes:
• PICU patients >=18 years of age
• PICU patients under the age of 18 years with a stay < 2 hours in the PICU or < 2 consecutive sets of vital signs consistent with life
• Patients admitted to PICU for palliative care
• Preterm infants post-gestational age 36 weeks

Palliative cases are excluded because the intention is that the patient will likely die in the ICU, skewing SMR calculations if the patients are included.

The other exclusions are consistent with the use of the PRISM 3 instrument for severity adjustment. The tool has not been validated in patients <36 weeks gestation, > or equal to 18 years, or if not in the PICU at least two hours/for two vital signs to be taken.

Further validation of these exclusions is not indicated.

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

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

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

2b4.1 Data/Sample *(Description of the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
PRISM 3 is a valid, realiable and internationally accepted risk measurement tool. The methodology and measure specifications have been published(1) and are available at https://portal.myvps.org/document/NQFMeasures.pdf

Calibration reassessment has been performed with plans for future model enhancement

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

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

2b4.4 If outcome or resource use measure is not risk adjusted, provide rationale and analyses to justify lack of adjustment:

2b5. Identification of Meaningful Differences in Performance. (*The performance measure scores were appropriately analyzed*

and discriminated meaningful differences in quality.)

2b5.1 Data/Sample *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
No sampling was done. However, the data available from the VPS system reveals that the severity adjusted length of stay among 80 participating PICUs ranged from 1.71 to 4.02 days in the third quarter of 2011. This indicates that there is unit specific variance. As numerators, denominators and all definitions are standardized with an IRR >96%, this variation reflects differences in care and not the measurement itself.

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

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

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

2b6.1 Data/Sample *(Describe the data or sample including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included)*:
Existing literature identifies the shortcomings of not severity adjusting LOS. Thus no attempt to report unadjusted LOS has been made.

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

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

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

2c.1 If measure is stratified for disparities, provide stratified results *(Scores by stratified categories/cohorts)*: N/A. This is consistent with published literature.

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

2.1-2.3 Supplemental Testing Methodology Information:

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

If the Committee votes No, STOP

3. USABILITY

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

C.1 Intended Purpose/ Use *(Check all the purposes and/or uses for which the measure is intended)*: Public Reporting, Quality Improvement (Internal to the specific organization)

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

3a. Usefulness for Public Reporting: H☐ M☐ L☐ I ☐
(*The measure is meaningful, understandable and useful for public reporting.*)

3a.1. Use in Public Reporting - disclosure of performance results to the public at large *(If used in a public reporting program, provide name of program(s), locations, Web page URL(s)).* If not publicly reported in a national or community program, state the reason AND plans to achieve public reporting, potential reporting programs or commitments, and timeline, e.g., within 3 years of endorsement: [*For Maintenance – If not publicly reported, describe progress made toward achieving disclosure of performance results to the public at large and expected date for public reporting; provide rationale why continued endorsement should be considered.*]
We are aware of no instances of public reporting of this measure at present. There is no barrier beyond interest of national or community programs for this data.

3a.2.Provide a rationale for why the measure performance results are meaningful, understandable, and useful for public reporting. If usefulness was demonstrated (e.g., focus group, cognitive testing), describe the data, method, and results: There is potential to enhance reporting of LOS data by eliminating misleading variation caused by not accounting for patient-level severity of illness. While not immediately intuitive, the concept of risk adjustment is widely used in other settings,e.g., standardized mortality ratio, and thus there exists no reason why this measure should or could not be used for public reporting.

3.2 Use for other Accountability Functions (payment, certification, accreditation). If used in a public accountability program, provide name of program(s), locations, Web page URL(s): None that we are aware of.

3b. Usefulness for Quality Improvement: H☐ M☐ L☐ I ☐
(*The measure is meaningful, understandable and useful for quality improvement.*)

3b.1. Use in QI. If used in quality improvement program, provide name of program(s), locations, Web page URL(s):
[*For Maintenance – If not used for QI, indicate the reasons and describe progress toward using performance results for improvement*].
This approach is used internally by multiple PICUs participating in the VPS clinical database. For instance, the 72-bed PICU at the Children´s Hospital of Wisconsin trends severity-adjusted LOS as part of its improvement program

3b.2. Provide rationale for why the measure performance results are meaningful, understandable, and useful for quality improvement. If usefulness was demonstrated *(e.g., QI initiative),* describe the data, method and results:
Use of severity adjusted LOS for QI purposes is superior to using an unadjusted LOS because of the elimination of misleading variation caused not accounting for by patient-level severity of illness.

Overall, to what extent was the criterion, *Usability*, met? H☐ M☐ L☐ I ☐
Provide rationale based on specific subcriteria:

4. FEASIBILITY

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

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

4a.1-2 How are the data elements needed to compute measure scores generated? *(Check all that apply).*
Data used in the measure are:
generated by and used by healthcare personnel during the provision of care, e.g., blood pressure, lab value, medical condition,
Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

4b. Electronic Sources: H☐ M☐ L☐ I ☐

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

4b.2 If ALL data elements are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using other than electronic sources: All necessary data may be available electronically if an organization has implemented an EHR. In the absence of an EHR, manual data collection would be required.

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

4c.1 Identify susceptibility to inaccuracies, errors, or unintended consequences of the measurement identified during testing and/or operational use and strategies to prevent, minimize, or detect. If audited, provide results:
Manual data abstraction with entry into a multi-institutional clinical PICU database (the VPS (myvps.org) has been completed for the variables used in this measure since 2002. Currently, 99 hospitals and 117 PICUs are abstracting and entering data with an aggregate interrater reliability of 96.78%

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

A.2 Please check if either of the following apply (*regarding proprietary measures*):
4d.1 Describe what you have learned/modified as a result of testing and/or operational use of the measure regarding data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, other feasibility/implementation issues *(e.g., fees for use of proprietary measures)*:
The data necessary for capturing length of stay in relatively simple and not burdensome. Data collection for the severity adjustment component is not significant but quite feasible. For instance, the group of 99 hospitals and 117 PICUs using the VPS database have collected these elements for more than 470,000 patient encounters between 2002 and 3rd Quarter of 2011

Finally, the elements needed for determining the SMR denominator are also used in NQF measure 0343

Overall, to what extent was the criterion, *Feasibility*, met? H☐ M☐ L☐ I ☐
Provide rationale based on specific subcriteria:

OVERALL SUITABILITY FOR ENDORSEMENT

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

If the Committee votes No, STOP.
If the Committee votes Yes, the final recommendation is contingent on comparison to related and competing measures.

5. COMPARISON TO RELATED AND COMPETING MEASURES

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

5.1 If there are related measures *(either same measure focus or target population)* or competing measures *(both the same measure focus and same target population*), list the NQF # and title of all related and/or competing measures:
0702 : Intensive Care Unit (ICU) Length-of-Stay (LOS)

5a. Harmonization

5a.1 If this measure has EITHER the same measure focus OR the same target population as NQF-endorsed measure(s): Are the measure specifications completely harmonized? Yes

5a.2 If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden:

5b. Competing Measure(s)
5b.1 If this measure has both the same measure focus and the same target population as NQF-endorsed measure(s): Describe why this measure is superior to competing measures *(e.g., a more valid or efficient way to measure quality);* OR provide a rationale for the additive value of endorsing an additional measure. *(Provide analyses when possible)*: The measures have complementary target populations with population-specific (and appropriate) risk adjustment tools. The requirement of population-specific tools precludes use of only one measure.

CONTACT INFORMATION
Co.1 Measure Steward (Intellectual Property Owner): Virtual PICU Systems, LLC, 4470 W Sunset Blvd, Suite 440, Los Angeles, California, 90027 Co.2 Point of Contact: Christine, Gall, cgall@myvps.org, 262-439-9640-
Co.3 Measure Developer if different from Measure Steward: NACHRI (Pedi-QS), 401 Wythe Street, Alexandria, Virginia, 22314 Co.4 Point of Contact: Ellen, Schwalenstocker, PhD, eschwalenstocker@nachri.org, 703-797-6045-
Co.5 Submitter: Christine, Gall, cgall@myvps.org, 262-439-9640-, Virtual PICU Systems, LLC
Co.6 Additional organizations that sponsored/participated in measure development: National Association of Children´s Hospitals and Related Institutions, Child Health Corporation of America, Medical Management Planning, VPS
Co.7 Public Contact: Christine, Gall, cgall@myvps.org, 262-439-9640-, Virtual PICU Systems, LLC

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.
Ad.2 If adapted, provide title of original measure, NQF # if endorsed, and measure steward. Briefly describe the reasons for adapting the original measure and any work with the original measure steward:
Measure Developer/Steward Updates and Ongoing Maintenance Ad.3 Year the measure was first released: 2008 Ad.4 Month and Year of most recent revision: Ad.5 What is your frequency for review/update of this measure? 3 years Ad.6 When is the next scheduled review/update for this measure? 01, 2012
Ad.7 Copyright statement:
Ad.8 Disclaimers:
Ad.9 Additional Information/Comments:
Date of Submission (*MM/DD/YY*): 10/18/2011

EXHIBIT C
FEES OR CHARGES TO OTHER ORGANIZATIONS

VPS Participant Fee Schedule

Total Annual Unit Admissions	Annual Unit VPS Participation Fee
<500	$15,625
500-999	$18,750
1,000-1,499	$21,875
1,500-2,000	$25,000
2,000	$31,250

VPS participation fees are based upon a sliding scale of unit admissions for a calendar year, and a one-time unlimited license fee of $13,000.

For each additional unit:

- The one-time license fee for a second unit will be adjusted dependent upon the continuity of the data collection staff across both participating units:
    - Same data collection teams: $13,000 (1st unit) + $11,700 (2nd unit, 10% adjustment) = $24,700 one-time fee
    - Different data collection teams: $13,000 (1st unit) + $12,350 (2nd unit, 5% adjustment) = $25,350 one-time fee
- The ongoing annual participation fee for a second unit (whichever is smaller) will be discounted 15%