NATIONAL QUALITY FORUM
Driving measurable health improvements together

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through NQF's Consensus Development Process (CDP). The information submitted by measure developers/stewards is included after the Brief Measure Information, Preliminary Analysis, and Pre-meeting Public and Member Comments sections.

To navigate the links in the worksheet: Ctrl + click link to go to the link; ALT + LEFT ARROW to return

Brief Measure Information

NQF #: 3450
Corresponding Measures:
Measure Title: Practice Environment Scale – Nursing Work Index (PES-NWI) (composite and five subscales) (previously NQF #0206)
Measure Steward: University of Pennsylvania, Center for Health Outcomes and Policy Research
sp.02. Brief Description of Measure: Practice Environment Scale-Nursing Work Index (PES-NWI) is a survey-based measure of the nursing practice environment completed by staff registered nurses; includes mean scores on index subscales and a composite mean of all subscale scores.
1b.01. Developer Rationale: The dissemination of the PES-NWI nationally and internationally assures that nurses' practice environments will be measured in consistent fashion across different health systems to develop evidence guiding policy and management decisions. The benefit of using the PES-NWI measure for health care organizations is that organizations provide better quality patient care through improved work environments.

sp.12. Numerator Statement: Continuous Variable Statement: For surveys completed by Registered Nurses (RN):
12a) Mean score on a composite of all subscale scores
12b) Mean score on Nurse Participation in Hospital Affairs (survey item numbers 5, 6, 11, 15, 17, 21, 23, 27, 28)
12c) Mean score on Nursing Foundations for Quality of Care (survey item numbers 4, 14, 18, 19, 22, 25, 26, 29, 30, 31)
12d) Mean score on Nurse Manager Ability, Leadership, and Support of Nurses (survey item numbers 3, 7, 10, 13, 20)
12e) Mean score on Staffing and Resource Adequacy (survey item numbers 1, 8, 9, 12)
12f) Mean score on Collegial Nurse-Physician Relations (survey item numbers 2, 16, 24)
12g) Three category variable indicating favorable, mixed, or unfavorable practice environments: favorable = four or more subscale means exceed 2.5; mixed = two or three subscale means exceed 2.5; unfavorable = zero or one subscales exceed 2.5.
sp.14. Denominator Statement: Staff RNs

sp.16. Denominator Exclusions: Not applicable

Measure Type: Structure
sp.28. Data Source: Instrument-Based Data
sp.07. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: 8/5/2009
Most Recent Endorsement Date: 6/11/2019

Preliminary Analysis: Maintenance of Endorsement

To maintain NQF endorsement, endorsed measures are evaluated periodically to ensure that the measure still meets the NQF endorsement criteria ("maintenance"). The emphasis for maintaining endorsement is focused on how effective the measure is for promoting improvements in quality. Endorsed measures should have some experience from the field to inform the evaluation. The emphasis for maintaining endorsement is noted for each criterion.

Criteria 1: Importance to Measure and Report

1a. Evidence

Maintenance measures – less emphasis on evidence unless there is new information or change in evidence since the prior evaluation.

1a. Evidence. The evidence requirements for a ***structure, process or intermediate outcome*** measure are that it is based on a systematic review (SR) and grading of the body of empirical evidence where the specific focus of the evidence matches what is being measured. For measures derived from patient report, evidence also should demonstrate that the target population values the measured process or structure and finds it meaningful.

The developer provides the following description for this measure:

- This is a maintenance, instrument-based, structure measure that is specified at the facility-level of analysis.
- The developer reports that this measure is a survey-based measure of the nursing working environment. The survey is completed by staff registered nurses and includes mean scores on index subscales and a composite mean of all subscale scores.
- The developer provides a logic model that depicts four concept ovals: system, interventions, outcomes, and client. The developer indicates that the system concept encompasses individual, organization, and group, and the "client" concept encompasses individual, family, and community. The ovals are connected by arrows, indicating that interventions do not influence outcomes directly, but rather through their association with system and client characteristics. System and client factors do exhibit a direct association.
- However, the developer does not provide any further detail regarding how nursing work environment applies within the logic model.

The developer provides the following evidence for this measure:

- Systematic Review of the evidence specific to this measure? ☒ Yes ☐ No

- Quality, Quantity and Consistency of evidence provided? ☒ Yes ☐ No
- Evidence graded? ☐ Yes ☒ No

Summary of prior review in 2018:

- The developer provided a summary of several systematic literature reviews, including at least one review and meta-analysis of the evidence connecting hospital nurses' work environments to patient outcomes. Results of that review noted that better work environments were associated with lower odds of negative outcomes and higher odds of positive outcomes.
- The evidence supports the linkage between better nurse work environments and patient outcomes including—but not limited to—mortality, readmissions, complications, infections, and nurse-rated quality and safety.
- During the Standing Committee review, the Standing Committee commented that culture in the workplace is not mentioned in the PES-NWI. The developer agreed culture in the workplace is important and referenced selective items in the Agency for Healthcare Research & Quality's Surveys on Patient Safety Culture. The developer noted, however, that these tools are not directly targeted towards nurses.
- Overall, the Standing Committee agreed that the measure meets the evidence criterion.

Changes to evidence from last review

☐ The developer attests that there have been no changes in the evidence since the measure was last evaluated.

☒ The developer provided updated evidence for this measure:

- In the current submission, the developer reports that there are 15 new empirical publications with evidence for the PES-NWI. The developer concludes that these papers build upon previously submitted literature demonstrating the significant association of this work environment instrument with safety outcomes.
- The developer also cites three studies in support of this measure
    - One study found that improvements in work environment by one standard deviation led to a decrease in the odds of unfavorable quality care and patient safety by factors ranging from 0.82 to 0.97. The study concludes that improvements within hospital work environments, nurse staffing, and educational composition of nurses coincide with improvements in quality of care and patient safety.
    - The second study concludes that better work environments were associated with lower odds of negative nurse outcomes (average OR of 0.71), poor safety or quality ratings (average OR of 0.65), and negative patient outcomes (average OR of 0.93), but higher odds of patient satisfaction (OR of 1.16).
    - The third study is a 10-year study of 23,650 nurses surveyed in 2006 and 14,935 surveyed in 2016 in four-state survey of random samples of licensed nurses found that changes in the hospital work environment and staffing influence missed care.

Questions for the Committee:

- *Based on the evidence provided by the developer, does the Standing Committee agree the evidence supports that the subscale concepts are associated with improvements in outcomes at the facility level?*

Guidance from the Evidence Algorithm

(Box 1) Not a health outcome, but a structure (instrument-based) measure with systematic review → (Box 3) Evidence is not → (Box 7) Empirical evidence is provided but without grading → (Box 8) Developer summarizes available body of evidence → (Box 9) → Systematic review concludes moderate certainty of evidence.

Preliminary rating for evidence: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

1b. Gap in Care/Opportunity for Improvement and Disparities

Maintenance measures – increased emphasis on gap and variation

1b. Performance Gap. The performance gap requirements include demonstrating quality problems and opportunity for improvement.

- The developer references data from 16 studies published since the measure's last review in 2018. It is unclear for some of studies provided by the developer whether the results were reported for the measure as specified, as the data may not have been aggregated to the facility-level of analysis. For instance, Table 3 (page 21 of developer submission), both tables on (page 23), Table 1 (page 24) notes scores for nurses and intensive care units, but it is unclear if those scores were aggregated to the facility level.
- One study, published in 2019, provides PES-NWI rates across 642 hospitals, however the time period for these data were not reported. This study reported scores ranging from 2.51 – 2.92 across all subscales, and a score of 2.69 for the composite (mean) score. Variance around these point estimates was not provided.

Disparities

- No disparities data were provided. The developer states that disparities data are not applicable to this measure.

Questions for the Committee:

- *Is there a continued gap in care that warrants a national performance measure?*
- *No disparities data were provided, are you aware of evidence that disparities exist in this area of healthcare?*

Preliminary rating for opportunity for improvement: ☐ **High** ☐ **Moderate** ☐ **Low** ☒ **Insufficient**

Rationale: While the developer cites several studies demonstrating variation in the measure using the PEW-NWI, with sizeable standard deviations, it is not clear in the submission whether there are data after 2018 that show a continued gap.

Committee Pre-evaluation Comments:

1a. Evidence

- I am not an expert in survey methods or in this survey, so please weigh my comments accordingly. There seems to be good evidence that the survey results correlate to important quality outcomes. It is hard to determine if this is a correlation vs causation, but important regardless.
- Indirect evidence supports improved work culture improves patient safety but direct evidence lacking.
- structure measure
- structure measure

- Yes. PES-NWI is a structural measure. Since 2018, there are 15 new publications with evidence of the PES-NWI's impact on safety outcomes and the patient's experience
- This is a structure measure that has been endorsed since 2009. Nurses are frontline workers who provide direct, critical care to patients. It is therefore logical to think that a positive nursing working environment has considerable impact on the quality of care and safety as well as patient outcomes. This issue has been even more critical during the pandemic when there is a wide shortage of nursing workforce. The preliminary rating of evidence is moderate.
- The data support the need for A measure (although not strong evidence that this survey will measure what it is that will lead to change in outcomes),and updated studies have been shared (15); moderate
- na
- reasonably strong evidence linking work environment and patient outcomes
- Preliminary review moderate evidence.
- Adequate evidence supports measure.
- The measure developer has provided recent research that substantiates that improved workplace practice environment results in improved quality outcomes for patients.
- There seems to be some evidence in the literature that supports the value of a survey-based structure measure, but it's not particularly strong.

1b. Gap in Care/Opportunity for Improvement and Disparities

- I did not see good data around variability in performance
- Very unclear how this measure is performing
- insufficient evidence
- insufficient evidence
- Descriptive statistics (i.e. mean and SD) were provided from multiple studies showing performance variation. Differences were noted between hospitals with better vs. worse nursing resources, rural vs. urban and Magnet, aspiring Magnet vs. non-Magnet facilities. Develop reports that disparities data is not applicable--likely because this is a facility level measure. Evidence provided does note differences in performance between
- The performance gap is not well demonstrated at the facility level.
- Insufficient
- no concerns
- it's not clear how significant is the performance gap based on the data shown
- Insufficient
- Disparity data are minimal, and a performance gap is demonstrated by only one study.
- Not clear that there are opportunities for improvement based on this measure. The developer states that disparities data are not applicable. The studies submitted since the last review are not aggregated to facility level of analysis.
- It is unclear how much of a performance gap there is, and no data on disparities was offered. Intuitively, it makes sense that work culture and environment play a part in how care is provided and outcomes are achieved. It's just not clear how strong a link there is.

Criteria 2: Scientific Acceptability of Measure Properties

Complex measure evaluated by Scientific Methods Panel? ☐ **Yes** ☒ **No**

Evaluators: Staff

2a. Reliability: Specifications and Testing

For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented.

For maintenance measures – less emphasis if no new testing data provided.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers.

Specifications:

- The developer states there have been no changes to the specifications since the measure's last endorsement in 2018, though the measure was previously endorsed at both the facility level of analysis as well as the clinician: group/practice level but is no longer seeking endorsement at the clinician: group practice level.
- The developer reported that a minimum of 30 responses per year are required to establish a minimum sample and recommend that hospitals survey all eligible nurses.
    - It is unclear from the cited literature whether the testing data include this minimum response size.

Reliability Testing:

- In 2018, this measure was reviewed by the NQF Scientific Methods Panel (SMP). At that time, the developer provided testing data for both the Patient/Encounter level and the Accountable Entity level. The SMP gave the measure high ratings for both reliability and validity, which the Standing Committee voted to accept.
- In the recent submission, the developer provides reliability testing conducted at the Patient/Encounter Level:
    - The developer cites a meta-analysis (Zangaro and Jones, 2019) as reliability testing in which the authors reviewed 51 studies and computed Cronbach's alpha to examine the reliability of the 31 items in the instrument or the items in each of the five subscales.
    - In the 2019 meta-analysis, 30 of the 51 studies reported reliability coefficients for the sub-scales, which ranged as follows:
        - Nurse Participation in Hospital Affairs (0.71-0.93)
        - Nursing Foundations for Quality of Care (0.56-0.91)
        - Nurse Manager Ability, Leadership, and Support of Nurses (0.63-0.92)
        - Staffing and Resource Adequacy (0.65-0.88)
        - Collegial Nurse–Physician Relations (0.62-0.91)
        - Mean reliability was estimated at 0.946.

Questions for the Committee regarding reliability:

- *Do you have any concerns that the measure cannot be consistently implemented (i.e., are measure specifications adequate)?*
- *The Standing Committee should consider whether the cited studies have applied the minimum sample requirement of 30 surveys.*

Preliminary rating for reliability: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

2b. Validity: Validity testing; Exclusions; Risk-Adjustment; Meaningful Differences; Comparability; Missing Data

For maintenance measures – less emphasis if no new testing data provided.

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality.

2b2-2b6. Potential threats to validity should be assessed/addressed.

Validity Testing

- Validity testing conducted at the Accountable Entity Level:
    - The developer presents data from the literature on empirical validity testing for the hypothesized relationships of the PES-NWI to patient health outcomes and quality and safety, the effect sizes and the significance levels.
    - Research studies calculated odds ratios (OR), Beta coefficients (β), or incident rate ratios (IRR) for the association of better PEW-NWI scores and the following outcomes:
        - Patient Record Outcomes
            - 30-day Mortality OR: 0.86 ($p<0.001$) and 0.95 ($p<0.001$)
            - Failure to Rescue OR: 0.88 ($p<0.001$) and 0.94 ($p<0.001$)
            - Catheter-associated urinary tract infections β: -1.34 ($p<0.05$)
            - Restraint rate β: -0.088 ($p=0.045$)
            - Length of stay IRR: 0.96 ($p=0.003$) and 0.97 ($p<0.01$)
            - Readmission within 7 days OR: 0.96 ($p<0.05$)
            - Readmission within 30 days OR: 0.97 ($p<0.05$)
            - Percent of long-stay nursing home residents with pressure ulcers β: 0.90 ($p=0.02$)
            - Number of hospitalizations per nursing home resident year β: 0.08 ($p=0.05$)
        - Nurse Reported Adverse Outcomes
            - NR fall with injury OR: 0.50 ($p<0.01$)
            - NR medication errors OR: 0.61 ($p<0.01$)
            - NR pressure ulcer OR: 0.54 ($p<0.01$)
            - NR UTI OR: 0.61 ($p<0.01$)
        - Patient Satisfaction
            - Nurses communicated well β: 2.62 ($p<0.001$)
            - Patient rates hospital highly β: 5.13 ($p<0.002$)
            - Patient would definitely recommend β: 6.08 ($p<0.001$)
        - Nurse Reported Measures
            - Quality of End of Life Care OR: 0.45 ($p<0.001$)
    - The developer states that there are no norms for these tests but that all hypothesized relationships were confirmed.

Meaningful Differences

- No concerns. The developer uses descriptive statistics and peer comparisons for benchmarking, and descriptive and inferential statistical analyses in empirical studies for determining both statistically significant and clinically/practically meaningful differences in performance measure scores among measured entities.

Missing Data

- No concerns. The developer reports identical means and standard deviations in a comparison of nurse responses from the survey compared to a follow-up nonrespondent survey.

Questions for the Committee regarding validity:

- *Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?*

Preliminary rating for validity: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:

2a. Reliability

- 2a1. Reliability-Specifications
  - survey sampling information would be helpful -I did not see it, but may have been specified
  - A lot of concerns give that the measure is not used in a clearly representative sample
  - moderate
  - moderate
  - No concerns
  - The developer states that 30 responses a year are required to establish a minimum sample size. In the cited meta-data analysis by Zangaro and Jones (2019), their cited sample size ranging from 35 to 33,845.
  - All are clearly defined
  - no changes
  - reasonable, no concerns
  - No concerns on consistent implementation.
  - None.
  - In this round the developer provided testing for the patient/encounter level. ounter level but it is unclear if the facilities have enough respondents across the board to be
  - The reliability testing appears to be appropriate and acceptable
- 2a2. Reliability-Testing
  - I am not sure if there are enough surveys gathered at each site, if they are probability weighted and if minority opinions are sought out.
  - Yes
  - no
  - no
  - No, multiple studies site 51 with high reliability coefficients for each of the 5 domains, mean reliability 0.946.
  - A meta-analysis by Zangaro and Jones (2019) was cited as reliability testing. The mean reliability was estimated at 0.946. Preliminary rating on reliability by the NQF staff is moderate.
  - No concerns
  - no concerns
  - reasonable, no concerns

- No
- No, as long as a minimum sample size is maintained.
- It is unclear if the cited studies have applied the minimum sample requirement of 30 surveys.
- It is still unclear if minimum sample size was addressed

2b. Validity

- 2b1. Validity -Testing
  - I am concerned about selection bias. Unhappy staff and nurses are unlikely to fill out surveys. I am also not sure if the number of surveys needed was adequately powered. The developer forms were difficult to follow and sparse on the information needed
  - Yes, the results presented are not from a sample that is clearly representative of all nursing homes
  - no
  - no
  - No concerns. Validity was assessed by comparing PES-NWI to other quality outcomes (i.e. 30 day Mortality, CAUTI, Readmissions, HCAHPS (Nurse communication, Pt. rating and Pt. likelihood to recommend).
  - No concerns on testing, although performance gap could be better illustrated at the facility level.
  - No concerns
  - no concerns
  - seems reasonable
  - No
  - No.
  - The developer uses descriptive statistics and peer comparisons for bench-marking.
  - no concerns

2b2-2b6. Potential threats to validity

- 2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment)
  - I don't think risk adjustment is needed, but weighting the survey respondent characteristics probably should be done
  - Unable to assess
  - NA
  - NA
  - Risk adjust may not be appropriate for this measure.
  - No concerns.
  - Not risk adjusted
  - not risk adjusted
  - N/A
  - n/a
  - no comment.
  - n/a
  - There is no risk-adjustment
- 2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)
  - selection bias, inadequate sample size, broad based sampling, intercation with other hospital characteristics (city under-resourced vs private, elective orthopedic hospitals etc) regional differences
  - Potentially
  - no
  - no
  - No. NDNQI reports a response rate of 68% overall at the nursing unit level. 2 studies reported no significant differences across response rates
  - In a comparison, the developer shows that missing data have means and standard de insignificant impacts on means and standard deviations.

- No - very small amount of missing data
- no concerns
- it's not clear to me what is a meaningful difference on the scale
- n/a
- No issues.
- No standardized benchmarks for this data.
- No

Criterion 3. Feasibility

Maintenance measures – no change in emphasis – implementation issues may be more prominent

3. Feasibility is the extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

- The developer reported that data for the measure are generated through the Practice Environment Scale-Nursing Work Index (PES-NWI) Survey
- The developer noted that the survey can be collected through electronic survey software or via paper.
- The developer also reported that the measure is an eMeasure in the National Database of Nursing Quality Indicators, the Veterans Administration, and the Military Hospital system.
- The developer reported that a minimum of 30 responses per year are required to establish a minimum sample but recommend that hospitals survey all eligible nurses.

Questions for the Committee:

- *Are the required data elements routinely generated and used during care delivery?*
- *Are the required data elements available in electronic form, e.g., EHR or other electronic sources?*
- *Is the data collection strategy ready to be put into operational use?*

Preliminary rating for feasibility: ☐ **High** ☒ **Moderate** ☐ **Low** ☐ **Insufficient**

Committee Pre-evaluation Comments:

3. Feasibility

- It is a survey, it is unclear how onerous it is for individuals or facilities
- Not clear how feasibile it is to require this of staff, making it difficult to say if you get a sample that is representative
- none
- none
- No concerns. Requires participation in NDNQI to implement if not part of VA (???)--costly for some, but not a regulatory program.
- No concerns, and feasibility is rated as moderate.
- Low - None of data elements are routinely generated - would be another survey for staff
- no concerns
- no concerns
- Survey based. No concerns.
- The data are not routinely collected in practice but are generated by electronic survey.
- This is an emeasure that is also reported in the National Database of Nursing Quality Indicators. A minmum of 30 responses per year are required but the recommendation is for hospitals to survey all eligible nurses.

- No concerns

Criterion 4: Use and Usability

Maintenance measures – increased emphasis – much greater focus on measure use and usefulness, including both impact/improvement and unintended consequences

4a. Use (4a1. Accountability and Transparency; 4a2. Feedback on measure)

4a. Use evaluates the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4a.1. Accountability and Transparency. Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

Current uses of the measure

Publicly reported?	☒ Yes	☐ No			
Current use in an accountability program?	☒ Yes	☐ No	☐ UNCLEAR		
Planned use in an accountability program?	☐ Yes	☐ No	☒ NA		

Accountability program details

- The measure is used for public reporting, professional certification or recognition program, quality improvement with benchmarking, internal quality improvement, state of Colorado annual reporting from all hospitals with 100 or more beds, and the 2019 Magnet Hospital Accreditation program.

4a.2. Feedback on the measure by those being measured or others. Three criteria demonstrate feedback: 1) those being measured have been given performance results or data, as well as assistance with interpreting the measure results and data; 2) those being measured and other users have been given an opportunity to provide feedback on the measure performance or implementation; 3) this feedback has been considered when changes are incorporated into the measure

Feedback on the measure by those being measured or others

- The developer reports that typical feedback from about the measure is that a reduction in length and testing for use in non-hospital settings is desired.
- The developer states that the measure has not been modified to date although reduction in survey length is planned for a future endorsement period.

***Questions for the Committee**:*

- *How have (or can) the performance results be used to further the goal of high-quality, efficient healthcare?*
- *How has the measure been vetted in real-world settings by those being measured or others?*

Preliminary rating for Use: ☒ Pass ☐ No Pass

4b. Usability (4a1. Improvement; 4a2. Benefits of measure)

4b. Usability evaluates the extent to which audiences (e.g., consumers, purchasers, providers, policymakers) use or could use performance results for both accountability and performance improvement activities.

4b.1 Improvement. Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated.

Improvement results

- The developer refers to performance results provided in section 1b of the measure submission showing slight improvements in average performance year over year.
    - However, concerns exist related to whether the studies cited are providing performance results of the measure as specified, using the minimum recommended number of responders.
- The developer cites a study (Aiken et al. 2018) that reports data from a sample of 535 hospitals in which work environment, as measure by the PES-NWI, improved in 21 percent of hospitals, remained stable in 71 percent of hospitals, and worsened in eight percent.

4b2. Benefits vs. harms. Benefits of the performance measure in facilitating progress toward achieving high-quality, efficient healthcare for individuals or populations outweigh evidence of unintended negative consequences to individuals or populations (if such evidence exists).

Unexpected findings (positive or negative) during implementation

- The developer reports that scoring and reporting is not always performed consistently.
- The developer also reports variation in the unit of analysis for reporting, specifically nurse, nursing unit and organizational levels

Potential harms

- No potential harms were reported by the developer.

***Questions for the Committee**:*

- *How can the performance results be used to further the goal of high-quality, efficient healthcare?*
- *The Standing Committee should consider whether the information provided shows improvement of the measure as specified over time.*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

Preliminary rating for Usability: ☐ **High** ☐ **Moderate** ☐ **Low** ☒ **Insufficient**

Rationale:

- The developer refers to performance results provided in section 1b of the measure submission also show slight improvements in average performance year over year; however, concerns exist related to whether the studies cited are actually showing improvement on the measure over time, and are providing performance results of the measure as specified using the recommended minimum number of surveys.

Committee Pre-evaluation Comments:

4a. Use

- Some hospitals improved and some deteriorated
- Seems like it is being used by some facilities
- yes
- yes
- Facilities participating in survey received trended descriptive statistics about their performance.
- According to the developer, "the measure is used for public reporting, professional certification or recognition program, quality improvement with benchmarking, internal quality improvement, state of Colorado annual reporting from all hospitals with 100 or more beds, and the 2019 Magnet Hospital Accreditation program." There is no public reporting geared towards healthcare consumers and patients, though. I would encourage the developer to think about how healthcare consumers and patients can be informed about this information when they decide where to seek better care.
- Has not been modified despite concerns about length
- no concerns
- no concerns
- In use.
- Known to improve patient safety, although issues remain about length of survey and how the results impact changes.
- the data is not currently publicly reported or used in an accountability program however will be submitted to CMS for MIPS and MUC.
- It is being used in various accountability programs. Feedback has been considered.

4b. Usability

- minimal harm
- Unclear
- Insufficient evidence
- Insufficient evidence
- None reported.
- Using a sample of 535 hospitals, Aiken et al. (2018) found work environment improvement in 21% of hospitals and 71% remained stable, measured by the PES-NWI. Scoring and reporting although were not always performed consistently. Preliminary rating on usability is insufficient. The concern is whether such an improvement is over years and whether the performance results used the recommended minimum number of surveys. I think this is a legitimate concern.
- Did not acknowledge the harm in additional time/surveying of staff.
- no concerns
- agree with staff it's not clear if there's been improvement
- No concerns on benefits vs. harms; however, insufficient prelim rating.
- The are no apparent harms.
- The measure is publicly reported and is use for quality improvement, state of Colorado annual reporting for all hospitals with 100 or more beds, and the 2019 Magnet Accreditation Program.
- No unintended consequences. Streamlining the tool may provide more robust data for analysis.

Criterion 5: Related and Competing Measures

Related measures

- 0204: Skill mix (Registered Nurse [RN], Licensed Vocational/Practical Nurse [LVN/LPN], unlicensed assistive personnel [UAP], and contract)
- 0205: Nursing Hours per Patient Day

Harmonization

- The developer reports that the measures have been harmonized to the extent possible.

Committee Pre-evaluation Comments:

5: Related and Competing Measures

- n/a
- This seems reasonable
- 2 other measures but developer reports harmonization has occurred
- 2 other measures but developer reports harmonization has occurred
- No
- No concerns.
- Did not reflect on the difference between the skill mix/nursing hours per day and specifically what/how this adds
- Yes and yes
- 204 and 205
- No comment
- Related measure are 0204 Skill mix and 0205 Nursing hours per patient day. These do not measure the same thing.
- no concerns

Public and NQF Member Comments (Submitted as of June 14, 2022)

Member Expression of Support

- This is clarifying information as a public comment on measure #3450, which I steward. The clarifying information is submitted as part of the pre-evaluation commenting period of the spring 2022 Patient Safety Consensus Development Process.
    - Regarding the staff's preliminary ratings assigned in the Preliminary Analysis, please note the following:
    - For **Criteria 1. Importance to Measure and Report**.
        - 1a. Evidence, The Analysis notes "However, the developer does not provide any further detail regarding how nursing work environment applies within the logic model." My reply is "The work environment is considered an organizational concept within the system. The work environment is considered to moderate the relationship between an intervention and an outcome. Or stated another way, the effect of an intervention depends on the context of the work environment."
    - Under **Changes to evidence from last review,** there is an error: It states "In the current submission, the developer reports that there are 15 new empirical publications with evidence for the PES-NWI." The correct number is 35 new empirical publications.
        - 1b. Gap in Care/Opportunity for Improvement / 1b. Performance Gap: Bullet 2 states: "However, the time period for these data were not reported." The clarification is: "These data were collected in 2005 through 2008 sequentially in four large states." Bullet 2 also states "Variance around these point estimates was not provided." The clarification is: "In Lake, Riman, & Sloane (2020), Table 2 on page 2159 reports the PES-NWI mean across a panel of 458 hospitals in 2006 and 2016. The means and SDs were:

        - 2006 2.70 (0.22)
        - 2016 2.77 (0.25)
    - These values demonstrate that although the work environment has improved modestly over the ten year period, there is even greater variation in 2016 across hospitals than there was in 2006."
- 1b. Disparities. The analysis notes: "The developer states that disparities data are not applicable to this measure." The clarification is: "There is one study that demonstrates poorer PES-NWI scores in hospitals that serve disproportionately more patients of Black race:
    - 1. Lake et al (2015) in Health Services Research, in data from 2008, shows in Table 3 on page 386 this PES-NWI mean and SD distribution across a sample of 98 hospitals nationally classified into categories of low, medium, and high percentages of very low birthweight infants of Black race:
        - Low: 3.16 (0.27)
        - Medium: 3.07 (0.21)
        - High: 2.95 (0.24)
        - These differences were statistically significant p = .004.
        - I had not included this information because I am not sure if this is the proper interpretation of disparities for structure measures."
- For **Criteria 2: Scientific Acceptability of Measure Properties**
- 2.a.2. **Reliability testing**. Specifications Bullet 2 notes: "It is unclear from the cited literature whether the testing data include this minimum response size." The clarification is "In Zangaro & Jones (2019) of the 51 studies included in the reliability generalization meta-analysis, Table 2 on pages 1665 - 1667 shows a range of respondents of 35 to 33,845. Thus, all of these studies meet the minimum sample size."
- Regarding **Questions for the Committee regarding reliability:** Bullet 2 states "The Standing Committee should consider whether the cited studies have applied the minimum sample requirement of 30 surveys." The response is "see above comment: all 51 studies exceeded the minimum requirement."
- For **Criterion 4: Use and Usability**
    - 4a.1.**Accountability program details.** Here is an additional program detail that was not listed: "The Leapfrog group plans to begin surveys using the PES-NWI in the 2023 survey year for payors and health plans to include in their value-based purchasing programs."
    - 4b.1 **Improvement.** Under **Improvement Results.** Bullet 1 states "concerns exist...minimum recommended number of responders." Clarifying comment is: "The minimum was reached in 51 studies compiled for the Zangaro & Jones (2019)meta-analysis, suggesting that this minimum is routinely met."
    - 4b.2. **Benefits versus Harms** includes the statement "(if such evidence exists)." The clarifying comment is: "There has been no evidence of unintended negative consequences to individuals or populations from use of the measure."
    - **Preliminary Rating for Usability** X Insufficient
        - **Rationale** "concerns exist related to whether the studies cited are actually showing improvement on the measure over time, and are providing performance results of the measure as specified using the recommended minimum number of surveys." Clarification is "The two panel studies (Lake et al (2020); Sloane et al (2018) of 452 hospitals were designed to show changes

in the same group of hospitals over a ten year period." and "As per Zangaro & Jones (2019), we assert that the minimum is routinely met."

Scientific Acceptability Evaluation

RELIABILITY: SPECIFICATIONS

1. **Have measure specifications changed since the last review? ☒ Yes ☐ No**
2. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented? ☒ Yes ☐ No**
3. **Briefly summarize any changes to the measure specifications and/or concerns about the measure specifications.**
   - The developer states there have been no changes to the specifications since last review, but when the measure was last reviewed (as #0206) it was previously endorsed at both the facility level of analysis as well as the clinician: group/practice level but is no longer seeking endorsement at the clinician: group practice level.

RELIABILITY: TESTING

4. **Did the developer conduct new reliability testing? ☒ Yes ☐ No**

   4a. If no, summarize the Standing Committee's previous feedback:

   4b. If yes, describe any differences between the new and old testing and summarize any relevant Standing Committee feedback from the previous review:
   - A 2019 meta-analysis examined the reliability generalization of the PES-NWI using 51 studies and representing 80,563 subjects. The mean reliability was estimated at 0.946.
5. **Reliability testing level: ☐ Accountable-Entity Level ☒ Patient/Encounter Level ☐ Neither**
6. **Reliability testing was conducted with the data source and level of analysis indicated for this measure:**

o **☒ Yes ☐ No**

7. If accountable-entity level and/or patient/encounter level reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ Yes ☐ No
8. **Assess the method(s) used for reliability testing:**
   - The developer reports that reliability testing was conducted at the patient/encounter level.
   - The developer cites a meta-analysis (Zangaro, 2019) as reliability testing in which the authors reviewed 51 studies and computed Cronbach's alpha to examine the reliability of the 31 items in the instrument or the items in each of the five subscales.
9. **Assess the results of reliability testing**
   - In the 2019 meta-analysis, 30 of the 51 studies reported reliability coefficients for the sub-scales, which ranged as follows:
     - Nurse Participation in Hospital Affairs (0.71-0.93)
     - Nursing Foundations for Quality of Care (0.56-0.91)
     - Nurse Manager Ability, Leadership, and Support of Nurses (0.63-0.92)
     - Staffing and Resource Adequacy (0.65-0.88)
     - Collegial Nurse–Physician Relations (0.62-0.91)
   - Mean reliability was estimated at 0.946.

- Based on this data, the developer states that the instrument composite and subscales show high internal consistency and reliability.

10. Was the method described and appropriate for assessing the proportion of variability due to real differences among measured entities? **NOTE:** If multiple methods used, at least one must be appropriate.

    ☒ **Yes** ☐ **No** ☐ **Not applicable**

11. Was the method described and appropriate for assessing the reliability of ALL critical data elements?

    ☒ **Yes** ☐ **No** ☐ **Not applicable** (patient/encounter level testing was not performed)

12. **OVERALL RATING OF RELIABILITY** (taking into account precision of specifications and all testing results):

    ☐ **High** (NOTE: Can be HIGH only if accountable-entity level testing has been conducted)

    ☒ **Moderate** (NOTE: Moderate is the highest eligible rating if accountable-entity level testing has not been conducted)

    ☐ **Low** (NOTE: Should rate LOW if you believe specifications are NOT precise, unambiguous, and complete or if testing methods/results are not adequate)

    ☐ **Insufficient** (NOTE: Should rate INSUFFICIENT if you believe you do not have the information you need to make a rating decision)

13. **Briefly explain rationale for the rating of OVERALL RATING OF RELIABILITY and any concerns you may have with the approach to demonstrating reliability.**

    Specifications are precise, unambiguous, and complete (Box 1) → Empirical reliability testing was conducted using the measure as specified (Box 2) → Empirical validity testing of was not conducted at the accountable entity level (facility) (Box 4) → Reliability testing was conducted on all critical data elements (Box 8) → The method described was appropriate for assessing the reliability of all encounter level elements (Box 9) → Based on the reliability statistic there is high confidence that the data elements used in the measure are reliable (Box 10a) → Rate as MODERATE

    - No concerns. Moderate is the highest possible rating for this measure since testing was not conducted at the accountable entity level.

VALIDITY: TESTING

14. **Did the developer conduct new validity testing? ☒ Yes ☐ No**

    14a. If no, summarize the Standing Committee's previous feedback:

    - Not applicable.

    14b. If yes, describe any differences between the new and old testing and summarize any relevant Standing Committee's feedback from the previous review:

    - The developer compiled a table of publications that examine the hypothesized relationship of PES-NWI with specific patient health outcomes and quality and safety (including effect sizes and significance levels) to examine the predictive validity of PES-NWI.

15. **Validity testing level (check all that apply):**

    ☐ **Accountable-Entity Level** ☐ **Patient or Encounter-Level** ☐ **Both**

    NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

16. **If patient/encounter level validity testing was provided, was the method described and appropriate for assessing the accuracy of ALL critical data elements? NOTE:** Data element validation from the literature is acceptable.

    ☐ **Yes**

    ☐ **No**

☒ **Not applicable** (patient/encounter level testing was not performed)

17. **Method of establishing validity at the accountable-entity level:**

☐ **Face validity**

☒ **Empirical validity testing at the accountable-entity level**

☐ **N/A (accountable-entity level testing not conducted)**

18. **Was the method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships?**

☒ **Yes**

☐ **No**

☐ **Not applicable** (accountable-entity level testing was not performed)

19. **Assess the method(s) for establishing validity**

- The developers present data from the literature on empirical validity testing for the hypothesized relationships of the PES-NWI to patient health outcomes and quality and safety, the effect sizes and the significance levels.

20. **Assess the results(s) for establishing validity**

- Research studies calculated odds ratios (OR), Beta coefficients (β), or incident rate ratios (IRR) for the association of better PEW-NWI scores and the following outcomes:
    - Patient Record Outcomes
        - 30-day Mortality OR: 0.86 ($p<0.001$) and 0.95 ($p<0.001$)
        - Failure to Rescue OR: 0.88 ($p<0.001$) and 0.94 ($p<0.001$)
        - Catheter-associated urinary tract infections β: -1.34 ($p<0.05$)
        - Restraint rate β: -0.088 ($p=0.045$)
        - Length of stay IRR: 0.96 ($p=0.003$) and 0.97 ($p<0.01$)
        - Readmission within 7 days OR: 0.96 ($p<0.05$)
        - Readmission within 30 days OR: 0.97 ($p<0.05$)
        - Percent of long-stay nursing home residents with pressure ulcers β: 0.90 ($p=0.02$)
        - Number of hospitalizations per nursing home resident year β: 0.08 ($p=0.05$)
    - Nurse Reported Adverse Outcomes
        - NR fall with injury OR: 0.50 ($p<0.01$)
        - NR medication errors OR: 0.61 ($p<0.01$)
        - NR pressure ulcer OR: 0.54 ($p<0.01$)
        - NR UTI OR: 0.61 ($p<0.01$)
    - Patient Satisfaction
        - Nurses communicated well β: 2.62 ($p<0.001$)
        - Patient rates hospital highly β: 5.13 ($p<0.002$)
        - Patient would definitely recommend β: 6.08 ($p<0.001$)
    - Nurse Reported Measures
        - Quality of End of Life Care OR: 0.45 ($p<0.001$)
- The developer states that there are no norms for these tests but that all hypothesized relationships were confirmed.

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

21. **Please describe any concerns you have with measure exclusions.**

- There are no denominator exclusions.

22. **Risk Adjustment**

22a. **Risk-adjustment method**

☒ None (only answer Question 20b and 20e) ☐ Statistical model ☐ Stratification

☐ Other method assessing risk factors (please specify)

22b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☒ Yes ☐ No ☐ Not applicable

22c. **Social risk adjustment:**

22c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

22c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

22c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

22d.**Risk adjustment summary:**

22d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

22d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

22d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

22d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

22d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

22e**. Assess the risk-adjustment approach**

- Not applicable.

23. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**
    - No concerns. The developer uses descriptive statistics and peer comparisons for benchmarking, and descriptive and inferential statistical analyses in empirical studies for determining both statistically significant and clinically/practically meaningful differences in performance measure scores among measured entities.

24. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    - Not applicable.

25. **Please describe any concerns you have regarding missing data.**
    - No concerns. The developer reports identical means and standard deviations in a comparison of nurse responses from the survey compared to a nonrespondent survey and concludes that missing data does not impact the validity of the measure.

ADDITIONAL RECOMMENDATIONS

26. **If you have listed any concerns in this form, do you believe these concerns warrant further discussion by the multi-stakeholder Standing Committee? If so, please list those concerns below.**
    - N/A

27. **OVERALL RATING OF VALIDITY taking into account the results and scope of all testing and analysis of potential threats.**

☐ **High** (NOTE: Can be HIGH only if score-level testing has been conducted)

☒ **Moderate** (NOTE: Moderate is the highest eligible rating if score-level testing has NOT been conducted)

☐ **Low** (NOTE: Should rate LOW if you believe that there **are** threats to validity and/or relevant threats to validity were **not assessed OR** if testing methods/results are not adequate)

☐ **Insufficient** (NOTE: For instrument-based measures and some composite measures, testing at both the score level and the data element level **is required**; if not conducted, should rate as INSUFFICIENT.)

28. **Briefly explain rationale for rating of OVERALL RATING OF VALIDITY and any concerns you may have with the developers' approach to demonstrating validity.**

Threats to validity were empirically assessed (Box 1) → Empirical validity testing was conducted using the measure as specified (Box 2) → Empirical validity testing was conducted at the accountable entity level (Box 5) → Testing method was appropriate (Box 6) → Based on the empirical validity testing results, there is moderate confidence that the accountable entity data are a valid indicator of quality (Box 7b) → Rate as MODERATE.

Developer Submission

Criteria 1: Importance to Measure and Report

Extent to which the specific measure focus is evidence-based, important to making significant gains in healthcare quality, and improving health outcomes for a specific high-priority (high-impact) aspect of healthcare where there is variation in or overall less-than-optimal performance. Measures must be judged to meet all sub criteria to pass this criterion and be evaluated against the remaining criteria

1ma.01. Indicate whether there is new evidence about the measure since the most recent maintenance evaluation. If yes, please briefly summarize the new evidence, and ensure you have updated entries in the Evidence section as needed.

[Response Begins]
Yes
[Yes Please Explain]
Since the prior endorsement in Nov 2018 there are 35 new empirical publications reporting evidence from the instrument. These papers substantiate and expand the prior literature demonstrating the significant association of this work environment instrument with safety outcomes.

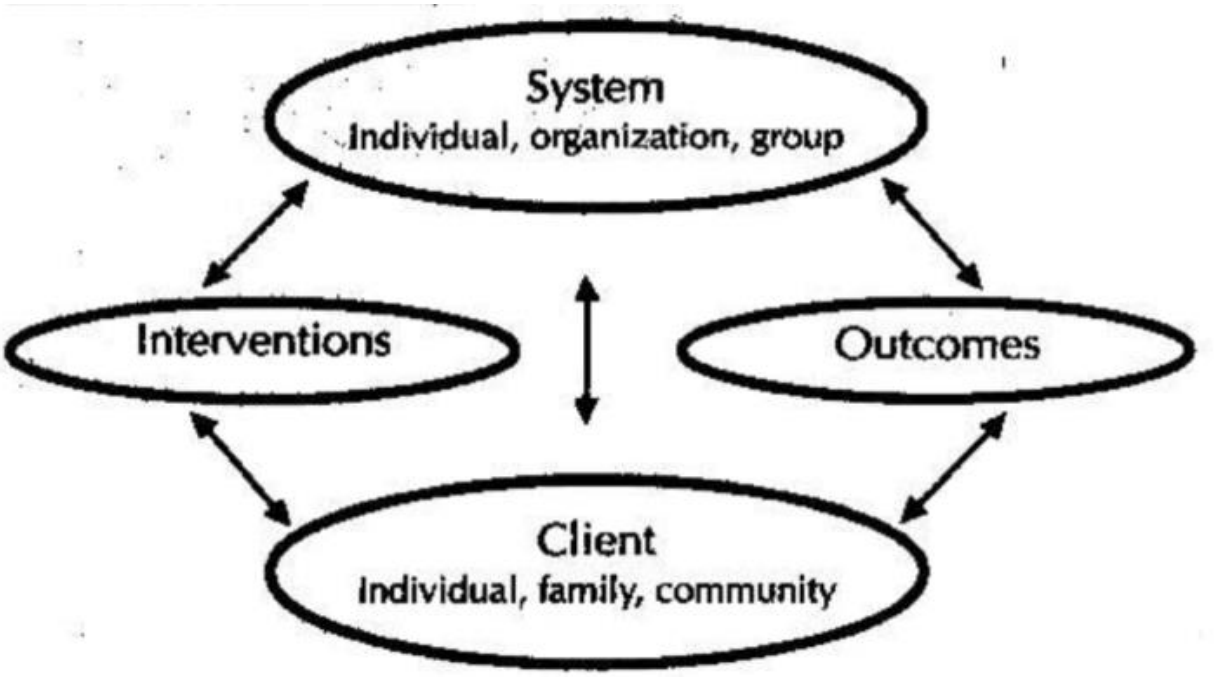
[Response Ends]

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Measure and Report: Evidence section. For example:
2021 Submission:
Updated evidence information here.
2018 Submission:
Evidence from the previous submission here.

1a.01. Provide a logic model.

Briefly describe the steps between the healthcare structures and processes (e.g., interventions, or services) and the patient's health outcome(s). The relationships in the diagram should be easily understood by general, non-technical audiences. Indicate the structure, process or outcome being measured.

[Response Begins]
We place the nursing practice environment in the system characteristics within the Quality Health Outcomes Model (Mitchell et al. 1998), which postulates that interventions are mediated by system and client characteristics in influencing health outcomes.



This image displays four concept ovals: system, interventions, outcomes, and client. System encompasses individual, organization, and group. Client encompasses individual, family, and community. The ovals are connected with arrows, indicating that interventions do not influence outcomes directly, but rather through their association with system and client characteristics. Furthermore, system and client factors do exhibit a direct association.

Mitchell, P. H., Ferketich, S., Jennings, B. M., & American Academy of Nursing Expert Panel on Quality Health Care. (1998). Quality health outcomes model. Image: Journal of Nursing Scholarship, 30(1), 43-46.

[Response Ends]

1a.02. Select the type of source for the systematic review of the body of evidence that supports the performance measure.

A systematic review is a scientific investigation that focuses on a specific question and uses explicit, prespecified scientific methods to identify, select, assess, and summarize the findings of similar but separate studies. It may include a quantitative synthesis (meta-analysis), depending on the available data.

[Response Begins]

Other (specify)

[Other (specify) Please Explain]

Published systematic reviews, meta-analysis, and longitudinal panel studies.

[Response Ends]

If the evidence is not based on a systematic review, skip to the end of the section and do not complete the repeatable question group below. If you wish to include more than one systematic review, add additional tables by clicking "Add" after the final question in the group.

Evidence - Systematic Reviews Table (Repeatable)

Group 1 - Evidence - Systematic Reviews Table

1a.03. Provide the title, author, date, citation (including page number) and URL for the systematic review.

[Response Begins]
Spring 2022 information:
Wei, H., et al. (2018). "The state of the science of nurse work environments in the United States: A systematic review." International Journal of Nursing Sciences 5(3): 287-300.

[Response Ends]

1a.04. Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the systematic review.

[Response Begins]
The systematic review concluded that "Healthier work environments lead to more satisfied nurses who will result in better job performance and higher quality of patient care, which will subsequently improve healthcare organizations' financial viability."
[Response Ends]

1a.05. Provide the grade assigned to the evidence associated with the recommendation, and include the definition of the grade.

[Response Begins]
Level 4 - case-control or cohort study
[Response Ends]

1a.06. Provide all other grades and definitions from the evidence grading system.

[Response Begins]
Levels of Evidence
Level 1 - Systematic review & meta-analysis of randomized controlled trials; clinical guidelines based on systematic reviews or meta-analyses
Level 2 - One or more randomized controlled trials
Level 3 - Controlled trial (no randomization)
Level 4 - Case-control or cohort study
Level 5 - Systematic review of descriptive & qualitative studies
Level 6 - Single descriptive or qualitative study
Level 7 - Expert opinion
Source: Melnyk, B.M. & Fineout-Overholt, E. (2011). *Evidence-based practice in nursing and healthcare: A guide to best practice.* Philadelphia: Lippincott, Williams & Wilkins.

[Response Ends]

1a.07. Provide the grade assigned to the recommendation, with definition of the grade.

[Response Begins]
Level 4 - case-control or cohort study
[Response Ends]

1a.08. Provide all other grades and definitions from the recommendation grading system.

[Response Begins]
Levels of Evidence
Level 1 - Systematic review & meta-analysis of randomized controlled trials; clinical guidelines based on systematic reviews or meta-analyses
Level 2 - One or more randomized controlled trials

Level 3 - Controlled trial (no randomization)
Level 4 - Case-control or cohort study
Level 5 - Systematic review of descriptive & qualitative studies
Level 6 - Single descriptive or qualitative study
Level 7 - Expert opinion
Source: Melnyk, B.M. & Fineout-Overholt, E. (2011). *Evidence-based practice in nursing and healthcare: A guide to best practice.* Philadelphia: Lippincott, Williams & Wilkins.

[Response Ends]

1a.09. Detail the quantity (how many studies) and quality (the type of studies) of the evidence.

[Response Begins]
54 studies. Descriptive correlational studies.
[Response Ends]

1a.10. Provide the estimates of benefit, and consistency across studies.

[Response Begins]
Quotation from Discussion: "This review shows that a healthy work environment plays a significant role in healthcare delivery and is fundamental in providing high quality
patient care."
An example study result of the benefit is "Patients who were taken care of in hospitals with poor work environments had 16% less chance of surviving in-hospital cardiac arrests than those in hospitals with respectable work environments. (McHugh et al 2016."

[Response Ends]

1a.11. Indicate what, if any, harms were identified in the study.

[Response Begins]
There have been no harms reported related to nurse work environments as measured by this instrument.
[Response Ends]

1a.12. Identify any new studies conducted since the systematic review, and indicate whether the new studies change the conclusions from the systematic review.

[Response Begins]
Since the end date of the systematic review was Dec 2017, there have been 197 additional studies. These studies do not change the conclusions of the systematic review.
[Response Ends]

1a.13. If source of evidence is NOT from a clinical practice guideline, USPSTF, or systematic review, describe the evidence on which you are basing the performance measure.

[Response Begins]
2021 evidence:
Sloane, D. M., et al. (2018). "Effect of Changes in Hospital Nursing Resources on Improvements in Patient Safety and Quality of Care: A Panel Study." **Med Care 56**(12): 1001-1008.
Lake E.T., Sanders J., Duan, R., Riman K., Schoenauer K, & Chen Y. (2019). A meta-analysis of the associations between the nurse work environment in hospitals and four sets of outcomes. Medical Care 57(5):353-361.
Lake ET, Riman KA, Sloane DM. (2020). Improved work environments and staffing lead to less missed nursing care: A panel study. Journal of Nursing Management DOI: 10.1111/jonm.12970 Epub 2020 March 12
2018 evidence:

Wei, H., et al. (2018). "The state of the science of nurse work environments in the United States: A systematic review." International Journal of Nursing Sciences 5(3): 287-300.

[Response Ends]

1a.14. Briefly synthesize the evidence that supports the measure.

[Response Begins]
2021 evidence:
Sloane et al 2018:
After taking into account cross-sectional differences between hospitals, differences among nurses within hospitals, and potential confounding variables, changes within hospitals in nursing resources were associated with significant changes in quality of care and patient safety. Improvements in work environment of 1 SD decrease odds of unfavorable quality care and patient safety by factors ranging from 0.82 to 0.97. Improvements within hospitals in work environments, nurse staffing, and educational composition of nurses coincide with improvements in quality of care and patient safety.
Lake et al 2019:
Consistent, significant associations between the work environment and all outcome classes were identified. Better work environments were associated with lower odds of negative nurse outcomes (average OR of 0.71), poor safety or quality ratings (average OR of 0.65), and negative patient outcomes (average OR of 0.93), but higher odds of patient satisfaction (OR of 1.16). The nurse work environment warrants attention to promote health care quality, safety, and patient and clinician wellbeing.
Lake et al 2020:
Over the 10-year period, most hospitals exhibited improved work environments, better nurse staffing and more missed care. In hospitals with improved work environments or nurse staffing, the prevalence and frequency of missed care decreased significantly. The effect on missed care of changes in the work environment was greater than that of nurse staffing. Changes in the hospital work environment and staffing influence missed care. Implications for Nursing Management: Modifications in the work environment and staffing are strategies to mitigate care compromise. Nurse managers should investigate work settings in order to identify weaknesses.
2018 evidence:
Wei, et al 2018:
This systematic review synthesized evidence from 54 studies published from January 2005 through December 2017. The researchers concluded that: "Healthier work environments lead to more satisfied nurses who will result in better job performance and higher quality of patient care, which will subsequently improve healthcare organizations' financial viability."

[Response Ends]

1a.15. Detail the process used to identify the evidence.

[Response Begins]
A literature search following the PRISMA guideline of five databases: MEDLINE via PubMed, CINAHL, PsycINFO, Nursing and Allied Health, and the Cochrane Library.
[Response Ends]

1a.16. Provide the citation(s) for the evidence.

[Response Begins]
2021 evidence:
Sloane, D. M., et al. (2018). "Effect of Changes in Hospital Nursing Resources on Improvements in Patient Safety and Quality of Care: A Panel Study." **Med Care 56**(12): 1001-1008.
Lake E.T., Sanders J., Duan, R., Riman K., Schoenauer K, & Chen Y. (2019). A meta-analysis of the associations between the nurse work environment in hospitals and four sets of outcomes. Medical Care 57(5):353-361.
Lake ET, Riman KA, Sloane DM. (2020). Improved work environments and staffing lead to less missed nursing care: A panel study. Journal of Nursing Management DOI: 10.1111/jonm.12970 Epub 2020 March 12
2018 evidence:

Wei, H., et al. (2018). "The state of the science of nurse work environments in the United States: A systematic review." International Journal of Nursing Sciences 5(3): 287-300.

[Response Ends]

1b.01. Briefly explain the rationale for this measure.

Explain how the measure will improve the quality of care, and list the benefits or improvements in quality envisioned by use of this measure.

[Response Begins]
The dissemination of the PES-NWI nationally and internationally assures that nurses' practice environments will be measured in consistent fashion across different health systems to develop evidence guiding policy and management decisions. The benefit of using the PES-NWI measure for health care organizations is that organizations provide better quality patient care through improved work environments.
[Response Ends]

1b.02. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis.

Include mean, std dev, min, max, interquartile range, and scores by decile. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities include. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]
October 2021 Maintenance
These 16 studies, published since the previous maintenance cycle of October 2018, demonstrate a continuing gap in performance since the last time this was endorsed. They are listed in chronological order. A reference list is included.

Knupp, A. M., Patterson, E. S., Ford, J. L., Zurmehly, J., & Patrick, T. (2018). Associations among nurse fatigue, individual nurse factors, and aspects of the nursing practice environment. *JONA: The Journal of Nursing Administration*, *48*(12), 642-648.

This is Table 2, from the above publication, which reports descriptive statistics (mean, SD, minimum, maximum) for PES-NWI subscales, demonstrating variation or gap in care in this sample.
TABLE 2 Descriptive Statistics for Three PES-NWI Subscales

*	Mean	SD	Min	Max
Nurse manager ability, leadership, and support	2.84	0.72	1	4
Staffing and resource adequacy	2.64	0.84	1	4
Collegial nurse-physician relations	2.80	0.84	1	4

* Cell intentionally left empty
Abbreviations: Min = minimum value; max = maximum value. Scale is 1 (strongly disagree) to 4 (strongly agree).

Hallowell, S. G., Rogowski, J. A., & Lake, E. T. (2019). How nurse work environments relate to the presence of parents in neonatal intensive care. *Advances in neonatal care: official journal of the National Association of Neonatal Nurses*, *19*(1), 65

The practice environment was highly rated in the NICU sample, where the average PES-NWI composite score was more than 3.0 (mean = 3.06), with a range of 2.42 to 3.97.

Smith, J. G., Plover, C. M., McChesney, M. C., & Lake, E. T. (2019). Isolated, small, and large hospitals have fewer nursing resources than urban hospitals: Implications for rural health policy. Public Health Nursing, 36(4), 469-477.

This is Table 5, from the above publication, which reports descriptive statistics for PES-NWI subscales in hospitals across more urban or rural settings, demonstrating statistically significant variation or gap in care in this sample.

TABLE 5 PES-NWI: rural versus urban (hospital-level) (n = 642 hospitals)

*	**Overall n = 642**	**Urban n = 566**	**Large n = 49**	**Small n = 18**	**Isolated n = 9**	**p-value within and between groups**
Participation in hospital affairs	2.53	2.53	2.45	2.55	**2.36**	0.09
Nursing foundations for quality care	2.92	2.93	2.83	2.89	**2.75**	0.00**
Nurse manager ability, leadership, and support of nurses	2.59	2.59	**2.56**	2.72	2.60	0.29
Staffing and resource adequacy	2.51	**2.50**	**2.50**	2.82	2.67	0.00**
Collegial nurse-physician relations	2.91	2.91	2.85	3.00	**2.84**	0.08
Composite PES-NWI score	2.69	2.69	**2.64**	2.80	**2.64**	0.10

* Cell intentionally left empty

Note: Bolded numbers indicate the lowest value compared across rurality.

Smith, J. G., Plover, C. M., McChesney, M. C., & Lake, E. T. (2019). Rural hospital nursing skill mix and work environment associated with frequency of adverse events. *SAGE open nursing*, *5*, 2377960819848246.

This is Table 3, from the above publication, which reports descriptive statistics for PES-NWI subscales in a sample of 76 rural hospitals, demonstrating variation or gap in care in this sample.

TABLE 3 Rural Hospital Nurse Work Environment (n=76 Hospitals)[†]

Independent Variable	**Mean**	**SD**	**Min**	**Max**
Nurse participation in hospital affairs	2.46	.31	1.65	3.15
Nursing foundations for quality care	2.84	.23	2.27	3.35
Nurse manager ability, leadership, and support of nurses	2.60	.37	1.75	3.50
Staffing and resource adequacy	2.60e	.38	1.80	3.60
Collegial nurse-physician relations	2.89	.26	2.20	3.67
Nurse work environment composite	2.68	.26	2.15	3.35

* Cell intentionally left empty

Note: minimum values refer to hospitals with the lowest PES_NWI subscale and composite scores in sample. Maximum values refer to hospitals with the highest PES-NWI subscale and composite scores in sample.

Campbell, C. M., Prapanjaroensin, A., Anusiewicz, C. V., Baernholdt, M., Jones, T., & Patrician, P. A. (2020). Variables associated with missed nursing care in Alabama: A cross-sectional analysis. *Journal of Nursing Management*, *28*(8), 2174-2184.

This is Table 3, from the above publication, which reports descriptive statistics for PES-NWI subscales in 950 nurses in the state of Alabama, demonstrating variation or gap in care in this sample.

TABLE 3 Descriptive statistics for subscales and composite scores of PES-NWI (N=950)

Factors	**Mean**	**SD**	**Min**	**Max**	*
Collegial Nurse-Physician Relations	3.07	0.69	1	4	*
Nurse Foundations for Quality of Care	3.07	0.60	1	4	*
Nurse Manager Ability, Leadership, and Support of Nurses	2.82	0.85	1	4	*
Nurse Participation in Hospital Affairs	2.71	0.73	1	4	*
Staffing and Resource Adequacy	2.51	0.88	1	4	*

Factors	Mean	SD	Min	Max	*
Composite Score	2.83	0.62	1	4	*

* Cell intentionally left empty

Lake, E. T., French, R., O'Rourke, K., Sanders, J., & Srinivas, S. K. (2020). Linking the work environment to missed nursing care in labour and delivery. *Journal of Nursing Management*, *28*(8), 1901-1908.

This is Table 2, from the above publication, which reports descriptive statistics for PES-NWI subscales in 1,297 maternity nurses, demonstrating variation or gap in care in this sample.

TABLE 2 Descriptive and psychometric statistics for the Practice Environment Scale of the Nursing Work Index[†]

Subscale	Mean	SD	Alpha	ICC
Staffing and resource adequacy	2.74	0.74	0.83	0.68
Nurse manager ability, leadership, and support of nurses	2.64	0.84	0.82	0.73
Nursing foundations for quality of care	3.06	0.52	0.81	0.58
Nurse participation in hospital affairs	2.62	0.65	0.80	0.64
Collegial nurse-physician relations	3.03	0.70	0.86	0.54
Composite	2.84	0.56	0.85	0.65

Abbreviations: ICC, intraclass correlation coefficient; SD, standard deviation; n = 1,297 nurses.

Lake, E. T., Staiger, D. O., Cramer, E., Hatfield, L. A., Smith, J. G., Kalisch, B. J., & Rogowski, J. A. (2020). Association of patient acuity and missed nursing care in US neonatal intensive care units. *Medical care research and review*, *77*(5), 451-460.

This is Table 2, from the above publication, which reports descriptive statistics for PES-NWI composite in 320 neonatal intensive care units, demonstrating variation or gap in care in this sample.

TABLE 2 Descriptive Statistics for the Work Environment.[†]

*	N	M	SD	Min	Max
Nurse work environment composite score	303	2.94	0.25	2.07	3.63

* Cell intentionally left empty

Karakachian, A., Colbert, A., Hupp, D., & Berger, R. (2021). Caring for victims of child maltreatment: Pediatric nurses' moral distress and burnout. *Nursing ethics*, *28*(5), 687-703.

The overall mean for the five subscales were (1) adequate staffing, 2.91; (2) manager leadership, 3.12; (3) foundation of quality of care, 3.17; (4) hospital affairs participation, 3.08; and (5) nurse–physician relationship, 3.99.

Lasater, K. B., McHugh, M., Rosenbaum, P. R., Aiken, L. H., Smith, H., Reiter, J. G., ... & Silber, J. H. (2021). Valuing Hospital investments in nursing: multistate matched-cohort study of surgical patients. *BMJ quality & safety*, *30*(1), 46-55.

This table from the above publication, reports descriptive statistics for the PES-NWI composite in hospitals classified as having better or worse nursing resources, demonstrating variation or gap in care in this sample.

Nurse work environment	Mean (SD)
Better nursing resources	3.01 (0.19)
Worse nursing resources	2.68 (0.22)

Montgomery, A. P., Azuero, A., & Patrician, P. A. (2021). Psychometric properties of Copenhagen Burnout Inventory among nurses. *Research in nursing & health*, *44*(2), 308-318.

This table from the above publication reports descriptive statistics for PES-NWI subscales and composite, demonstrating variation or gap in care in this sample.

PES-NWI	Mean	SD
Nurse Participation in Hospital Affairs	2.73	0.72
Nursing Foundations for Quality of Care	3.08	0.60
Nurse Manager, Ability, Leadership, and Support of Nurses	2.82	0.86
Staffing and Resource Adequacy	2.50	0.87
Collegial Nurse-Physician Relations	3.07	0.70
PES-NWI Composite Score	2.84	0.62

Sano, R., Schiffman, R. F., Shoji, K., & Sawin, K. J. (2018). Negative consequences of providing nursing care in the neonatal intensive care unit. *Nursing Outlook*, *66*(6), 576-585.
Nurse-physician collegiality: Mean (SD): 3.34 (0.57)
Smith, J. G., Rogowski, J. A., & Lake, E. T. (2020). Missed care relates to nurse job enjoyment and intention to leave in neonatal intensive care. *Journal of Nursing Management*, *28*(8), 1940-1947.
Nurse work environment: Mean (SD): 2.95 (0.27)
Baernholdt, M., Jones, T. L., Anusiewicz, C. V., Campbell, C. M., Montgomery, A., & Patrician, P. A. (2022). Development and Testing of the Quality Improvement Self-efficacy Inventory. *Western Journal of Nursing Research*, *44*(2), 159-168.

This is Table 1, from the above publication, which reports descriptive statistics for PES-NWI subscales and composite in 886 nurses, demonstrating variation or gap in care in this sample.

Table 1. Scores on the Practice Environment Scale of the Nursing Work Index and QISEI Instrument Ratings (N=886).

Variable	n	Mean	SD
Practice environment scale of the nursing work index (range: 1-4)	*	*	*
Nurse participation in hospital affairs (9 items)	**885**	**2.72**	**0.72**
Nursing foundations for quality of care (10 items)	**885**	**3.08**	**0.60**
Nurse manager ability, leadership, and support of nurses	**885**	**2.82**	**0.86**
Staffing and resource adequacy (4 items)	**885**	**2.50**	**0.87**
Collegial nurse-physician relations (3 items)	**885**	**3.07**	**0.69**
Composite Score	**885**	**2.84**	*

* Cell intentionally left empty

Brom, H., Carthon, M. B., Sloane, D., McHugh, M., & Aiken, L. (2021). Fewer Readmissions and Shorter Length of Stay Among Adults with Ischemic Stroke When Cared for in Hospitals with Better Nurse Work Environments. *Research in nursing & health*, *44* https://doi.org/10.1002/nur.22121

This is Table 2, from the above publication, which reports descriptive statistics for PES-NWI composite in 543 hospitals, demonstrating variation or gap in care in this sample.
TABLE 2 Hospital characteristics where ischemic stroke patients receive care†

Characteristic	Hospitals (n=543)	*	Poor (n=213, 39.2%)	*	Mixed (n=126, 23.2%)	*	Best (n=204, 37.6%)	*
*	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Work environment	2.7	0.2	2.5	0.1	2.7	0.1	3.0	0.1

* Cell intentionally left empty

Djukic, M., Jun, J., & Fletcher, J. (2021). An Examination of the Factors Associated With Implementation of Evidence-Based Management Practices for Improving Nurse Work Environments. *Worldviews on Evidence-Based Nursing*, *18*(2), 129-137.

This table from the above publication which reports descriptive statistics for PES-NWI subscales and composite, demonstrating variation or gap in care in this sample.

Variable	**n**	**M**	**SD**
Staff nurses participation in hospital affairs	166	3.57	.68
Providing strong foundation for nursing care quality	164	3.94	.64
Nurse manager ability, leadership	170	4.42	.56
Staffing and resource adequacy	167	4.10	.76
Promoting nurse-physician relations	171	4.32	.71

Farag, A., Moon, C., & Xiao, Q. (2021). Work and Personal Characteristics Associated with Sleep Behavior Among Acute Care Nurses. *Journal of Nursing Regulation*, *12*(1), 40-51.
This table from the above publication reports the percent of nurses who agree that organizational attributes are present for PES-NWI subscales, demonstrating variation or gap in care in this sample.

Work environment (n = 1,134). Percent who agree or strongly agree	*	*
Good leadership support	744	65.6
Good nurse-physician relationship	908	80.7
Good staffing and resources adequacy	688	60.7

References:

1. Baernholdt M, Jones TL, Anusiewicz CV, Campbell CM, Montgomery A, Patrician PA. Development and Testing of the Quality Improvement Self-efficacy Inventory. *Western Journal of Nursing Research.* 2022;44(2):159-168.
2. Brom H, Brooks Carthon JM, Sloane D, McHugh M, Aiken L. Better nurse work environments associated with fewer readmissions and shorter length of stay among adults with ischemic stroke: A cross-sectional analysis of United States hospitals. *Res Nurs Health.* 2021;44(3):525-533.
3. Campbell CM, Prapanjaroensin A, Anusiewicz CV, Baernholdt M, Jones T, Patrician PA. Variables associated with missed nursing care in Alabama: A cross-sectional analysis. *Journal of Nursing Management.* 2020;28(8):2174-2184.
4. Djukic M, Jun J, Fletcher J. An Examination of the Factors Associated With Implementation of Evidence-Based Management Practices for Improving Nurse Work Environments. *Worldviews on Evidence-Based Nursing.* 2021;18(2):129-137.
5. Farag A, Moon C, Xiao Q. Work and Personal Characteristics Associated With Sleep Behavior Among Acute Care Nurses. *Journal of Nursing Regulation.* 2021;12(1):40-51.
6. Hallowell SG, Rogowski JA, Lake ET. How nurse work environments relate to the presence of parents in neonatal intensive care. *Adv Neonatal Care.* 2017.
7. Karakachian A, Colbert A, Hupp D, Berger R. Caring for victims of child maltreatment: Pediatric nurses' moral distress and burnout. *Nursing ethics.* 2021;28(5):687-703.
8. Knupp AM, Patterson ES, Ford JL, Zurmehly J, Patrick T. Associations among nurse fatigue, individual nurse factors, and aspects of the nursing practice environment. *JONA: The Journal of Nursing Administration.* 2018;48(12):642-648.
9. Lake ET, Staiger DO, Cramer E, et al. Association of patient acuity and missed nursing care in US neonatal intensive care units. *Med Care Res Rev.* 2020;77(5):451-460.
10. Lake ET, French R, O'Rourke K, Sanders J, Srinivas SK. Linking the work environment to missed nursing care in labour and delivery. *Journal of Nursing Management.* 2020;28(8):1901-1908.
11. Lasater KB, McHugh M, Rosenbaum PR, et al. Valuing hospital investments in nursing: multistate matched-cohort study of surgical patients. *BMJ Qual Saf.* 2021;30(1):46-55.
12. Montgomery AP, Azuero A, Patrician PA. Psychometric properties of Copenhagen Burnout Inventory among nurses. *Res Nurs Health.* 2021;44(2):308-318.
13. Sano R, Schiffman RF, Shoji K, Sawin KJ. Negative consequences of providing nursing care in the neonatal intensive care unit. *Nursing Outlook.* 2018;66(6):576-585.
14. Smith JG, Plover CM, McChesney MC, Lake ET. Isolated, small, and large hospitals have fewer nursing resources than urban hospitals: Implications for rural health policy. *Public Health Nursing.* 2019;36(4):469-477.

15. Smith JG, Rogowski JA, Lake ET. Missed care relates to nurse job enjoyment and intention to leave in neonatal intensive care. *Journal of Nursing Management.* 2019;28:1940-1947.
16. Smith JG, Plover CM, McChesney MC, Lake ET. Rural hospital nursing skill mix and work environment associated with frequency of adverse events. *SAGE open nursing.* 2019;5:2377960819848246.

[Response Ends]

1b.03. If no or limited performance data on the measure as specified is reported above, then provide a summary of data from the literature that indicates opportunity for improvement or overall less than optimal performance on the specific focus of measurement. Include citations.

[Response Begins]
Maintenance of Endorsement (October 2021):
Aiken et al (2018) demonstrated that over a one-decade period in a sample of 535 hospitals, the work environment, as measured by the PES-NWI, improved in 21% of hospitals, remained stable in 71%, and worsened in 8%, signifying variability across hospitals and time and ability for the instrument to detect changes, both positive and negative.
Maintenance of Endorsement (October 2018):
Add three systematic reviews and meta-analysis since 2012.
Since 2012 there have been two new systematic reviews: Swiger et al. (2017) and Lee & Scott (2018). Here are summaries demonstrating opportunity for improvement.
Swiger, P. A., Patrician, P. A., Miltner, R. S. S., Raju, D., Breckenridge-Sproat, S., & Loan, L. A. (2017). The Practice Environment Scale of the Nursing Work Index: an updated review and recommendations for use. International journal of nursing studies, 74, 76-84. Retrieved from https://www.sciencedirect.com/science/article/pii/S0020748917301281
The literature review aimed to provide an updated review and usage recommendations for the Practice Environment Scale of the Nursing Work Index. Researchers included 46 articles from 28 countries between 2010 and 2016 that focused on the relationships between the Practice Environment Scale of the Nursing Work Index and patient, nurse, or organizational outcomes. Most studies indicated significant findings between effects of nurse practice environments on outcomes. The instrument has remained largely unchanged since its development and frequency of usage continues to be high.
This excerpt from Swiger et al. page 79 notes a performance gap in the literature:
"2.6.1. Reported PES-NWI scores
Sixteen articles (35%) reported composite PES-NWI scores, based on the 4-point Likert scale, which ranged from 2.30 to 3.07. The lowest composite score came from a study with a relatively low sample size (n =301) investigating turnover intention of registered nurses in the Eastern Caribbean who worked on medical, surgical, medical-surgical, or obstetric units (Lansiquot et al., 2012). The highest score came from a hospital in Australia that was in the process of seeking Magnet recognition (Walker et al., 2010). In studies where a sample was identified as having been collected from nurses working in Magnet facilities, the reported composite score ranged from 2.92 to 3.00 (Kutney-Lee et al., 2015; Ma and Park, 2015). Collective subscale and composite score ranges from 3 studies reporting scores from Magnet, emerging or aspiring Magnet, and non-Magnet facilities can be found in Table 1; the Staffing and Resource adequacy remains the lowest subscale for all three groups, confirming the finding from the Warshawsky and Havens (2011) review.
This Table 1 from Swiger et al. presents score ranges from three articles demonstrating lower scores in non-Magnet hospitals, middling scores in Emerging Magnet Hospitals, and higher scores in Magnet Hospitals
Table 1 Reported Score Ranges (n = 3 articles).
PES-NWI Measure Reported Mean Score Range (SD)
Subscale
Non-Magnet Scores Emerging/Aspiring Magnet Scores Magnet Hospital Scores

1. Nurse Participation in Hospital Affairs 2.34 - 2.87 2.49 - 3.06 2.76 - 3.01
2. Nursing Foundations for Quality of Care 2.82 - 3.11 2.98 - 3.19 3.09 - 3.20
3. Nurse Manager Ability,Leadership,& Support of Nurses 2.41 - 3.00 2.48 - 3.17 2.72 - 3.07
4. Staffing and Resource Adequacy 2.07 - 2.62 2.31 - 2.88 2.65 - 2.88
5. Collegial Nurse-Physician Relations 2.78 - 2.99 2.85 - 3.06 2.99 - 3.07

Composite 2.51 - 2.92 2.62– 3.07 2.92 - 3.00
Lee, S. E., & Scott, L. D. (2018). Hospital nurses' work environment characteristics and patient safety outcomes: A literature review. Western journal of nursing research, 40(1), 121-145. Retrieved from http://journals.sagepub.com/doi/full/10.1177/0193945916666071
The literature review conducted by Lee and Scott evaluated associations between hospital nurses' work environment characteristics and patient safety outcomes. Researchers searched five databases and reviewed 18 studies published in

English between 1999 and 2016. Most studies did not include a definition for work environment, and patient outcomes were measured using different variables and instruments. The relationship between nurses' work environment characteristics and patient safety outcomes were inconsistent between studies. These 18 studies, some of which overlap with the earlier Warshawsky & Havens (2011) study, demonstrate a performance gap which was linked to patient safety outcomes

Maintenance of Endorsement (October 2012)

Warshawsky, N. E., & Havens, D. S. (2011). Global use of the practice environment scale of the nursing work index. Nursing Research, 60(1), 17. Retrieved from http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3021172/

[Response Ends]

1b.04. Provide disparities data from the measure as specified (current and over time) by population group, e.g., by race/ethnicity, gender, age, insurance status, socioeconomic status, and/or disability.

Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included. Include mean, std dev, min, max, interquartile range, and scores by decile. For measures that show high levels of performance, i.e., "topped out", disparities data may demonstrate an opportunity for improvement/gap in care for certain sub-populations. This information also will be used to address the sub-criterion on improvement (4b) under Usability and Use.

[Response Begins]

Disparities not applicable to this measure.

[Response Ends]

1b.05. If no or limited data on disparities from the measure as specified is reported above, then provide a summary of data from the literature that addresses disparities in care on the specific focus of measurement. Include citations. Not necessary if performance data provided in above.

[Response Begins]

Disparities not applicable to this measure.

[Response Ends]

Criteria 2: 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. Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.

spma.01. Indicate whether there are changes to the specifications since the last updates/submission. If yes, update the specifications in the Measure Specifications section of the Measure Submission Form, and explain your reasoning for the changes below.

[Response Begins]
No
[Response Ends]

spma.02. Briefly describe any important changes to the measure specifications since the last measure update and provide a rationale.

For annual updates, please explain how the change in specifications affects the measure results. If a material change in specification is identified, data from re-testing of the measure with the new specifications is required for early maintenance review.

For example, specifications may have been updated based on suggestions from a previous NQF CDP review.

[Response Begins]
Maintenance of Endorsement (Oct 2021): There have been no changes.
[Response Ends]

sp.01. Provide the measure title.

Measure titles should be concise yet convey who and what is being measured (see What Good Looks Like).

[Response Begins]
Practice Environment Scale - Nursing Work Index (PES-NWI) (composite and five subscales) (previously NQF#0206 - Undergoing Maintenance)
[Response Ends]

sp.02. Provide a brief description of the measure.

Including type of score, measure focus, target population, timeframe, (e.g., Percentage of adult patients aged 18-75 years receiving one or more HbA1c tests per year).

[Response Begins]
Practice Environment Scale-Nursing Work Index (PES-NWI) is a survey-based measure of the nursing practice environment completed by staff registered nurses; includes mean scores on index subscales and a composite mean of all subscale scores.
[Response Ends]

sp.04. Check all the clinical condition/topic areas that apply to your measure, below.

Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.
Please do not select:

Surgery: General

[Response Begins]
Other (specify)
[Other (specify) Please Explain]

All acute care settings

[Response Ends]

sp.05. Check all the non-condition specific measure domain areas that apply to your measure, below.

[Response Begins]
Safety
[Response Ends]

sp.06. Select one or more target population categories.

Select only those target populations which can be stratified in the reporting of the measure's result.
Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.
Please do not select:

Populations at Risk: Populations at Risk

[Response Begins]
Adults (Age >= 18)
Children (Age < 18)
Populations at Risk: Veterans
[Response Ends]

sp.07. Select the levels of analysis that apply to your measure.

Check ONLY the levels of analysis for which the measure is SPECIFIED and TESTED.
Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.
Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]
Facility
[Response Ends]

sp.08. Indicate the care settings that apply to your measure.

Check ONLY the settings for which the measure is SPECIFIED and TESTED.
[Response Begins]
Inpatient/Hospital
[Response Ends]

sp.09. Provide a URL link to a web page specific for this measure that contains current detailed specifications including code lists, risk model details, and supplemental materials.

Do not enter a URL linking to a home page or to general information. If no URL is available, indicate "none available".

[Response Begins]
none available
[Response Ends]

sp.11. Attach the data dictionary, code table, or value sets (and risk model codes and coefficients when applicable). Excel formats (.xlsx or .csv) are preferred.

Attach an excel or csv file; if this poses an issue, contact staff. Provide descriptors for any codes. Use one file with multiple worksheets, if needed.

[Response Begins]

No data dictionary/code table – all information provided in the submission form

[Response Ends]

sp.12. State the numerator.

Brief, narrative description of the measure focus or what is being measured about the target population, i.e., cases from the target population with the target process, condition, event, or outcome).
DO NOT include the rationale for the measure.

[Response Begins]

Continuous Variable Statement: For surveys completed by Registered Nurses (RN):
12a) Mean score on a composite of all subscale scores
12b) Mean score on Nurse Participation in Hospital Affairs (survey item numbers 5, 6, 11, 15, 17, 21, 23, 27, 28)
12c) Mean score on Nursing Foundations for Quality of Care (survey item numbers 4, 14, 18, 19, 22, 25, 26, 29, 30, 31)
12d) Mean score on Nurse Manager Ability, Leadership, and Support of Nurses (survey item numbers 3, 7, 10, 13, 20)
12e) Mean score on Staffing and Resource Adequacy (survey item numbers 1, 8, 9, 12)
12f) Mean score on Collegial Nurse-Physician Relations (survey item numbers 2, 16, 24)
12g) Three category variable indicating favorable, mixed, or unfavorable practice environments: favorable = four or more subscale means exceed 2.5; mixed = two or three subscale means exceed 2.5; unfavorable = zero or one subscales exceed 2.5.

[Response Ends]

sp.13. Provide details needed to calculate the numerator.

All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, time period for data collection, specific data collection items/responses, code/value sets.
Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]

Included Populations:

- Registered Nurses with direct patient care responsibilities for 50% or greater of their shift
- All hospital units
- Full time, part time, and flex / pool RNs employed by the hospital

Excluded Populations

- New hires of less than 3 months
- Agency, traveler or contract nurses
- Nurses in management or supervisory roles with direct patient care responsibilities less than 50% of their shift, whose primary responsibility is administrative in nature

Data Elements by Subscale (with survey question/item number)

Nurse Participation in Hospital Affairs
PES-NWI Career Development (5)
PES-NWI Participation in Policy Decisions (6)
PES-NWI Chief Nursing Officer Visibility (11)
PES-NWI Chief Nursing Officer Authority (15)
PES-NWI Advancement Opportunities (17)
PES-NWI Administration Listens and Responds (21)
PES-NWI Staff Nurses Hospital Governance (23)

PES-NWI Nursing Committees (27)
PES-NWI Nursing Administrators Consult (28)

Nursing Foundations for Quality of Care
PES-NWI Continuing Education (4)
PES-NWI High Nursing Care Standards (14)
PES-NWI Philosophy of Nursing (18)
PES-NWI Nurses Are Competent (19)
PES-NWI Quality Assurance Program (22)
PES-NWI Preceptor Program (25)
PES-NWI Nursing Care Model (26)
PES-NWI Patient Care Plans (29)
PES-NWI Continuity of Patient Assignments (30)
PES-NWI Nursing Diagnosis (31)

Nurse Manager Ability, Leadership, and Support of Nurses
PES-NWI Supportive Supervisory Staff (3)
PES-NWI Supervisors Learning Experiences (7)
PES-NWI Nurse Manager and Leader (10)
PES-NWI Recognition (13)
PES-NWI Nurse Manager Backs up Staff (20)

Staffing and Resource Adequacy
PES-NWI Adequate Support Services (1)
PES-NWI Time to Discuss Patient Problems (8)
PES-NWI Enough Nurses for Quality Care (9)
PES-NWI Enough Staffing (12)

Collegial Nurse-Physician Relations
PES-NWI Nurse and Physician Relationships (2)
PES-NWI Nurse and Physician Teamwork (16)
PES-NWI Collaboration (24)

Composite Score
Mean of subscale scores

Three Category Variable
Favorable = four or more subscale means exceed 2.5
Mixed = two or three subscale means exceed 2.5
Unfavorable = zero or one subscales exceed 2.5
[Response Ends]

sp.14. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]
Staff RNs
[Response Ends]

sp.15. Provide details needed to calculate the denominator.

All information required to identify and calculate the target population/denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets.
Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]
The target population is staff registered nurses. The denominator is calculated as the number of eligible staff RNs in the facility. The time period is typically three or four weeks for an eligible nurse to complete the survey. Specific data collection items are answers to each of the 31 survey items.
To calculate a subscale score, the numerator is the sum of responses (values of 1 to 4 in Likert categories) for all items in a subscale. The denominator is the number of items in the subscale. The quotient is the subscale score, which is a simple average. Higher values indicate greater agreement that desirable organizational attributes are present in the current job, which yields higher scores for the instrument. The composite is calculated as the average value of all the subscales.
[Response Ends]

sp.16. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]
Not applicable
[Response Ends]

sp.17. Provide details needed to calculate the denominator exclusions.

All information required to identify and calculate exclusions from the denominator such as definitions, time period for data collection, specific data collection items/responses, code/value sets – Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format at sp.11.

[Response Begins]
Not applicable
[Response Ends]

sp.18. Provide all information required to stratify the measure results, if necessary.

Include the stratification variables, definitions, specific data collection items/responses, code/value sets, and the risk-model covariates and coefficients for the clinically-adjusted version of the measure when appropriate. Note: lists of individual codes with descriptors that exceed 1 page should be provided in an Excel or csv file in required format in the Data Dictionary field.

[Response Begins]
12a) Mean score on a composite of all subscale scores
12b) Mean score on Nurse Participation in Hospital Affairs (survey item numbers 5, 6, 11, 15, 17, 21, 23, 27, 28)
12c) Mean score on Nursing Foundations for Quality of Care (survey item numbers 4, 14, 18, 19, 22, 25, 26, 29, 30, 31)
12d) Mean score on Nurse Manager Ability, Leadership, and Support of Nurses (survey item numbers 3, 7, 10, 13, 20)
12e) Mean score on Staffing and Resource Adequacy (survey item numbers 1, 8, 9, 12)
12f) Mean score on Collegial Nurse-Physician Relations (survey item numbers 2, 16, 24)
12g) Three category variable indicating favorable, mixed, or unfavorable practice environments: favorable = four or more subscale means exceed 2.5; mixed = two or three subscale means exceed 2.5; unfavorable = zero or one subscales exceed 2.5.
[Response Ends]

sp.19. Select the risk adjustment type.

Select type. Provide specifications for risk stratification and/or risk models in the Scientific Acceptability section.
[Response Begins]
No risk adjustment or risk stratification
[Response Ends]

sp.20. Select the most relevant type of score.

Attachment: If available, please provide a sample report.
[Response Begins]

Continuous variable, e.g. average
[Response Ends]

sp.21. Select the appropriate interpretation of the measure score.

Classifies interpretation of score according to whether better quality or resource use is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score
[Response Begins]
Better quality = Higher score
[Response Ends]

sp.22. Diagram or describe the calculation of the measure score as an ordered sequence of steps.

Identify the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period of data, aggregating data; risk adjustment; etc.

[Response Begins]

1. Start processing.
2. Check Survey Date
    a. If the Survey Date is missing or invalid the case will proceed to a Measure Category Assignment of X and will be rejected. Stop processing.
    b. If Survey Date is valid, continue and proceed to initialization.
3. Initialization. Initialize NurseParticipationScore to 0; NursingFoundationScore to 0; NurseMgrAbilityScore to 0; StaffingScore to 0; RelationsScore to 0; TotalScore to 0; ExceedCounter to 0. Continue and proceed to PES-NWI Career Development.
4. Check PES-NWI Career Development
    a. If the PES-NWI Career Development is missing or zero, the case will proceed to PES-NWI Participation in Policy Decisions.
    b. If the PES-NWI Career Development equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Career Development to the NurseParticipationScore and proceed to PES-NWI Participation in Policy Decisions.
5. Check PES-NWI Participation in Policy Decisions
    a. If the PES-NWI-Participation in Policy Decisions is missing or zero, the case will proceed to PES-NWI Chief Nursing Officer Visibility.
    b. If the PES-NWI Participation in Policy Decisions equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Participation in Policy Decisions to the NurseParticipationScore and proceed to PES-NWI Chief Nursing Officer Visibility.
6. Check PES-NWI Chief Nursing Officer Visibility
    a. If the PES-NWI- Chief Nursing Officer Visibility is missing or zero, the case will proceed to PES-NWI Chief Nursing Officer Authority.
    b. If the PES-NWI Chief Nursing Officer Visibility equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Chief Nursing Officer Visibility to the NurseParticipationScore and proceed to PES-NWI Chief Nursing Officer Authority.
7. Check PES-NWI Chief Nursing Officer Authority
    a. If the PES-NWI- Chief Nursing Officer Authority is missing or zero, the case will proceed to PES-NWI Advancement Opportunities.
    b. If the PES-NWI Chief Nursing Officer Authority equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Chief Nursing Officer Authority to the NurseParticipationScore and proceed to PES-NWI Advancement Opportunities.
8. Check PES-NWI Advancement Opportunities
    a. If the PES-NWI- Advancement Opportunities is missing or zero, the case will proceed to PES-NWI Administration Listens and Responds.
    b. If the PES-NWI Advancement Opportunities equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Advancement Opportunities to the NurseParticipationScore and proceed to PES-NWI Administration Listens and Responds.
9. Check PES-NWI Administration Listens and Responds
    a. If the PES-NWI Administration Listens and Responds is missing or zero, the case will proceed to PES-NWI Staff Nurses Hospital Governance.

    b. If the PES-NWI Administration Listens and Responds equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Administration Listens and Responds to the NurseParticipationScore and proceed to PES-NWI Staff Nurses Hospital Governance.
10. Check PES-NWI Staff Nurses Hospital Governance
    a. If the PES-NWI- Staff Nurses Hospital Governance is missing or zero, the case will proceed to PES-NWI Nursing Committees.
    b. If the PES-NWI Staff Nurses Hospital Governance equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Staff Nurses Hospital Governance to the NurseParticipationScore and proceed to PES-NWI Nursing Committees.
11. Check PES-NWI Nursing Committees
    a. If the PES-NWI Nursing Committees is missing or zero, the case will proceed to PES-NWI Nursing Administrators Consult.
    b. If the PES-NWI Nursing Committees equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Nursing Committees to the NurseParticipationScore and proceed to PES-NWI Nursing Administrators Consult.
12. Check PES-NWI Nursing Administrators Consult
    a. If the PES-NWI Nursing Administrators Consult is missing or zero, the case will proceed to calculate mean score on Nurse-Participation in Hospital Affairs.
    b. If the PES-NWI Nursing Administrators Consult equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Nursing Administrators Consult to the NurseParticipationScore and proceed to calculate mean score on Nurse-Participation in Hospital Affairs.
13. Calculate Mean Score on Nurse-Participation in Hospital Affairs. Mean Score of Nurse-Participation in Hospital Affairs equals mean of NurseParticipationScore. Assign the calculated mean score to NSC-12b. Continue and proceed to PES-NWI Continuing Education.
14. Check PES-NWI Continuing Education
    a. If the PES-NWI Continuing Education is missing or zero, the case will proceed to PES-NWI High Nursing Care Standards.
    b. If the PES-NWI Continuing Education equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Continuing Education to the NurseFoundationScore and proceed to PES-NWI High Nursing Care Standards.
15. Check PES-NWI High Nursing Care Standards
    a. If the PES-NWI High Nursing Care Standards is missing or zero, the case will proceed to PES-NWI Philosophy of Nursing.
    b. If the PES-NWI High Nursing Care Standards equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI High Nursing Care Standards to the NurseFoundationScore and proceed to PES-NWI Philosophy of Nursing.
16. Check PES-NWI Philosophy of Nursing
    a. If the PES-NWI Philosophy of Nursing is missing or zero, the case will proceed to PES-NWI Nurses Are Competent.
    b. If the PES-NWI Philosophy of Nursing equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Philosophy of Nursing to the NurseFoundationScore and proceed to PES-NWI Nurses Are Competent.
17. Check PES-NWI Nurses Are Competent
    a. If the PES-NWI Nurses Are Competent is missing or zero, the case will proceed to PES-NWI Quality Assurance Program.
    b. If the PES-NWI Nurses Are Competent equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Nurses Are Competent to the NurseFoundationScore and proceed to PES-NWI Quality Assurance Program.
18. Check PES-NWI Quality Assurance Program
    a. If the PES-NWI Quality Assurance Program is missing or zero, the case will proceed to PES-NWI Preceptor Program.
    b. If the PES-NWI Quality Assurance Program equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Quality Assurance Program to the NurseFoundationScore and proceed to PES-NWI Preceptor Program.
19. Check PES-NWI Preceptor Program
    a. If the PES-NWI Preceptor Program is missing or zero, the case will proceed to PES-NWI Nursing Care Model.

b. If the PES-NWI Preceptor Program equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Preceptor Program to the NurseFoundationScore and proceed to PES-NWI Nursing Care Model.

20. Check PES-NWI Nursing Care Model
    a. If the PES-NWI Nursing Care Model is missing or zero, the case will proceed to PES-NWI Patient Care Plans.
    b. If the PES-NWI Nursing Care Model equals 1, 2, 3, or 4, add the allowable value scored for Nursing Care Model to the NurseFoundationScore and proceed to PES-NWI Patient Care Plans.
21. Check PES-NWI Patient Care Plans
    a. If the PES-NWI Patient Care Plans is missing or zero, the case will proceed to PES-NWI Continuity of Patient Assignments.
    b. If the PES-NWI Patient Care Plans equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Patient Care Plans to the NurseFoundationScore and proceed to PES-NWI Continuity of Patient Assignments
22. Check PES-NWI Continuity of Patient Assignments
    a. If the PES-NWI Continuity of Patient Assignments is missing or zero, the case will proceed to PES-NWI Nursing Diagnosis.
    b. If the PES-NWI Continuity of Patient Assignments equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Continuity of Patient Assignments to the NurseFoundationScore and proceed to PES-NWI Nursing Diagnosis.
23. Check PES-NWI Nursing Diagnosis
    a. If the PES-NWI Nursing Diagnosis is missing or zero, the case will proceed to calculate mean score on Nursing Foundations for Quality of Care.
    b. If the PES-NWI Nursing Diagnosis equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Nursing Diagnosis to theNurseFoundationScore and proceed to calculate mean score on Nursing Foundations for Quality of Care.
24. Calculate Mean Score on Nursing Foundations for Quality of Care. Mean Score of Nursing Foundations for Quality of Care equals mean of NurseFoundationScore. Assign the calculated mean score to NSC-12c. Continue and proceed to PES-NWI Supportive Supervisory Staff.
25. Check PES-NWI Supportive Supervisory Staff
    a. If the PES-NWI Supportive Supervisory Staff is missing or zero, the case will proceed to PES-NWI Supervisors Learning Experience.
    b. If the PES-NWI Supportive Supervisory Staff equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Supportive Supervisory Staff to the NurseMgrAbilityScore and proceed to PES-NWI Supervisors Learning Experience.
26. Check PES-NWI Supervisors Learning Experience
    a. If the PES-NWI Supervisors Learning Experience is missing or zero, the case will proceed to PES-NWI Nurse Manager and Leader.
    b. If the PES-NWI Supervisors Learning Experience equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Supervisors Learning Experience to the NurseMgrAbilityScore and proceed to PES-NWI Nurse Manager and Leader.
27. Check PES-NWI Nurse Manager and Leader
    a. If the PES-NWI Nurse Manager and Leader is missing or zero, the case will proceed to PES-NWI Recognition.
    b. If the PES-NWI Nurse Manager and Leader equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Nurse Manager and Leader to the NurseMgrAbilityScore and proceed to PES-NWI Recognition.
28. Check PES-NWI Recognition
    a. If the PES-NWI Recognition is missing or zero, the case will proceed to PES-NWI Nurse Manager Backs up Staff
    b. If the PES-NWI Recognition equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Recognition to the NurseMgrAbilityScore and proceed to PES-NWI Nurse Manager Backs up Staff.
29. Check PES-NWI Nurse Manager Backs up Staff
    a. If the PES-NWI Nurse Manager Backs up Staff is missing or zero, the case will proceed to calculate mean score on Nurse Manager Ability, Leadership, and Support of Nurses.
    b. If the PES-NWI Nurse Manager Backs up Staff equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Nurse Manager Backs up Staff to the NurseMgrAbilityScore and proceed to calculate mean score on Nurse Manager Ability, Leadership, and Support of Nurses.
    Calculate Mean Score on Nurse Manager Ability, Leadership, and Support of Nurses. Mean Score of

Nurse Manager Ability, Leadership, and Support of Nurses equals mean of NurseMgrAbilityScore. Assign the calculated mean score to NSC-12d. Continue and proceed to PES-NWI Adequate Support Services.

30. Check PES-NWI Adequate Support Services
    a. If the PES-NWI Adequate Support Services is missing or zero, the case will proceed to PES-NWI Time to Discuss Patient Problems.
    b. If the PES-NWI Adequate Support Services equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Adequate Support Services to the StaffingScore and proceed to PES-NWI Time to Discuss Patient Problems.
31. Check PES-NWI Time to Discuss Patient Problems
    a. If the PES-NWI Time to Discuss Patient Problems is missing or zero, the case will proceed to PES-NWI Enough Nurses for Quality Care.
    b. If the PES-NWI Time to Discuss Patient Problems equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Time to Discuss Patient Problems to the StaffingScore and proceed to PES-NWI Enough Nurses for Quality Care.
32. Check PES-NWI Enough Nurses for Quality Care
    a. If the PES-NWI Enough Nurses for Quality Care is missing or zero, the case will proceed to PES-NWI Enough Staffing.
    b. If the PES-NWI Enough Nurses for Quality Care equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Enough Nurses for Quality Care to the StaffingScore and proceed to PES-NWI Enough Staffing.
33. Check PES-NWI Enough Staffing
    a. If the PES-NWI Enough Staffing is missing or zero, the case will proceed to calculate mean score on Staffing and Resource Adequacy.
    b. If the PES-NWI Enough Staffing equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Enough Staffing to the StaffingScore and proceed to calculate mean score on Staffing and Resource Adequacy.
34. Calculate Mean Score on Staffing and Resource Adequacy. Mean Score of Staffing and Resource Adequacy equals mean of StaffingScore. Assign the calculated mean score to NSC-12e. Continue and proceed to PES-NWI Nurse and Physician Relationships.
35. Check PES-NWI Nurse and Physician Relationships
    a. If the PES-NWI Nurse and Physician Relationships is missing or zero, the case will proceed to PES-NWI Nurse and Physician Teamwork.
    b. If the PES-NWI Nurse and Physician Relationships equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Nurse and Physician Relationships to the RelationsScore and proceed to PES-NWI Nurse and Physician Teamwork.
36. Check PES-NWI Nurse and Physician Teamwork
    a. If the PES-NWI Nurse and Physician Teamwork is missing or zero, the case will proceed to PES-NWI Collaboration.
    b. If the PES-NWI Nurse and Physician Teamwork equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Nurse and Physician Teamwork to the RelationsScore and proceed to PES-NWI Collaboration.
37. Check PES-NWI Collaboration
    a. If the PES-NWI Collaboration is missing or zero, the case will proceed to calculate mean score on Collegial Nurse-Physician Relations.
    b. If the PES-NWI Collaboration equals 1, 2, 3, or 4, add the allowable value scored for PES-NWI Collaboration to the RelationsScore and proceed to calculate mean score on Collegial Nurse-Physician Relations.
38. Calculate Mean Score on Collegial Nurse-Physician Relations. Mean Score of Collegial Nurse-Physician Relations equals mean of RelationsScore. Assign the calculated mean score to NSC-12f. Continue and proceed to calculate the Total Score on composite of all subscale scores.
39. Calculate Total Score on a composite of all subscale scores. Total Score of a composite of all subscale scores equals the sum of NurseParticipationScore, NursingFoundationScore, NurseMgrAbilityScore, StaffingScore, and RelationsScore. Continue and proceed to calculate Mean Score on a composite of all subscale scores.
40. Calculate Mean Score on a composite of all subscale scores. Mean Score of a composite of all subscale scores equals the mean of Total Score on a composite of all subscale scores. Assign the calculated mean score to NSC-12a. Continue and proceed to Mean Score on NurseParticipationScore.
41. Check Mean Score on NurseParticipationScore
    a. If the score of Mean Score on NurseParticipationScore is less than or equal to 2.5, the case will proceed to Mean Score on NursingFoundationScore.

b. If the score of Mean Score on NurseParticipationScore is greater than 2.5, add 1 to ExceedCounter and proceed to Mean Score on NursingFoundationScore.

42. Check Mean Score on NursingFoundationScore
    a. If the score of Mean Score on NursingFoundationScore is less than or equal to 2.5, the case will proceed to Mean Score on NurseMgrAbilityScore.
    b. If the score of Mean Score on NursingFoundationScore is greater than 2.5, add 1 to ExceedCounter and proceed to Mean Score on NurseMgrAbilityScore.
43. Check Mean Score on NurseMgrAbilityScore
    a. If the score of Mean Score on NurseMgrAbilityScore is less than or equal to 2.5, the case will proceed to Mean Score on StaffingScore.
    b. If the score of Mean Score on NurseMgrAbilityScore is greater than 2.5, add 1 to ExceedCounter and proceed to Mean Score on StaffingScore.
44. Check Mean Score on StaffingScore
    a. If the score of Mean Score on StaffingScore is less than or equal to 2.5, the case will proceed to Mean Score on RelationsScore.
    b. If the score of Mean Score on StaffingScore is greater than 2.5, add 1 to ExceedCounter and proceed to Mean Score on RelationsScore.
45. Check Mean Score on RelationsScore
    a. If the score of Mean Score on RelationsScore is less than or equal to 2.5, the case will proceed to ExceedCounter.
    b. If the score of Mean Score on RelationsScore is greater than 2.5, add 1 to ExceedCounter and proceed to ExceedCounter.
46. Check ExceedCounter
    a. If ExceedCounter is greater than or equal to 4, the case will proceed to a Measure Category Assignment of “Favorable”. Stop processing.
    b. If ExceedCounter is greater than or equal to 2 and less than 4, the case will proceed to a Measure Category Assignment of “Mixed”. Stop processing.
    c. If ExceedCounter is greater than or equal to 0 and less than 2, the case will proceed to a Measure Category Assignment of “Unfavorable”. Stop processing.

[Response Ends]

sp.23. Attach a copy of the instrument (e.g. survey, tool, questionnaire, scale) used as a data source for your measure, if available.

[Response Begins]
Copy of instrument is attached.
[Response Ends]

sp.24. Indicate the responder for your instrument.

[Response Begins]
Clinician
[Response Ends]

sp.25. If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

[Response Begins]
For public reporting, the specific sampling approach is a random sample of 50 direct care staff registered nurses. With an anticipated response rate of 60%, the publicly reported measure would be based on 30 or more responses. The minimum of 30 is based on The Joint Commission´s established minimum for comparative results to be calculated to represent the hospital. Satisfactory estimates of PES hospital scores have been obtained with fewer than 30 responses (Lake & Friese, 2006). Nevertheless, a larger sample improves the precision of the results. While a random sample may be used at the hospital-level, it is recommended that hospitals survey all eligible nurses to allow all nurses the opportunity to complete the practice environment survey instrument.
[Response Ends]

sp.26. Identify whether and how proxy responses are allowed.

[Response Begins]
they are not allowed
[Response Ends]

sp.27. Survey/Patient-reported data.

Provide instructions for data collection and guidance on minimum response rate. Specify calculation of response rates to be reported with performance measure results.

[Response Begins]
According to Lake and Friese (2006) the minimum number of completed surveys per hospital for satisfactory estimates is 15, therefore considering a typical response rate of 60%, a random sample of at least 25 nurses needs to be surveyed annually. For purposes of public reporting the measure a minimum of 30 completed surveys is desired, therefore hospitals that choose to sample should sample a minimum of 50 nurses annually. While a random sample may be used at the hospital-level, it is recommended that hospitals survey all eligible nurses to allow all nurses the opportunity to complete the practice environment survey instrument.
[Response Ends]

sp.28. Select only the data sources for which the measure is specified.

[Response Begins]
Instrument-Based Data
[Response Ends]

sp.29. Identify the specific data source or data collection instrument.

For example, provide the name of the database, clinical registry, collection instrument, etc., and describe how data are collected.

[Response Begins]
Practice Environment Scale-Nursing Work Index (PES-NWI) Survey
[Response Ends]

sp.30. Provide the data collection instrument.

[Response Begins]
Available in attached appendix in Question 1 of the Additional Section
[Response Ends]

2ma.01. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Reliability - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:
Current Submission:
Updated testing information here.
Previous Submission:
Testing from the previous submission here.

[Response Begins]
Yes
[Response Ends]

2ma.02. Indicate whether additional empirical reliability testing at the accountable entity level has been conducted. If yes, please provide results in the following section, Scientific Acceptability: Validity - Testing. Include information on all testing conducted (prior testing as well as any new testing).

Please separate added or updated information from the most recent measure evaluation within each question response in the Scientific Acceptability sections. For example:
Current Submission:
Updated testing information here.
Previous Submission:
Testing from the previous submission here.

[Response Begins]
Yes
[Response Ends]

2ma.03. For outcome, patient-reported outcome, resource use, cost, and some process measures, risk adjustment/stratification may be conducted. Did you perform a risk adjustment or stratification analysis?

[Response Begins]
Yes
[Response Ends]

2ma.04. For maintenance measures in which risk adjustment/stratification has been performed, indicate whether additional risk adjustment testing has been conducted since the most recent maintenance evaluation. This may include updates to the risk adjustment analysis with additional clinical, demographic, and social risk factors.

Please update the Scientific Acceptability: Validity - Other Threats to Validity section.

Note: This section must be updated even if social risk factors are not included in the risk adjustment strategy.

[Response Begins]
No additional risk adjustment analysis included
[Response Ends]

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 fields in the Scientific Acceptability sections of the Measure Submission Form.

- Measures must be tested for all the data sources and levels of analyses that are specified. If there is more than one set of data specifications or more than one level of analysis, contact NQF staff about how to present all the testing information in one form.
- All required sections must be completed.
- For composites with outcome and resource use measures, Questions 2b.23-2b.37 (Risk Adjustment) also must be completed.
- If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), Questions 2b.11-2b.13 also must be completed.
- An appendix for supplemental materials may be submitted (see Question 1 in the Additional section), but there is no guarantee it will be reviewed.
- Contact NQF staff with any questions. Check for resources at the Submitting Standards webpage.
- For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for the 2021 Measure Evaluation Criteria and Guidance.

Note: The information provided in this form is intended to aid the Standing Committee and other stakeholders in understanding to what degree the testing results for this measure meet NQF's evaluation criteria for testing.

2a. Reliability testing demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For instrument-based measures (including PRO-PMs) and composite performance measures, reliability should be demonstrated for the computed performance score.
2b1. Validity testing demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For instrument based measures (including PRO-PMs) and composite performance measures, validity should be demonstrated for the computed performance score.
2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure;
AND
If patient preference (e.g., informed decision-making) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately).
2b3. For outcome measures and other measures when indicated (e.g., resource use):

- an evidence-based risk-adjustment strategy (e.g., risk models, risk stratification) is specified; is based on patient factors (including clinical and social risk factors) that influence the measured outcome and are present at start of care; 14,15 and has demonstrated adequate discrimination and calibration

OR

- rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for identification of statistically significant and practically/clinically meaningful 16 differences in performance;
OR
there is evidence of overall less-than-optimal performance.
2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.
2b6. Analyses identify the extent and distribution of missing data (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders) and how the specified handling of missing data minimizes bias.
2c. For composite performance measures, empirical analyses support the composite construction approach and demonstrate that:
2c1. the component measures fit the quality construct and add value to the overall composite while achieving the related objective of parsimony to the extent possible; and
2c2. the aggregation and weighting rules are consistent with the quality construct and rationale while achieving the related objective of simplicity to the extent possible.
(if not conducted or results not adequate, justification must be submitted and accepted)

Definitions
Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).
Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.
Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.
Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.
Risk factors that influence outcomes should not be specified as exclusions.

With large enough sample sizes, small differences that are statistically significant may or may not be practically or clinically meaningful. The substantive question may be, for example, whether a statistically significant difference of one percentage point in the percentage of patients who received smoking cessation counseling (e.g., 74 percent v. 75 percent) is clinically meaningful; or whether a statistically significant difference of $25 in cost for an episode of care (e.g., $5,000 v.$5,025) is practically meaningful. Measures with overall less-than-optimal performance may not demonstrate much variability across providers.

Please separate added or updated information from the most recent measure evaluation within each question response in the Importance to Scientific Acceptability sections. For example:
2021 Submission:
Updated testing information here.
2018 Submission:
Testing from the previous submission here.

2a.01. Select only the data sources for which the measure is tested.

[Response Begins]
Instrument-Based Data
[Response Ends]

2a.02. If an existing dataset was used, identify the specific dataset.

The dataset used for testing must be consistent with the measure specifications for target population and healthcare entities being measured; e.g., Medicare Part A claims, Medicaid claims, other commercial insurance, nursing home MDS, home health OASIS, clinical registry).

[Response Begins]
The reliability testing was conducted in empirical studies conducted by researchers.
[Response Ends]

2a.03. Provide the dates of the data used in testing.

Use the following format: "MM-DD-YYYY - MM-DD-YYYY"

[Response Begins]
The studies in the meta-analysis described below were published from 2002 to 2017.
[Response Ends]

2a.04. Select the levels of analysis for which the measure is tested.

Testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan.
Please refrain from selecting the following answer option(s). We are in the process of phasing out these answer options and request that you instead select one of the other answer options as they apply to your measure.
Please do not select:

- *Clinician: Clinician*
- *Population: Population*

[Response Begins]
Facility
[Response Ends]

2a.05. List the measured entities included in the testing and analysis (by level of analysis and data source).

Identify the number and descriptive characteristics of measured entities included in the analysis (e.g., size, location, type); if a sample was used, describe how entities were selected for inclusion in the sample.

[Response Begins]
2021 submission
The most comprehensive source of instrument reliability data is from a 2019 publication, in which the authors conducted a meta-analysis to examine the reliability generalization of the PES-NWI. They conducted a meta-analysis of 51 studies representing 80,563 subjects. They estimated a mean reliability for U.S. samples of .946.
Reference: Zangaro, G. A. and K. Jones (2019). "Practice Environment Scale of the Nursing Work Index: A reliability generalization meta-analysis." Western Journal of Nursing Research 41(11): 1658-1684.
2018 submission
Twenty studies have evaluated the psychometric performance of the PES-NWI since its development.

[Response Ends]

2a.06. Identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis), separated by level of analysis and data source; if a sample was used, describe how patients were selected for inclusion in the sample.

If there is a minimum case count used for testing, that minimum must be reflected in the specifications.

[Response Begins]
2021 submission
In the Zangaro, et al. (2019, p. 1664) publication noted in 2a.05, the number and descriptive characteristics of the nurses were as follows:
"The average age across studies, based on 35 studies (n = 63,664) that reported the mean age of the sample was 39 years. Gender was reported in 41 studies and most of the participants were females (>90%). The average number of years of experience was 14.35 years, based on reports in 22 studies. Education was reported differently across studies; however, in 21 reports, >50% of the sample had a bachelor's degree or higher."
In addition, the 51 studies comprised 49 samples from hospitals, and one each from a long-term care facility and a hemodialysis facility.

[Response Ends]

2a.07. If there are differences in the data or sample used for different aspects of testing (e.g., reliability, validity, exclusions, risk adjustment), identify how the data or sample are different for each aspect of testing.

[Response Begins]
The differences in the samples used for different aspects of testing are that, in general, most empirical studies report reliability information. Fewer report validity information. However, those that report validity information are a subset of all that report reliability information.
[Response Ends]

2a.08. List the social risk factors that were available and analyzed.

For example, patient-reported data (e.g., income, education, language), proxy variables when social risk data are not collected from each patient (e.g. census tract), or patient community characteristics (e.g. percent vacant housing, crime rate) which do not have to be a proxy for patient-level data.

[Response Begins]
Social risk factors are not applicable to this instrument, which surveys nurses about their work conditions.
[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.07 check patient or encounter-level data; in 2a.08 enter "see validity testing section of data elements"; and enter "N/A" for 2a.09 and 2a.10.

2a.09. Select the level of reliability testing conducted.

Choose one or both levels.
[Response Begins]
Patient or Encounter-Level (e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements)
[Response Ends]

2a.10. For each level of reliability testing checked above, describe the method of reliability testing and what it tests.

Describe the steps—do not just name a method; what type of error does it test; what statistical analysis was used.

[Response Begins]
The method of reliability testing is Cronbach's alpha, to examine the internal consistency of the 31 items in the instrument or the items in each of the five subscales. This method tests if the items are internally consistent. It is at the nurse-level.
[Response Ends]

2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?

For example, provide the percent agreement and kappa for the critical data elements, or distribution of reliability statistics from a signal-to-noise analysis. For score-level reliability testing, when using a signal-to-noise analysis, more than just one overall statistic should be reported (i.e., to demonstrate variation in reliability across providers). If a particular method yields only one statistic, this should be explained. In addition, reporting of results stratified by sample size is preferred (pg. 18, NQF Measure Evaluation Criteria).

[Response Begins]
See results in 2a.05.
In addition, in total, 30 of the 51 studies reported a reliability coefficient for each individual subscale in the PES-NWI. The reliability coefficients for the subscales ranged as follows: Nurse Participation in Hospital Affairs (.71-.93); Nursing Foundations for Quality of Care (.56-.91); Nurse Manager Ability, Leadership, and Support of Nurses (.63-.92); Staffing and Resource Adequacy (.65-.88); and Collegial Nurse–Physician Relations
(.62-.91).

[Response Ends]

2a.12. Interpret the results, in terms of how they demonstrate reliability.

(In other words, what do the results mean and what are the norms for the test conducted?)

[Response Begins]
The results indicate the instrument composite and subscales are highly reliable in terms of internal consistency reliability.
[Response Ends]

2b.01. Select the level of validity testing that was conducted.

[Response Begins]
Empirical validity testing
[Response Ends]

2b.02. For each level of testing checked above, describe the method of validity testing and what it tests.

Describe the steps—do not just name a method; what was tested, e.g., accuracy of data elements compared to authoritative source, relationship to another measure as expected; what statistical analysis was used.

[Response Begins]
October 2018 maintenance.
Title: Table reporting concurrent and predictive validity of the PES-NWI.
This table displays the hypothesized relationships of the instrument to patient health outcomes and quality and safety, the effect sizes and the significance levels.

Outcomes	**Hypothesized relationship with PES-NWI**	**Research study**	**Statistical evidence or associations**
Patient Record Outcomes	*	*	*
Mortality	significant negative relationship	*	*
*	*	Aiken et al (2008)	30 day inpatient mortality (-, OR= 0.91)
*	*	Aiken et al (2011) b	30 day inpatient mortality (-, OR= 0.93)
*	*	Cho et al (2014)	30 day inpatient mortality (-, OR= 0.52)
*	*	Friese et al (2008)	- Mortality
*	*	Nicely et al (2013)	30 day inpatient mortality (-, OR= 0.89)
*	*	Kelly (2014)	30 day inpatient mortality (-, OR= 0.97)
Hospital readmission	significant negative relationship	*	*
*	*	Gardner et al (2007)	— Hospitalizations
*	*	Ma and Park (2015)	30 day readmission (-, OR= 0.97)
*	*	McHugh et al (2016)	30 day readmission (-, OR= 0.84)
Complications	significant negative relationship	*	*
*	*	Friese et al (2008)	- Complications
Failure to rescue	significant negative relationship	*	*
*	*	Aiken et al (2008)	Failure to rescue (-, OR= 0.91)
*	*	Aiken et al (2011) b	Failure to rescue (-, OR= 0.93)
*	*	Friese et al (2008)	- Failure to rescue
*	*	Nicely et al (2013)	Failure to rescue (-, OR= 0.90)
Discharged without breastmilk	significant negative (?) relationship	*	*
*	*	Hallowell et al (2016)	Discharge on human milk and composite (+, $\beta = 0.04$) = Adjusted $R2 = 0.37$

Outcomes	Hypothesized relationship with PES-NWI	Research study	Statistical evidence or associations
*	*	Lake (2016)	Discharged without breastmilk (-, OR= 0.92)
Nurse Reported Adverse Outcomes	*	*	*
Total formal adverse event reports	significant positive relationship?	*	*
*	*	Kirwan et al (2013)	+ Total formal adverse event reports
NR nosocomial infection	significant negative relationship	*	*
*	*	Kutney-Lee, Lake, et al (2009)	- Nosocomial infections
*	*	Lake (2015)	Nosocomial infections (-, OR= 0.85)
*	*	Spence Laschinger and Leiter (2006)	- Nosocomial infections
NR falls	significant negative relationship	*	*
*	*	Cho et al (2016)	Falls with injury and composite (-, OR = 0.68)
*	*	Kutney-Lee, Lake, et al (2009)	- Falls with injury
*	*	Prezerakos et al (2015)	Patient falls and composite (-, OR= 0.02)
*	*	Spence Laschinger and Leiter (2006)	- Falls
NR medication errors	significant negative relationship	*	*
*	*	Cho et al (2016)	Medication errors and composite (-, OR=0.55)
*	*	Manojlovich and DeCicco (2007)	- Medication errors
*	*	Spence Laschinger and Leiter (2006)	- Medication errors
NR catheter-associated sepsis	significant negative relationship	*	*
*	*	Manojlovich and DeCicco (2007)	- Catheter-associated sepsis
NR pressure ulcer	significant negative relationship	*	*
*	*	Cho et al (2016)	Pressure ulcers and composite (-, OR = 0.61)
*	*	Choi and Staggs (2014)	Unit acquired pressure ulcers SRA (-, OR = 0.78)

Outcomes	Hypothesized relationship with PES-NWI	Research study	Statistical evidence or associations
*	*	Flynn et al (2010)	Pressure ulcers and composite (-, β=0.37)
*	*	Ma and Park (2015)	Hospital acquired pressure ulcers and composite (-, OR= 0.73)
NR UTI	significant negative relationship	*	*
*	*	Kelly (2013)	NR UTI (-, OR= 0.80)
NR bloodstream infection	significant negative relationship	*	*
*	*	Kelly (2013)	NR bloodstream infection (-, OR=0.77)
NR pneumonia	significant negative relationship	*	*
*	*	Kelly (2013)	NR pneumonia (-, OR= 0.80)
NR central line infection	significant negative relationship	*	*
*	*	Lake (2016)	NR central line infection (-, OR= 0.89)
Patient Satisfaction	*	*	*
Patient safety climate	significant positive relationship	*	*
*	*	Armstrong and Laschinger (2006)	+ Patient safety climate
*	*	Armstrong et al (2009)	+ Patient safety climate
Perceived quality of care	significant positive relationship	*	*
*	*	Gardner et al (2009)	+ Perceived quality of care
Nurses communicated well	significant positive relationship	*	*
*	*	Aiken et al (2012)	Nurses communicated well (+, OR=1.11)
*	*	You et al (2013)	Nurses communicated well (+, OR= 1.30)
Patient rates hospital highly	significant positive relationship	*	*
*	*	Aiken et al (2012)	Patient rates hospital highly (+, OR= 1.16)
*	*	Kutney-Lee et al (2015)	Patient rates hospital highly (+, OR= 1.17)
*	*	You et al (2013)	Patient rates hospital highly (+, OR= 1.29)
Patient satisfaction	significant positive relationship	*	*

Outcomes	Hypothesized relationship with PES-NWI	Research study	Statistical evidence or associations
*	*	Boev (2012)	Patient satisfaction and Nurse Manager Ability and Support of Nurses (+, β= 0.424)
*	*	Kutney-Lee, McHugh, et al (2009)	+ Patient satisfaction (HCAHPS)
*	*	Tei-Tominaga and Sato (2016)	Patient satisfaction and NPR (+, OR= 0.144)

* Cell intentionally left empty
[Response Ends]

2b.03. Provide the statistical results from validity testing.

Examples may include correlations or t-test results.

[Response Begins]
October 2021 maintenance.
Title: Table of Validity Testing of the PES-NWI.
This table displays the hypothesized relationships of the instrument to patient health outcomes and quality and safety, the effect sizes and the significance levels. What is notable in the 2021 results is a larger set of outcomes that have been linked statistically to better work environments.
This table reports results from regression models estimating the effect of a better work environment, as measured by the PES-NWI, on the specified outcome variable.

Outcomes	Hypothesized relationship with PES-NWI	Research study	Statistical evidence or associations
Patient Record Outcomes	*	*	OR or β [95% CI]
Mortality	significant negative relationship	*	*
*	*	Schlack (2021)	30-day mortality (-, OR = 0.86 [0.81, 0.92] <0.001)
*	*	Kang (2020)	30-day mortality (-, OR = .95 [0.94, 0.97]) *p* < .001
Failure to rescue	significant negative relationship	*	*
*	*	Schlack (2021)	(-,OR = 0.88 [0.83, 0.95] <0.001)
*	*	Kang (2020)	(-, OR = .94 [0.93, 0.96]) p < .001
Catheter-associated urinary tract infections	significant negative relationship	*	*
*	*	Landerfelt (2020)	(-, β = -1.34 (SE = 0.54) *P* < .05
Restraint rate	significant negative relationship	*	*
*	*	Olds (2021)	(-, β = −0.088, [−0.178–−0.014] *p* = 0.045
Length of stay	significant negative relationship	*	*

Outcomes	Hypothesized relationship with PES-NWI	Research study	Statistical evidence or associations
*	*	*	IRR = incident rate ratio
*	*	Schlack (2021)	(-, IRR = 0.96 (0.94, 0.99) *p* = 0.003
*	*	Brom (2021)	(-, IRR = 0.97 [0.95–0.99] *p* < .01.
Readmission within 7 days	significant negative relationship	*	*
*	*	Brom (2021)	(-, OR = 0.96 [0.93–0.99] p < .05.
Readmission within 30 days	significant negative relationship	*	*
*	*	Brom (2021)	(-, OR = 0.97 (0.94–0.99] p < .05.
Percent of long-stay nursing home residents with pressure ulcers	significant negative relationship	*	*
*	*	(White, 2020)	(-, β = 0.90 [1.64, 0.17] *p* = 0.02
Number of hospitalizations per nursing home resident year	significant negative relationship	*	*
*	*	(White, 2020)	(-, β = 0.08 [0.15,0.001] *p* = .05
Nurse Reported Adverse Outcomes	*	*	*
NR fall with injury	significant negative relationship	*	*
*	*	Smith (2019)	(-, OR = 0.50 [0.41–0.61] *p*<.01
NR medication errors	significant negative relationship	*	*
*	*	Smith (2019)	(-, OR = 0.61 [0.52–0.73] *p*<.01
NR pressure ulcer	significant negative relationship	*	*
*	*	Smith (2019)	(-, OR = 0.54 [0.43–0.67] *p*<.01)
NR UTI	significant negative relationship	*	*
*	*	Smith (2019)	(-, OR = 0.61 [0.51–0.73] *p*<.01)
Patient Satisfaction	*	*	*
Nurses communicated well	significant positive relationship	*	*
*	*	Brooks Carthon (2021)	β = 2.62 (2.10, 3.15), *P* <.001
*	*	*	*
Patient rates hospital highly	significant positive relationship	*	*
*	*	Brooks Carthon (2021)	β = 5.13 (4.17, 6.11), *P* <.001

Outcomes	Hypothesized relationship with PES-NWI	Research study	Statistical evidence or associations
Patient would definitely recommend	significant positive relationship	*	*
*	*	Brooks Carthon (2021)	β = 6.08 (4.98, 7.19), *P* <.001
Nurse Reported Measures	*	*	*
Quality of End of Life Care (graded C, D, or F)	significant negative relationship	*	*
*	*	Lasater (2019)	(-, OR = 0.45 (0.40–0.52) *P* < .001)

* Cell intentionally left empty

References:

1. Brom H, Brooks Carthon JM, Sloane D, McHugh M, Aiken L. Better nurse work environments associated with fewer readmissions and shorter length of stay among adults with ischemic stroke: A cross-sectional analysis of United States hospitals. *Res Nurs Health.* 2021;44(3):525-533.
2. Carthon JMB, Hatfield L, Brom H, et al. System-level improvements in work environments lead to lower nurse burnout and higher patient satisfaction. *Journal of nursing care quality.* 2021;36(1):7-13.
3. Kang XL, Brom HM, Lasater KB, McHugh MD. The association of nurse–physician teamwork and mortality in surgical patients. *Western journal of nursing research.* 2020;42(4):245-253.
4. Landerfelt PE, Lewis A, Li Y, Cimiotti JP. Nursing leadership and the reduction of catheter-associated urinary tract infection. *American Journal of Infection Control.* 2020;48(12):1546-1548.
5. Lasater KB, Sloane DM, McHugh MD, Aiken LH. Quality of end-of-life care and its association with nurse practice environments in US hospitals. *Journal of the American Geriatrics Society.* 2019;67(2):302-308.
6. Smith JG, Plover CM, McChesney MC, Lake ET. Rural hospital nursing skill mix and work environment associated with frequency of adverse events. *SAGE open nursing.* 2019;5:2377960819848246.
7. Schlak AE, Aiken LH, Chittams J, Poghosyan L, McHugh M. Leveraging the work environment to minimize the negative impact of nurse burnout on patient outcomes. *International Journal of Environmental Research and Public Health.* 2021;18(2):610.
8. Olds D, Cramer E. Predictors of physical restraint use on critical care units: an observational structural equation modeling approach. *Int J Nurs Stud.* 2021;118:103925.
9. White EM, Aiken LH, Sloane DM, McHugh MD. Nursing home work environment, care quality, registered nurse burnout and job dissatisfaction. *Geriatric Nursing.* 2020;41(2):158-164.

[Response Ends]

2b.04. Provide your interpretation of the results in terms of demonstrating validity. (i.e., what do the results mean and what are the norms for the test conducted?)

[Response Begins]

The results demonstrate concurrent and predictive validity. There are no norms for these tests.

[Response Ends]

2b.05. Describe the method for determining if statistically significant and clinically/practically meaningful differences in performance measure scores among the measured entities can be identified.

Describe the steps—do not just name a method; what statistical analysis was used? Do not just repeat the information provided in Importance to Measure and Report: Gap in Care/Disparities.

[Response Begins]

The method for determining both statistically significant and clinically/practically meaningful differences in performance measure scores among measured entities is by using descriptive statistics and peer comparisons for benchmarking and descriptive and inferential statistical analyses in empirical studies.

Benchmarking
National Database of Nursing Quality Indicators (developed by the American Nurses Association in 1994; purchased by Press Ganey Inc. in 2015). The instrument is used in an annual survey of hospital registered nurses in approximately 600 hospitals throughout the United States. These hospitals voluntarily participate in the annual survey for the purpose of benchmarking their values against their peer hospitals, such as academic medical centers, community hospitals, or hospitals in certain regions of the country.
The Leapfrog Group is a national network of employers and purchasers. Leapfrog conducts a free annual survey of all U.S. hospitals to assesses hospital safety, quality, and efficiency based on national performance measures that are of specific interest to health care purchasers and consumers. Results are free to the public at www.leapfroggroup.org/compare-hospitals. In 2018, 2,021 hospitals submitted a Leapfrog Hospital Survey, representing 54% of eligible hospitals and about 69% of all acute care hospitals beds. Leapfrog is initiating the measurement of this instrument #3450 in their members. Dr. Lake, measure steward, is a member of the Nursing Workforce expert panel for Leapfrog to advise on how to incorporate the PES-NWI into their annual survey. The panel has been meeting since September 2019. The expert panels can be viewed here: https://www.leapfroggroup.org/about/expert-panelists
Nursing Workforce is one of five NQF Safe practices that Leapfrog requires hospitals to report on. Leapfrog issues a Hospital Safety Grade that represents a hospital's performance on 28 different measures of patient safety. In addition, confidential benchmarking reports are used by hospitals to engage their leadership and staff and to guide quality improvement programs.
In addition, Leapfrog Hospital Survey data is used by payors and health plans in value-based purchasing programs.
The current plan as of Sept 2021 is to hold off on scoring and public reporting of the PES-NWI until the 2023 survey, given the current turmoil with COVID-19.

Empirical Research
The purpose of empirical research using this instrument is to determine how variation across hospitals in nurses' work environments relates to the quality and safety of patient care, to patient outcomes, and to nurses' job outcomes. A meta-analysis reporting these relationships was published in 2019 by the measure developer/steward Lake and colleagues: Lake E.T., Sanders J., Duan, R., Riman K., Schoenauer K, & Chen Y. (2019). A meta-analysis of the associations between the nurse work environment in hospitals and four sets of outcomes. *Medical Care* 57(5):353-361.
This publication documents the statistically significant and clinically/practically meaningful differences in performance scores across 21 independent samples spanning the period from the instrument's publication in 2002 through September 2018. As reported in the abstract: "Consistent, significant associations between the work environment and all outcome classes were identified. Better work environments were associated with lower odds of negative nurse outcomes (average OR of 0.71), poor safety or quality ratings (average OR of 0.65), and negative patient outcomes (average OR of 0.93), but higher odds of patient satisfaction (OR of 1.16)."
In these studies variation across hospitals is linked to variation across patients or nurses within those hospitals.
In the three years since measure endorsement (2018) there have been an additional 34 publications reporting empirical data from this instrument. These studies have not yet been summarized.

[Response Ends]

2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.

Examples may include number and percentage of entities with scores that were statistically significantly different from mean or some benchmark, different from expected; how was meaningful difference defined.

[Response Begins]
The evidence provided in 2b.06 demonstrates differences in performance across measured entities (hospitals) as follows: In hospitals with better work environments, as measured by higher scores on the PES-NWI, there are significantly lower odds of a patient experiencing 30-day mortality or failure to rescue (Schlak et al 2021, Kang et al 2020), catheter-associated urinary tract infections (Landerfelt et al 2020), which are quality indicators, as well as restraint use (Olds, 2021), and length of stay (Schlak et al 2021, Brom et al 2021), hospital readmission within 7 or 30 days (Brom et al 2021),

which are CMS pay-for-performance measures, and long stay nursing home residents suffering pressure ulcers or being hospitalized (White et al 2020).
In hospitals with better work environments, as measured by higher scores on the PES-NWI, we see a significant positive relationship to patient satisfaction as measured by HCAHPS measures: nurses communicated well, patient rates hospital highly and patient would definitely recommend the hospital (Brooks Carthon et al 2021).
Regarding nurse-reported measures, in better work environments, significantly fewer nurses report that falls with injury, medication errors, pressure ulcers, or urinary tract infections occur frequently (Smith et al 2019) or grade the quality of life care poorly (grade C, D, or F) (Lasater et al, 2019).

[Response Ends]

2b.07. Provide your interpretation of the results in terms of demonstrating the ability to identify statistically significant and/or clinically/practically meaningful differences in performance across measured entities.

In other words, what do the results mean in terms of statistical and meaningful differences?

[Response Begins]
The evidence from the meta-analysis (Lake et al 2019) shows differences in odds of patient mortality, readmission, satisfaction, as well as quality and safety of care and nurse burnout and intent-to-leave by a hospital's nurse work environment. The results regarding statistical and meaningful differences mean that the instrument detects performance differences across measured entities that can be addressed by hospital managers and leaders.
[Response Ends]

2b.08. Describe the method of testing conducted to identify the extent and distribution of missing data (or non-response) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and non-responders). Include how the specified handling of missing data minimizes bias.

Describe the steps—do not just name a method; what statistical analysis was used.

[Response Begins]
The method of testing to identify the extent and distribution of missing data (or non-response) depends on whether the testing was through benchmarking or empirical studies.
The National Database of Nursing Quality indicators reports the nursing-unit-level response rates for their annual RN survey as 68% overall. Hospitals offer incentives for nurses to complete the survey. The 68% response rate is considered high by today's standards, in which nurses are invited to do many surveys and response rates are typically poor.
The extent of missing data in empirical studies is reported on a study-by-study basis. The evidence from the many studies that have reported data from this instrument has not been compiled. In multiple studies conducted by the University of Pennsylvania Center for Health Outcomes and Policy Research, the differences between responders and non-responders to a survey including the PES-NWI were compared by conducting a "non-response" survey. A random sample of non-responders were invited to complete the survey and offered a monetary incentive (a $20 bill was included with the mailed survey). Then the answers of the former non-responders were compared with the original responders to determine if there was response bias in the PES-NWI data. The analysis revealed no significant differences between the original responders and the non-responders.
These results were reported in this working paper:
Smith, H. L. (2009) A double sample to minimize bias due to nonresponse in a mail survey. Population Studies Center Working Papers.
These results were also noted in this publication:
Lasater, K. B., et al. (2019). "A Methodology For Studying Organizational Performance: A Multistate Survey of Front-line Providers." Med Care 57(9): 742-749.

[Response Ends]

2b.09. Provide the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data.

For example, provide results of sensitivity analysis of the effect of various rules for missing data/non-response. If no empirical sensitivity analysis was conducted, identify the approaches for handling missing data that were considered and benefits and drawbacks of each).

[Response Begins]
From the Lasater, et al. (2019) paper noted in 2b.08, Table 2 reports "Comparison of Nurse Responses From the Main Survey and Nonrespondent Survey." They report results from the main survey (52,510 respondents) and the nonrespondent survey (1,168). The PES-NWI composite value for both groups had identical means and SDs: 2.2 (0.6). The test for significant difference was not significant (p = .501).
[Response Ends]

2b.10. Provide your interpretation of the results, in terms of demonstrating that performance results are not biased due to systematic missing data (or differences between responders and non-responders), and how the specified handling of missing data minimizes bias.

In other words, what do the results mean in terms of supporting the selected approach for missing data and what are the norms for the test conducted; if no empirical analysis was conducted, justify the selected approach for missing data.

[Response Begins]
The relatively high response rates obtained from surveys of nurses for benchmarking purposes combined with the evidence of no bias between responders and non-responders from the Lasater et al 2019 study (see 2b.08) support the interpretation that performance results are not biased due to differences between responders and non-responders.
[Response Ends]

Note: This item is directed to measures that are risk-adjusted (with or without social risk factors) OR to measures with more than one set of specifications/instructions (e.g., one set of specifications for how to identify and compute the measure from medical record abstraction and a different set of specifications for claims or eCQMs). It does not apply to measures that use more than one source of data in one set of specifications/instructions (e.g., claims data to identify the denominator and medical record abstraction for the numerator). Comparability is not required when comparing performance scores with and without social risk factors in the risk adjustment model. However, if comparability is not demonstrated for measures with more than one set of specifications/instructions, the different specifications (e.g., for medical records vs. claims) should be submitted as separate measures.

2b.11. Indicate whether there is more than one set of specifications for this measure.

[Response Begins]
No, there is only one set of specifications for this measure
[Response Ends]

2b.12. Describe the method of testing conducted to compare performance scores for the same entities across the different data sources/specifications.

Describe the steps—do not just name a method. Indicate what statistical analysis was used.

[Response Begins]
[Response Ends]

2b.13. Provide the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications.

Examples may include correlation, and/or rank order.

[Response Begins]
[Response Ends]

2b.14. Provide your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications.

In other words, what do the results mean and what are the norms for the test conducted.

[Response Begins]
[Response Ends]

2b.15. Indicate whether the measure uses exclusions.

[Response Begins]
N/A or no exclusions
[Response Ends]

2b.16. Describe the method of testing exclusions and what was tested.

Describe the steps—do not just name a method; what was tested, e.g., whether exclusions affect overall performance scores; what statistical analysis was used?

[Response Begins]
There are no exclusions.
[Response Ends]

2b.17. Provide the statistical results from testing exclusions.

Include overall number and percentage of individuals excluded, frequency distribution of exclusions across measured entities, and impact on performance measure scores.

[Response Begins]
N/A. There are no exclusions.
[Response Ends]

2b.18. Provide your interpretation of the results, in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results.

In other words, the value outweighs the burden of increased data collection and analysis. Note: If patient preference is an exclusion, the measure must be specified so that the effect on the performance score is transparent, e.g., scores with and without exclusion.

[Response Begins]
N/A. There are no exclusions.
[Response Ends]

2b.19. Check all methods used to address risk factors.

[Response Begins]
No risk adjustment or stratification
[Response Ends]

2b.20. If using statistical risk models, provide detailed risk model specifications, including the risk model method, risk factors, risk factor data sources, coefficients, equations, codes with descriptors, and definitions.

[Response Begins]
[Response Ends]

2b.21. If an outcome or resource use measure is not risk-adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (i.e., case mix) is not needed to achieve fair comparisons across measured entities.

[Response Begins]
This instrument is a nurse survey. We believe that all permanent staff nurses are qualified to evaluate their work environment without accounting for differences in nurse characteristics to achieve fair comparisons across measured hospitals.
[Response Ends]

2b.22. Select all applicable resources and methods used to develop the conceptual model of how social risk impacts this outcome.

[Response Begins]
[Response Ends]

2b.23. Describe the conceptual and statistical methods and criteria used to test and select patient-level risk factors (e.g., clinical factors, social risk factors) used in the statistical risk model or for stratification by risk.

Please be sure to address the following: potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10 or other statistical tests; correlation of x or higher. Patient factors should be present at the start of care, if applicable. Also discuss any "ordering" of risk factor inclusion; note whether social risk factors are added after all clinical factors. Discuss any considerations regarding data sources (e.g., availability, specificity).

[Response Begins]
[Response Ends]

2b.24. Detail the statistical results of the analyses used to test and select risk factors for inclusion in or exclusion from the risk model/stratification.

[Response Begins]
[Response Ends]

2b.25. Describe the analyses and interpretation resulting in the decision to select or not select social risk factors.

Examples may include prevalence of the factor across measured entities, availability of the data source, empirical association with the outcome, contribution of unique variation in the outcome, or assessment of between-unit effects and within-unit effects. Also describe the impact of adjusting for risk (or making no adjustment) on providers at high or low extremes of risk.

[Response Begins]
[Response Ends]

2b.26. Describe the method of testing/analysis used to develop and validate the adequacy of the statistical model or stratification approach (describe the steps—do not just name a method; what statistical analysis was used). Provide the statistical results from testing the approach to control for differences in patient characteristics (i.e., case mix) below. If stratified ONLY, enter "N/A" for questions about the statistical risk model discrimination and calibration statistics.

Validation testing should be conducted in a data set that is separate from the one used to develop the model.

[Response Begins]
[Response Ends]

2b.27. Provide risk model discrimination statistics.

For example, provide c-statistics or R-squared values.

[Response Begins]
[Response Ends]

2b.28. Provide the statistical risk model calibration statistics (e.g., Hosmer-Lemeshow statistic).

[Response Begins]
not applicable.
[Response Ends]

2b.29. Provide the risk decile plots or calibration curves used in calibrating the statistical risk model.

The preferred file format is .png, but most image formats are acceptable.

[Response Begins]
[Response Ends]

2b.30. Provide the results of the risk stratification analysis.

[Response Begins]
[Response Ends]

2b.31. Provide your interpretation of the results, in terms of demonstrating adequacy of controlling for differences in patient characteristics (i.e., case mix).

In other words, what do the results mean and what are the norms for the test conducted?

[Response Begins]
[Response Ends]

2b.32. Describe any additional testing conducted to justify the risk adjustment approach used in specifying the measure.

Not required but would provide additional support of adequacy of the risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed.

[Response Begins]
[Response Ends]

Criteria 3: Feasibility

Extent to which the specifications including measure logic, require data that are readily available or could be captured without undue burden and can be implemented for performance measurement.

3.01. Check all methods below that are used to generate the data elements needed to compute the measure score.

[Response Begins]
Other (Please describe)
[Other (Please describe) Please Explain]
Data elements are generated by survey of nurses.

[Response Ends]

3.02. Detail to what extent the specified data elements are available electronically in defined fields.

In other words, indicate whether data elements that are needed to compute the performance measure score are in defined, computer-readable fields.
[Response Begins]
Patient/family reported information (may be electronic or paper)
[Response Ends]

3.03. If ALL the data elements needed to compute the performance measure score are not from electronic sources, specify a credible, near-term path to electronic capture, OR provide a rationale for using data elements not from electronic sources.

[Response Begins]
This is nurse reported information. This measure is an eMeasure in the National Database of Nursing Quality Indicators, a voluntary benchmarking hospital network.
2021 Submission:
As above, this measure continues to be nurse-reported in the National Database of Nursing Quality Indicators. In addition, the Leapfrog Group plans to require the PES-NWI as a reported survey beginning in 2024.
2018 Submission:
This measure is also an eMeasure in the Veterans Administration and the Military Hospital system.
Here we support feasibility by presenting the numbers of hospitals, nursing units, and nurses, and response rates across years.
NDNQI VA
response rate # hospitals # units #nurses response rate # hospitals #nurses
2012 0.43 139 23,831
2013 0.72 574 11,264 206,978 0.43 138 24,166
2014 0.68 395 7,557 131,619 0.47 141 28,930
2015 0.69 453 9,168 157,531 0.52 141 33,446
2016 0.72 349 8,236 132,764 0.53 141 35,700
2017 0.70 384 8,520 147,568 0.54 141 37,305
2018 0.56 141 38,967

[Response Ends]

3.04. Describe any efforts to develop an eCQM.

[Response Begins]
An eCQM is not applicable to this measure. Presently there are no efforts to develop an eCQM. Nurses may answer the survey electronically but that is a hospital-specific decision.
[Response Ends]

3.06. Describe difficulties (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.

[Response Begins]
For hospitals in the NDNQI or the VA or in the state of Colorado that have decided to survey their nurses, they experience no difficulties with data collection, availability, sampling, confidentiality, time/cost of data collection or other feasibility /implementation issues. The amount of missing data is very small (less than 2%). The timing/frequency is typically annual.
[Response Ends]

Consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

3.07. Detail any fees, licensing, or other requirements to use any aspect of the measure as specified (e.g., value/code set, risk model, programming code, algorithm),

Attach the fee schedule here, if applicable.

[Response Begins]
none
[Response Ends]

Criteria 4: Use and Usability

Extent to which potential audiences (e.g., consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient healthcare for individuals or populations.

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.
NQF-endorsed measures are expected to be used in at least one accountability application within 3 years and publicly reported within 6 years of initial endorsement, in addition to demonstrating performance improvement.

4a.01. Check all current uses. For each current use checked, please provide:

Name of program and sponsor

URL

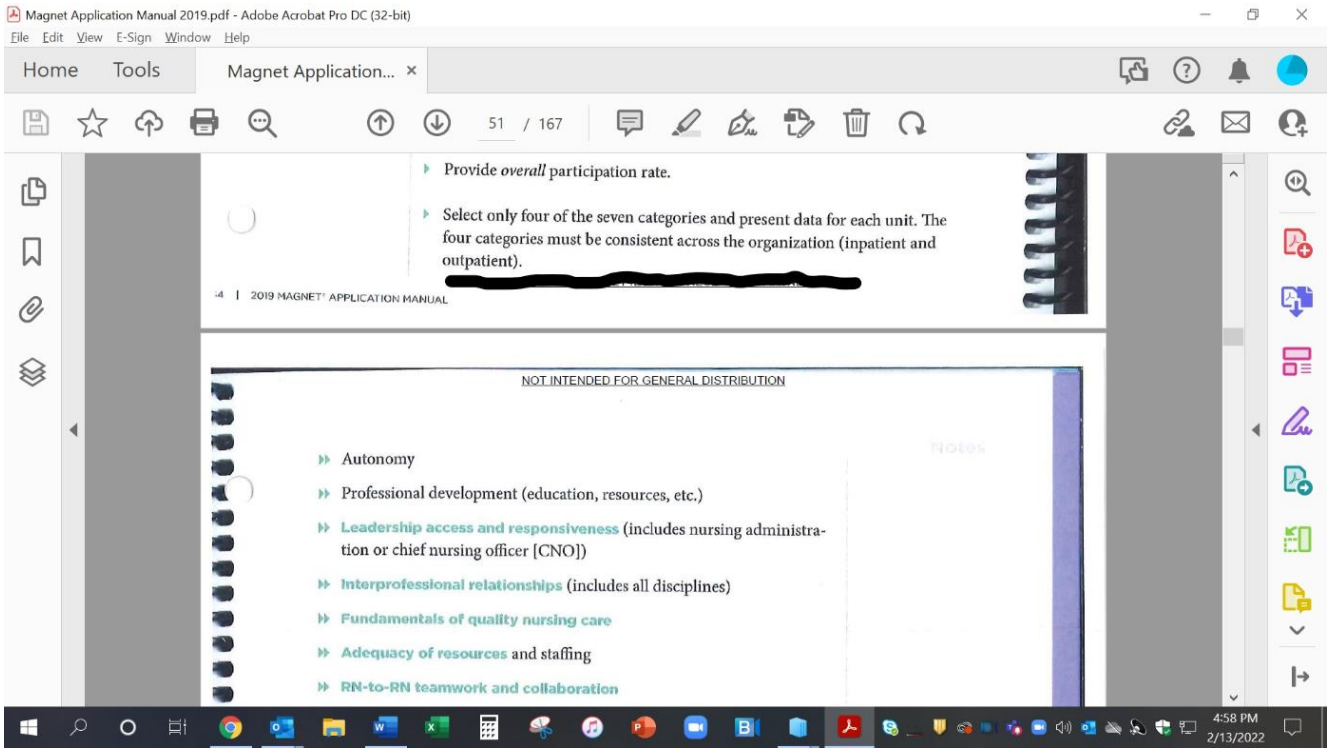
Purpose

Geographic area and number and percentage of accountable entities and patients included

Level of measurement and setting

[Response Begins]
Public Reporting
[Public Reporting Please Explain]
State of Colorado does annual reporting of PES-NWI from all hospitals with 100 or more beds:
http://www.cohospitalquality.org:8080/corda/dashboards/COLORADO_REPORT_CARD_BY_HOSPITAL/main.dashxml

Professional Certification or Recognition Program
[Professional Certification or Recognition Program Please Explain]
The 2019 Magnet Hospital Accreditation program requires measurement of multiple domains in the PES-NWI. EP2E0 requires the reporting of four of seven categories with data presented for each nursing unit. These seven categories include four subscales of the PES-NWI: interprofessional relationships, fundamentals of quality nursing care, adequacy of resources and staffing, and leadership access and responsiveness. See attached excerpt from the 2019 accreditation manual.
This image, entitled "Magnet Application Manual 2019 Excerpt", displays an excerpt from the magnet handbook indicating that at least four of seven categories must be provided to fulfill this magnet requirement. The seven categories include four that align with PES-NWI subscales: leadership access and responsiveness, interprofessional relationships, fundamentals of quality nursing care, and adequacy of resources and staffing.



Title: Excerpt from the Magnet Hospital Accreditation Manual

Quality Improvement with Benchmarking (external benchmarking to multiple organizations)

[Quality Improvement with Benchmarking (external benchmarking to multiple organizations) Please Explain]

The National Database of Nursing Quality Indicators provides benchmarking reports to all participating hospitals to compare across peer groups, such as by hospital size, teaching status, or region.

Quality Improvement (Internal to the specific organization)

[Quality Improvement (Internal to the specific organization) Please Explain]

Specific hospitals may use the survey results to conduct quality improvement. For example, in this study below, a hospital identified the nurse managers rated most highly on that subscale and these nurse managers mentored other nurse managers to improve their abilities.

Anderson, B. J., et al. (2010). "Listening to nursing leaders: Using national database of nursing quality indicators data to study excellence in nursing leadership." Journal of Nursing Administration 40(4): 182-187.

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Public reporting
Professional Certification or Recognition Program
Quality Improvement with Benchmarking (external benchmarking to multiple organizations)
Quality Improvement (internal to the specific organization)

[Response Ends]

4a.03. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing), explain why the measure is not in use.

For example, do policies or actions of the developer/steward or accountable entities restrict access to performance results or block implementation?

[Response Begins]

n/a
[Response Ends]

4a.04. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes: used in any accountability application within 3 years, and publicly reported within 6 years of initial endorsement.

A credible plan includes the specific program, purpose, intended audience, and timeline for implementing the measure within the specified timeframes. A plan for accountability applications addresses mechanisms for data aggregation and reporting.

[Response Begins]
n/a
[Response Ends]

4a.05. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

Detail how many and which types of measured entities and/or others were included. If only a sample of measured entities were included, describe the full population and how the sample was selected.

[Response Begins]
In the NDNQI, the VA, and the military hospitals, the performance results are shared in reports and dashboards with hospital managers to identify and address weaknesses in the nursing practice environments in their facilities. The number of facilities equaled 384 NDNQI hospitals in 2017 as well as all 141 VA and all 13 army hospitals nationally. The NDNQI is a national voluntary benchmarking database to track nursing quality indicators. All VA and army hospitals collect the measure data. Interpretation is provided by NDNQI site coordinators in each hospital and at the VA and army hospitals by their quality and safety staff.
[Response Ends]

4a.06. Describe the process for providing measure results, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

[Response Begins]
Measure results are reported annually to the facilities that complete the survey. The data provided are descriptive statistics as well as trends for the subscales and the composite score.
[Response Ends]

4a.07. Summarize the feedback on measure performance and implementation from the measured entities and others. Describe how feedback was obtained.

[Response Begins]
The NDNQI publishes monographs and holds conferences to provide exemplars of excellence to the facilities that collect, report, and evaluate the nursing-unit level data to assess the quality of nursing care and to transform the nursing units and improve outcomes. Press Ganey acquired NDNQI in 2015 and prepares annual strategic insight reports that are publicly available.
References to the monographs/reports:
Montalvo, I., & Dunton, N. (Eds.). (2007). Transforming nursing data into quality care: Profiles of quality improvement in US healthcare facilities. American Nurses Association.
Dunton, N., & Montalvo, I. (2009). Sustained improvement in nursing quality: Hospital performance on NDNQI indicators, 2007-2008. American Nurses Association.
Press Ganey Inc. 2017. Achieving Excellence: The Convergence of Safety, Quality, Experience and Caregiver Engagement http://healthcare.pressganey.com/2017-Strategic-Insights?s=White_Paper-PGPost
Example of a Conference:

Dunton, N., Staggs, V. & Potter, C. January 25, 2012. NDNQI Research Findings for the Advanced Site Coordinator. Preconference Workshop 002

[Response Ends]

4a.08. Summarize the feedback obtained from those being measured.

[Response Begins]
The increasing trend in completion of the measure in the NDNQI membership indicates that the measure is valued by the member facilities. The inclusion of the measure in hospital dashboards indicates the measure is valued for monitoring quality.
[Response Ends]

4a.09. Summarize the feedback obtained from other users.

[Response Begins]
The display of measure results in annual reports and manager dashboards demonstrates that the measure results are valuable to users.
[Response Ends]

4a.10. Describe how the feedback described has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

[Response Begins]
Typical feedback about the measure is that a reduction in length and testing for use in non-hospital settings is desired. The measure has not been modified to date although reduction in survey length is planned for a future endorsement period.
[Response Ends]

4b.01. You may refer to data provided in Importance to Measure and Report: Gap in Care/Disparities, but do not repeat here. Discuss any progress on improvement (trends in performance results, number and percentage of people receiving high-quality healthcare; Geographic area and number and percentage of accountable entities and patients included). If no improvement was demonstrated, provide an explanation. If not in use for performance improvement at the time of initial endorsement, provide a credible rationale that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

[Response Begins]
Maintenance of Endorsement (October 2021):
As reported in 1b, the use of the instrument by the NDNQI, the VA, and the military hospitals is for benchmarking and performance improvement.
The evidence presented above by Swiger et al (2017), shows that hospitals that have achieved Magnet recognition for meeting standards of nursing excellence have improved performance by having better work environments as compared to non-Magnet hospitals.
Additionally, this publication reports use of the instrument to improve nursing leadership, one subscale of the instrument:
Anderson, B. J., Manno, M., O´Connor, P., & Gallagher, E. (2010). Listening to nursing leaders: Using national database of nursing quality indicators data to study excellence in nursing leadership. Journal of Nursing Administration, 40(4), 182-187.
The article aims to examine nurse leadership qualities that create healthy work environments conducive to delivery of quality bedside care. The PES-NWI was used to assess qualities of exemplary nurse managers chosen by their staff. Researchers concluded that effective nurse leaders emphasized visibility, communication, and valued respect and empathy. These leadership strategies help to create healthy work environments that support nurse job satisfaction, nurse retention, and quality patient care delivery.

[Response Ends]

4b.02. Explain any unexpected findings (positive or negative) during implementation of this measure, including unintended impacts on patients.

[Response Begins]
As noted by Warshawsky and Havens, 2011 it is important that scoring and reporting of the PES-NWI be done consistently. There was inconsistency in reporting of subscales and composites across the many studies. There has also been variation in the unit of analysis for reporting, specifically nurse, nursing unit and organizational levels.
Maintenance of Endorsement (October 2021):
There have been no unexpected findings except for some non-significant results in some studies in the new literature reviews. Nonsignificant results may be related to small sample sizes in some studies.

[Response Ends]

4b.03. Explain any unexpected benefits realized from implementation of this measure.

[Response Begins]
Maintenance of Endorsement (October 2021)
The unexpected benefits have been the worldwide use of the measure, generating comparable evidence to improve nursing work environments globally, and thereby improve patient safety and quality outcomes.

[Response Ends]

Criteria 5: 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.

If you are updating a maintenance measure submission for the first time in MIMS, please note that the previous related and competing data appearing in question 5.03 may need to be entered in to 5.01 and 5.02, if the measures are NQF endorsed. Please review and update questions 5.01, 5.02, and 5.03 accordingly.

5.01. Search and select all NQF-endorsed related measures (conceptually, either same measure focus or target population).

(Can search and select measures.)
[Response Begins]
0204: Skill mix (Registered Nurse [RN], Licensed Vocational/Practical Nurse [LVN/LPN], unlicensed assistive personnel [UAP], and contract)
0205: Nursing Hours per Patient Day
[Response Ends]

5.02. Search and select all NQF-endorsed competing measures (conceptually, the measures have both the same measure focus or target population).

(Can search and select measures.)
[Response Begins]
[Response Ends]

5.03. If there are related or competing measures to this measure, but they are not NQF-endorsed, please indicate the measure title and steward.

[Response Begins]
N/A
[Response Ends]

5.04. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s), indicate whether the measure specifications are harmonized to the extent possible.

[Response Begins]
Yes
[Response Ends]

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

[Response Begins]
N/A
[Response Ends]

5.06. Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality). Alternatively, justify endorsing an additional measure.

Provide analyses when possible.

[Response Begins]

N/A
[Response Ends]