NATIONAL QUALITY FORUM
Driving measurable health improvements together

MEASURE WORKSHEET

This document summarizes the evaluation of the measure as it progresses through National Quality Forum's (NQF) Consensus Development Process (CDP). The information submitted by the 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 #: 3498e

Corresponding Measures:

Measure Title: Hospital Harm - Pressure Injury

Measure Steward: Centers for Medicare & Medicaid Services

sp.02. Brief Description of Measure: This electronic clinical quality measure (eCQM) assesses the proportion of inpatient hospitalizations for patients ages 18 years and older at the start of the encounter who suffer the harm of developing a new stage 2, stage 3, stage 4, deep tissue, or unstageable pressure injury.

1b.01. Developer Rationale:

Current (2022) Submission:

This safety eCQM captures the number of patients who experience harm in the form of a pressure injury, during their inpatient hospitalization. The incidence of pressure injuries in hospitalized patients has been estimated at 5.4 per 10,000 patient-days and the rate of hospital-acquired pressure injuries has been estimated at 8.4% (Li et al., 2020). While studies have identified a reduction in the incidence of pressure injuries from 1990 to 2017 (Siotos et al., 2022), other studies have found that pressure injuries are consistently underreported, with lower-stage pressure ulcers the least likely to be reported (Chen et al., 2022). Over 50% of reported pressure injuries in hospitals were Stage 2 or higher (Li et al., 2020). Hospital-acquired pressure injuries are serious events and one of the most common patient harms. Pressure injuries commonly cause local infection, osteomyelitis, anemia, and sepsis (Brem et al., 2010), in addition to causing significant depression, pain, and discomfort to patients (Gunningberg et al., 2011). Hospital-acquired pressure injuries are associated with 1.5 to 2.0 times greater risk of 30, 60, and 90-day readmissions (Wassel et al, 2020). Any Stage 3, Stage 4, or unstageable pressure ulcer acquired after admission/presentation to a healthcare setting is considered a serious reportable event by the National Quality Forum (NQF) (National Quality Forum, 2011).

Systematically assessing patients who develop new pressure injuries while in the hospital setting will provide hospitals with a reliable and timely measurement, to more reliably assess harm reduction efforts and modify their improvement efforts in near realtime. The intent of this measure is to incentivize greater achievements in reducing harms and enhance hospital performance on patient safety outcomes. Stage 2 pressure injuries constitute a very real patient harm that should be monitored and addressed; however, the relative level of harm is less than with Stage 3, Stage 4, Unstageable pressure injuries and potentially DTI. (National Pressure Ulcer Advisory Panel, 2019). The revised measure specification allows a 24-hour time window for accurate and timely identification of stage 2, 3, 4, or unstageable pressure injury present on admission. The revised measure specification allows a 72-hour time window for accurate and timely identification of deep tissue injury (DTI) because early diagnosis of DTI allows prompt identification of possible causes, initiation of treatment, and implementation of preventive strategies. Up to 72 hours can lapse between the precipitating pressure event and the onset of purple or maroon skin, so a longer time window is needed to exclude cases when the precipitating event occurred before the patient's admission. (Wound Management and Prevention, 2018).

Previous (2019) Submission:

This safety eCQM captures the number of patients who experience harm in the form of a pressure injury, during their inpatient hospitalization. Hospital-acquired pressure injuries are serious events and one of the most common patient harms. Pressure injuries commonly cause local infection, osteomyelitis, anemia, and sepsis (Brem, et al., 2010), in addition to causing significant depression, pain, and discomfort to patients (Gunningberg et al., 2011). Pressure injury is considered a serious reportable event by the National Quality Forum (NQF) (Centers for Medicare and Medicaid Services, 2015). CMS also established non-payment for pressure injury (National Quality Forum, 2016), and the rate of pressure injuries is considered an indicator of the quality of nursing care a hospital provides (National Quality Forum, 2005).

It is widely accepted that the risk of developing a pressure injury can be reduced through best practices such as frequent repositioning, proper skin care, and specialized cushions or beds (Berlowitz, et al., 2012). Systematically measuring patients who develop new pressure injuries while in the hospital setting will provide hospitals with a reliable and timely measurement, to more reliably assess harm reduction efforts and modify their improvement efforts in near real-time. This eCQM will fill a gap in measurement and provide incentives for hospitals' quality improvement. Although several pressure injury measures are currently in use, there are no electronic health record (EHR)-based measures intended for use in acute care hospitals. In addition, the intent of this measure is to incentivize greater achievements in reducing harms and enhance hospital performance on patient safety outcomes.

---

sp.12. Numerator Statement:

Inpatient hospitalizations for patients with a new deep tissue pressure injury (DTI) or stage 2, 3, 4, or unstageable pressure injury, as evidenced by any of the following:

- A diagnosis of DTI with the DTI not present on admission;
- A diagnosis of stage 2, 3, 4 or unstageable pressure injury with the pressure injury diagnosis not present on admission;
- A DTI found on exam greater than 72 hours after the start of the encounter; or
- A stage 2, 3, 4 or unstageable pressure injury found on exam greater than 24 hours after the start of the encounter.

sp.14. Denominator Statement: Inpatient hospitalizations where the patient is 18 years of age or older at the start of the encounter.

sp.16. Denominator Exclusions:

- Inpatient hospitalizations for patients with a DTI or stage 2, 3, 4 or unstageable pressure injury diagnosis present on admission.
- Inpatient hospitalizations for patients with a DTI found on exam within 72 hours of the start of the encounter.
- Inpatient hospitalizations for patients with a stage 2, 3, 4, or unstageable pressure injury found on exam within 24 hours of the start of the encounter.
- Inpatient hospitalizations for patients with diagnosis of a COVID-19 infection during the encounter.

Measure Type: Outcome

sp.28. Data Source:

Electronic Health Records

sp.07. Level of Analysis:

Facility

Criteria 1: Importance to Measure and Report

1a. Evidence

1a. Evidence. The evidence requirements for a ***health outcome*** measure include providing empirical data that demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service; if these data not available, data demonstrating wide variation in performance can be used, assuming the data are from a robust number of providers and the results are not subject to systematic bias. For measures derived from a patient report, the evidence also should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.

The developer provides the following description for this measure:

- This is a new electronic clinical quality measure (eCQM) outcome measure at the facility level that assesses the proportion of inpatient hospitalizations for patients ages 18 years and older at the start of the encounter who suffer the harm of developing a new stage 2, stage 3, stage 4, deep tissue, or unstageable pressure injury.
- The developer provides a logic model that depicts an increased monitoring of patients at risk for pressure injury, including risk and skin assessments, frequent repositioning, proper skin care, and specified cushions/beds leading to lower rates of pressure injuries acquired during hospitalization. The logic model then shows lower rates of pressure injuries acquired during hospitalization leading to lower rates of HAPI-associated infections, lower rates of sepsis, reduced pain, and reduced discomfort.
- This measure was originally submitted for NQF endorsement in 2019. The Patient Safety Standing Committee reviewed and recommended the measure during its measure evaluation in-person meeting, however, the developer chose to withdraw the measure prior to CSAC review stating "that they are considering substantive changes and assessing potential impacts."

Summary:

- The developer cites two evidence-based guidelines that outline prevention of pressure ulcers:
    - The developer cites three strong positive recommendation from the Prevention and Treatment of Pressure Ulcers/Injuries: Clinical Practice Guideline (2019) published by the European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel and Pan Pacific Pressure Injury Alliance.
        - Grade A recommendation to consider individuals with a Category/Stage I pressure injury to be at risk of developing a Category/Stage II or greater pressure injury.
        - Grade A recommendation to Inspect the skin of individuals at risk of pressure injuries to identify presence of erythema.
        - Grade A recommendation to develop and implement a structured, tailored and multi-faceted quality improvement program to reduce the incidence of pressure injuries at an organizational level.
    - The developer cites one strong recommendation from the Risk Assessment and Prevention of Pressure Ulcers: A Clinical Practice Guideline (2015) published by the American College of Physicians (ACP).
        - The ACP recommends that clinicians should choose advanced static mattresses or advanced static overlays in patients who are at an increased risk of developing pressure ulcers.

Question for the Standing Committee:

- *Is there at least one thing that the provider can do to achieve a change in the measure results?*

Questions for the Standing Committee:

- *What is the relationship between this measure and patient outcomes?*
- *How strong is the evidence for this relationship?*
- *Is the evidence directly applicable to the process of care being measured?*

Guidance From the Evidence Algorithm

Outcome measure (Box 1) -> Relationship between measured health outcome and one healthcare action is demonstrated by empirical data(box 2) -> Pass

Preliminary rating for evidence: ☒ **Pass** ☐ **No Pass**

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

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

- The developer provides data from 18 hospitals with varying bed size, geographic location, teaching status, urbanicity, and EHR systems for the full 2020 calendar year to demonstrate performance gap.
    - The developer notes hospitals' performance rate in pressure injury ranged from a minimum of 0 to a maximum of 2.02 per 100 qualified inpatient admissions.
- The developer reported a weighted average of 1.06 per 100 qualified inpatient admissions and a standard deviation of 0.56 per 100 qualified inpatient admissions across test sites.
- The developer reported an interquartile range of 0.63 per 100 qualified inpatient admissions.
- The developer concluded that that prior studies confirm that significant variation in rates of hospital acquired pressure injuries exists between hospitals within multi-hospital systems.

Disparities

- The developer reports the following trends for the subgroups of age, sex, ethnicity, and primary payer across all test sites and within the measure denominator population:
    - Age: Patients aged 65 or above were more likely to experience HA-PI than those 64 or younger.
    - Sex: Male patients had higher chance of experiencing hospital acquired (HA) PI than female patients.
    - Ethnicity: Non-Hispanic African Americans had a moderately higher chance of developing HA-PI than other ethnicities.
    - Primary Payer: Medicare beneficiaries were more likely than Medicaid beneficiaries or commercially insured patients to experience PI during hospitalization.

Questions for the Standing Committee:

- *Is there a gap in care that warrants a national performance measure?*

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

Criteria 2: Scientific Acceptability of Measure Properties

Complex measure evaluated by the Scientific Methods Panel (SMP)? ☐ **Yes** ☒ **No**

Evaluators: Staff

2a. Reliability: Specifications and Testing

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

- The submitted measure specification follows established technical specifications for electronic clinical quality measures (eCQMs) (Quality Data Model [QDM], health quality measure format [HQMF], and Clinical Quality Language [CQL]) as indicated in subcriterion 2a1.
- The submitted measure specification is fully represented and is not hindered by any limitations in the established technical specifications for eCQMs.

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

Specifications:

- Measure specifications are clear and precise.
- eCQMs as specified using the latest industry-accepted eCQM technical specifications: HQMF, QDM, CQL, and value sets vetted through the National Library of Medicine's (NLM) Value Set Authority Center (VSAC).

Reliability Testing:

- During the 2019 review, the SMP had no issues with reliability testing and passed this measure on reliability with a rating of high.
- Reliability testing conducted at the Accountable Entity Level:
    - The developer conducted a signal-to-noise (SNR) analysis as well as intra-class correlation coefficient (ICC) via the split-half sample approach on electronic health records from 128,323 qualified inpatient encounters at 18 hospitals with a range of 25 to 499 beds from January 2020 to December 2020.
    - The developer reported that SNR ranged from 0.86 to 1.00, with the mean and median equal to 0.96 and 0.97 respectively.
    - The developer reported the 100 estimated ICCs had a median of 0.99 and a mean ranging from 0.79 to 0.97.
    - The developer concluded that score level reliability is robust.

Guidance From the Reliability Algorithm

Specifications are precise, unambiguous, and complete (box 1) Yes -> Empirical reliability testing conducted using statistical testing (box 2) Yes -> Reliability testing conducted with computed performance measure scores (box 4) Yes -> Method was appropriate for assessing the proportion of variability due to real differences

among measured entities (box 5) Yes -> High certainty or confidence that the performance measure scores are reliable (box 6a) Yes -> High

Questions for the Standing Committee regarding reliability:

- *Do you have any concerns that the measure cannot be consistently implemented (i.e., are the measure specifications adequate)?*

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

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

2b2. Validity testing should demonstrate that 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

- As noted above, this measure was originally submitted for NQF endorsement in 2019. In their preliminary analyses, SMP subgroup 4 members did not reach consensus on the validity of the measure. There was some concern among reviewers about weak validity results in one of the tested datasets; reviewers suggested that inconsistent use of structured fields in EHRs raises concerns about data quality and documentation practices; however, the subgroup ultimately passed the measure on validity.
- The 2019 Patient Safety Standing Committee then reviewed and recommended the measure during its measure evaluation in-person meeting, however, the developer chose to withdraw the measure prior to CSAC review stating that they were considering "substantive changes and assessing potential impacts."
- Validity testing conducted at the Patient/Encounter Level:
    - To conduct data element validity testing, the developer compared data exported from the EHR to data manually abstracted from medical charts for a subsample of the measure population.
    - The developer states that empirical validity was calculated using frequency of missingness, percent match agreement, positive predicative value (PPV), sensitivity, negative predicative value (NPV), and specificity.
    - The developer found that all measure's data elements are consistently stored in the EHR and can be accurately exported for calculation.
    - The developer reported PPV results ranging from 0.97 to 1.0 as well as near perfect sensitivity, NPV, and specificity across measure components and sites.
    - The developer concluded that there was clear evidence that this measure, as currently specified, can detect true hospital acquired pressure injuries with high precision and that the measure will have very low false positives in implementation.
- Validity testing conducted at the Accountable Entity Level:
    - The developer states that measure score validity was conducted by assessing convergent validity to determine whether multiple measures are correlated.
    - The developer collected test site patient safety outcomes from five related infection measures (e.g., Central line-associated bloodstream infection (CLABSI), Catheter-associated urinary tract infection (CAUTI), Surgical Site Infection from colon surgery, Methicillin-resistant

Staphylococcus aureus (MRSA) bacteremia, and Clostridium difficile (C. diff) infection) on Hospital Care Compare then estimated Spearman's rank correlation coefficients.

- The developer also collected test site patient safety outcomes from a set of 12 related Hospital Consumer Assessment of Healthcare Providers and Systems (HCAHPS) measures (e.g., Staff responsiveness (star rating), Communication about medicine (star rating), Discharge information (star rating), Care transition (star rating), Overall rating of hospital (star rating), Nurse communication (linear mean score), Staff responsiveness (linear mean score), Communication about medicine (linear mean score), Discharge information (linear mean score), Care transition (linear mean score), Overall rating of hospital (linear mean score) on Hospital Care Compare then estimated Spearman's rank correlation coefficients.
- The developer states that positive (or negative) correlations provide support for the score level validity.
- The developer found that the rate of pressure injury correlates with several independently collected and NQF-endorsed measures of hospital harms, such as healthcare associated MRSA bloodstream and intestinal infections.
- The developer reports higher pressure injury rates are inversely related to patients' experiences with hospital care, particularly nursing components such as nurse communications, staff responsiveness, and discharge information.
- The developer concluded that construct validity for NQF 3498e at the hospital level is moderate, and correlational directions are largely in line with expectations.

Exclusions

- The developer reports that exclusions were tested by removing measure exclusion criterion one at a time from the logic and calculating the effect on the numerator and denominator, as well as the observed measure rate as a result. The developer also used parallel-form comparison to evaluate whether patients excluded from the denominator per the EHR truly met the clinical intent for exclusion.
- The developer concluded that all exclusions are necessary to reduce the measure's false positive rate and to prevent hospitals from being penalized by appropriate management of pre-existing or comorbid conditions, such as COVID-19.
- The developer notes that the COVID-19 exclusion is expected to be a temporary exclusion and will be reconsidered before the next endorsement cycle.

Risk Adjustment

- The measure is not risk-adjusted or stratified.

Meaningful Differences

- The developer reports that in order to identify meaningful differences, full denominator data was used to calculate the hospital-level measure performance rate and its 95 percent confidence interval for each of the 18 test sites. The developer then calculated the system-wide, weighted average measure performance rate across sites. Finally, the developer compared each test site's performance in pressure injury against the system-wide average and gauge if its performance deviates significantly from the weighted mean.
- The developer also estimated a linear regression model, relating the incidence of pressure injury to a set of hospital-specific indicators (or hospital-specific fixed effects) with a generalized T-test.

- The developer reports that testing data show that measure performance rates ranged from 0 to 2.02%.
- The developer also notes that several hospitals' performance rates are consistently below the system-wide average while a few others are above that mean.
- The developer concludes that regression results demonstrate that the measure can detect clinically meaningful differences in pressure injury across hospitals.

Missing Data

- To identify the extent and distribution of missing data, the developer compared data exported from the EHR to data manually abstracted from patients' medical charts for every patient included in the abstraction sample.
- For the initial population, the developer found only one data element missing for the Epic EHR and two data elements missing for the Cerner EHR.
- For the denominator exclusions, found zero data element missing for the Epic EHR and one data element missing for the Cerner EHR.
- The developer concluded that all measure's critical data elements are consistently stored in the EHR and can be accurately exported for calculation because the frequency of data missingness is zero for most test sites.

Comparability

- The measure only uses one set of specifications for this measure.

Questions for the Standing Committee regarding validity:

- *Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk adjustment approach, etc.)?*
- *Are the accuracy issues that are captured in the Feasibility Scorecard substantial enough to impact the validity of these data elements?*

Guidance From the Validity Algorithm

Potential threats to validity empirically assessed (box 1) Yes -> Empirical validity testing conducted (box 2) Yes -> validity testing conducted with computed performance measure scores (box 5) Yes -> Method described and appropriate for assessing conceptually and theoretically sound hypothesized relationships (box 6) Yes -> moderate certainty or confidence that the performance measure scores are a valid indicator of quality (box 7a) Yes -> Moderate

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

Criterion 3. Feasibility

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 states that the data elements for this measure are generated or collected by and used by healthcare personnel during the provision of care.

    - The developer also notes that all data elements are in defined fields in electronic health records.
- Using a simulated data set, the submission demonstrates that the evaluation of 100 percent of the measure logic can be automated.
- The Feasibility Scorecard assesses each data element across the following domains:
    - Availability – is the data element readily available in a structured format across electronic health record (EHR) systems?
    - Accuracy – is the information contained in the data correct?
    - Standards – is the data element coded using a nationally accepted terminology standard?
    - Workflow – is the data element routinely captured and used during care delivery?
- The developer has identified feasibility issues for the following data elements. For each data element, the developer was asked to provide additional context for the issue and a plan for addressing the issue: Physical Exam, Performed: Pressure Injury Deep Tissue and Physical Exam, Performed: Pressure Injury Stage 2, Stage 3, Stage 4 or Unstageable.
    - The developer notes that all hospitals used for feasibility assessment have the technical capability to record PI staging information within the EHR but 10% of the hospitals (two out of twenty) identified clinical documentation workflow inconsistencies which limit the ability to extract PI physical assessment documentation from the EHR in a structured format.
    - The developer states the following for both data elements: Workflow modifications would better enable capture (i.e., use the structured fields available through vendor system and already activated at the site to enter PI staging).The developer also notes that this would only require education, there would be no technical requirement.

Questions for the Standing 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?*
- *For data elements assessed to have feasibility issues, does the developer present a credible, near-term path to electronic collection?*

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

- Moderate – all identified feasibility issues have a core plan to address the issues and 100 percent coverage in simulated data unit tests (BONNIE)

Criterion 4: Use and Usability

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

4a. Use evaluates the extent to which audiences (e.g., consumers, purchasers, providers, and 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 they are not in use at the time of initial endorsement, then a credible plan for implementation within the specified time frames 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 developer stated that the measure is not currently used in an accountability program as the measure is under initial endorsement review, however, the measure was submitted to the 2022 Measures Under Consideration (MUC) list and will be reviewed by the Measure Applications Partnership (MAP) during the 2022-2023 review cycle. The developer states that CMS has sought MAP support for implementation in accountability programs such as Hospital Inpatient Quality Reporting and Promoting Interoperability Programs.

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; and (3) This feedback has been considered when changes are incorporated into the measure.

Feedback on the measure provided by those being measured or others

- The developer reports that feedback on the measure can be provided by users through the CMS Measures Management System posting, NQF endorsement review, Measures Application Partnership review, Proposed Rules published in the Federal Register, user community feedback through the QualityNet portal, and ongoing review by the Technical Advisory Panel.
- The developer states that implementation resources are provided through the CMS eCQI Resource Center and The ONC Project Tracking System for eCQMs included in CMS reporting programs.
- The developer notes that users are provided with a common place to transparently log, track, and discuss and clarify issues with eCQM implementation and logic interpretation.
- The developer states that as part of the measure rollout, CMS and The Joint Commission provide an annual webinar series for measured entities to review the measure specification, logic, and answer implementation questions.

***Questions for the Standing 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, and 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 states that because this is a new measure and no trend data is available.

4b2. Benefits versus harms. The 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 there were no unexpected findings found in the development or testing of this measure.

Potential harms

- The developer did not report any potential harms.

Additional Feedback:

- During the December 2022 discussion, MAP conditionally supported the measure for rulemaking pending endorsement by a consensus-based entity (CBE), with endorsement including a discussion of risk adjustment and stratification.
- A version of this measure was reviewed by the Measure Applications Partnership (MAP) for the Hospital IQR program and Medicare Promoting Interoperability Program for Eligible Hospitals (EHs) and Critical Access Hospitals (CAHs) (Medicare Promoting Interoperability Program) during the 2017-2018 pre-rulemaking cycle. The measure received a recommendation of "conditional support for rulemaking" pending NQF review and endorsement once the measure was fully tested. MAP supported the measure but had several concerns related to how the measure was specified.
    - MAP also cautioned about potential bias against facilities that do not have the expertise needed to accurately stage pressure injuries (e.g., certified wound care nurses).
    - MAP noted that risk adjustment may be necessary to ensure the measure does not disproportionately penalize facilities who may treat more complex patients (e.g., academic medical centers or safety net providers).
- This MUC is also submitted for the Medicare Promoting Interoperability Program.
    - MAP identified similar concerns to the ones identified during its 2017-2018 review. MAP noted the measure requires hospitals to have staff who are aware of pressure ulcers, are aware of their progression, and who have expertise staging ulcers.
    - MAP also noted that certain factors may impact a hospital's score on the measure, including the hospital's number of complex patients, whether it is a safety net hospital, and the proportion of patients with food insecurity. However, other MAP members cautioned against risk adjusting for those factors, as patients in those groups are the ones that need assessment the most.
    - The MAP Health Equity Advisory Group expressed no concerns regarding health equity and noted that the measure fills a quality gap.
- The MAP Rural Health Advisory Group cited concerns regarding data collection and the potential for rural providers to perform poorly on the measure relative to other providers due to staffing shortages.

***Questions for the Standing Committee**:*

- *How can the performance results be used to further the goal of high quality, efficient healthcare?*
- *Do the benefits of the measure outweigh any potential unintended consequences?*

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

Criterion 5: Related and Competing Measures

Related Measures

The developer did not note any NQF-endorsed related measures. However, they did note a non-NQF endorsed related measure, Patient Safety Indicator (PSI) 03: Pressure Ulcer Rate. The developer noted that PSI 03 is included in the NQF-endorsed composite measure, NQF #0531 PSI 90 Patient Safety and Adverse Events.

Harmonization

- The developer states that harmonization between PSI 03 and this measure are not necessary because there are differences between the measure and PSI 03, particularly the measure focus, target population, and the data sources used for each.
    - The developer further states that harmonization between this measure and NQF #0531 is not necessary because the only overlapping similarity between the measures is PSI 03, which the developer previously stated does not need harmonization as the outcome of the measure is different.
    - The developer continued noting that while the target populations for the measures are similar, the denominators are different.

Criteria 1: Importance to Measure and Report

1a. Evidence

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]

Current (2022) Submission

The incidence of pressure injuries in hospitalized patients has been estimated at 5.4 per 10,000 patient-days and the rate of hospital-acquired pressure injuries has been estimated at 8.4% (Li et al., 2020). Over 50% of reported pressure injuries in hospitals were Stage 2 or higher (Li et al., 2020). Over 50% of reported pressure injuries in hospitals were Stage 2 or higher (Li et al., 2020). Using the EHR data from 18 hospitals and in year 2020, we found that hospital-level measure performance rates ranged from 0.0% to 2.02% (for every 1,000 qualified hospital admissions there are 20 inpatient encounters where patients suffered Pressure Injury), with a system-wide, weighted average rate equal to 1.06%. Prior studies confirm that significant variation in rates of hospital acquired pressure injuries exists between hospitals (Rondinelli et al., 2018). Number of days to bed change has been significantly associated with an increase in pressure ulcer risk (OR, 2.89 [95% CI, 1.26-6.63]) and patients with a high nursing workload (i.e., patients who require more time from nurses at the bedside) have been found to reduce risk if staffing is adequate, i.e., 2:1 or 3:1 versus 5:1 or more (OR, 0.0916 [95% CI, 0.855-0.980]; p=0.011) (Bly et al., 2016; Cremasco et al., 2013).

Figure 1: PI Logic Model

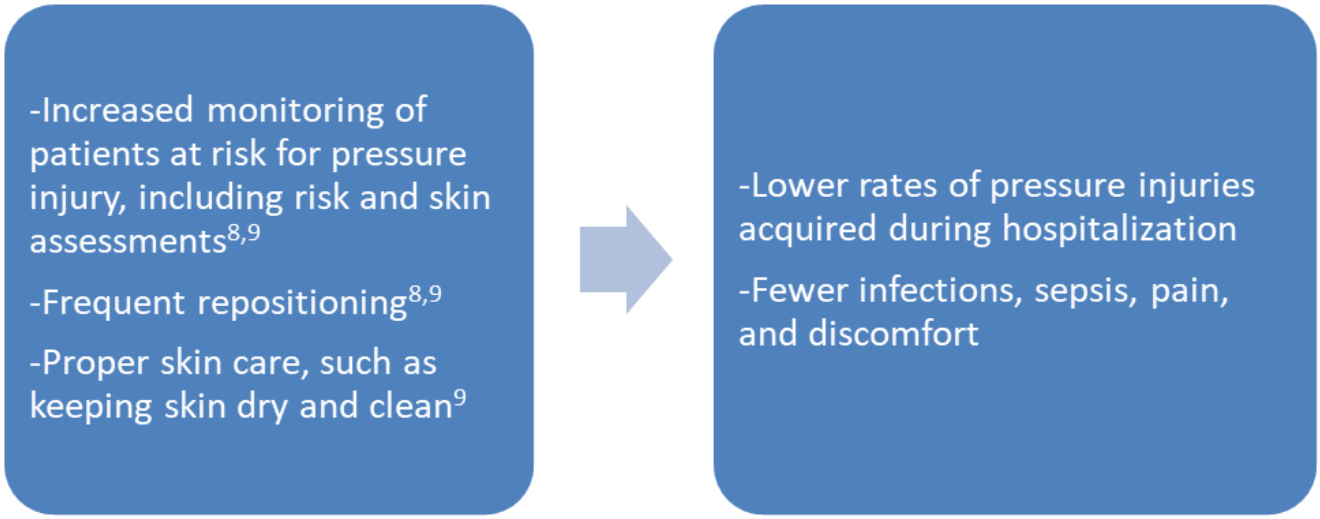
References:

1. Bly, D., Schallom, M., Sona, C., & Klinkenberg, D. (2016). A model of pressure, oxygenation, and perfusion risk factors for pressure ulcers in the intensive care unit. American Journal of Critical Care, 25(2), 156–154. https://doi.org/10.4037/ajcc2016840
2. Cremasco, M. F., Wenzel, F., Zanei, S. S. V., & Whitaker, I. Y. (2013). Pressure ulcers in the intensive care unit: The relationship between nursing workload, illness severity and pressure ulcer risk. Journal of Clinical Nursing, 22(15–16), 2183–2191. https://doi.org/10.1111/j.1365-2702.2012.04216.x

3. European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel and Pan Pacific Pressure Injury Alliance. Prevention and Treatment of Pressure Ulcers/Injuries: Clinical Practice Guideline. The International Guideline. Emily Haesler (Ed.). EPUAP/NPIAP/PPPIA: 2019. https://www.internationalguideline.com/
4. Li, Z., Lin, F., Thalib, L., & Chaboyer, W. (2020). Global prevalence and incidence of pressure injuries in hospitalised adult patients: A systematic review and meta-analysis. International Journal of Nursing Studies, Vol. 105. https://doi.org/10.1016/j.ijnurstu.2020.103546
5. Rondinelli, J., Zuniga, S., Kipnis, P., Kawar, L. N., Liu, V., & Escobar, G. J. (2018). Hospital-Acquired Pressure Injury: Risk-Adjusted Comparisons in an Integrated Healthcare Delivery System. Nurs Res, 67(1), 16–25. https://doi.org/10.1097/NNR.0000000000000258

Previous (2019) Submission

The goal of the Pressure Injury Electronic Clinical Quality Measure (eCQM) is to improve patient safety and prevent patients from acquiring a new pressure injury during their hospitalization. Pressure injuries, also called pressure ulcers, bed sores, or decubitus ulcers, are serious events and one of the most common patient harms. The injury can present as intact skin or an open ulcer, may be painful, and occurs from unrelieved pressure on the skin or in combination with shear force. Pressure injuries commonly lead to further patient harm, including local infection, osteomyelitis, anemia, and sepsis, in addition to causing significant depression, pain, and discomfort to patients.[1,2,3] The presence or development of a pressure injury can increase the length of a patient's hospital stay by an average of four days, which increases spending ranging from $20,900 to $151,700 per pressure injury.[4] Pressure injury is considered a serious reportable event by the National Quality Forum (NQF),[5] the CMS established non-payment for pressure injury,[6] and it is considered an indicator of the quality of nursing care a hospital provides.[7] It is well accepted that pressure injury can be reduced through best practices[8] such as frequent repositioning, proper skin care, and specialized cushions or beds.[9] The desired outcome for this eCQM is a reduction in rates of hospitalized patients who develop a new pressure injury. We define the harm as: a new stage 2, stage 3, stage 4 pressure injury, deep tissue pressure injury, or unstageable pressure injury during hospitalization.



Pressure Injury Logic Model

References:

1. Brem H MJ, Nierman D, et al. High Cost of Stage IV Pressure Ulcers. doi:10.1016/j.amjsurg.2009.12.021. American Journal of Surgery. 2010;200(4):473-477.
2. National Pressure Ulcer Advisory Panel. NPAUAP Pressure Injury Stages 2016; http://www.npuap.org/resources/educational-and-clinical-resources/npuap-pressure-injury-stages/.
3. Gunningberg L, Donaldson, N., Aydin, C., Idvall, E. Exploring variation in pressure ulcer prevalence in Sweden and the USA: Benchmarking in action. 18. 10.1111/j.1365-2753.2011.01702.x. Journal of evaluation in clinical practice. 2011: 904-910.
4. Bauer K, Rock K, Nazzal M, Jones O, Qu W. Pressure Ulcers in the United States' Inpatient Population From 2008 to 2012: Results of a Retrospective Nationwide Study. Ostomy Wound Manage. 2016;62(11):30-38.

5. National Quality Forum. List of SREs. 2016; http://www.qualityforum.org/Topics/SREs/List_of_SREs.aspx#sre4. Accessed January 13, 2017.
6. Centers for Medicare & Medicaid Services. Hospital-Acquired Conditions. 2015; https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HospitalAcqCond/Hospital-Acquired_Conditions.html. Accessed January 13, 2017.
7. National Quality Forum. National Voluntary Consensus Standards for Nursing-Sensitive Care: An Initial Performance Measure Set. Washington, D.C.: National Quality Forum; 2004.
8. Agency for Healthcare Research and Quality. Preventing Pressure Ulcers in Hospitals - A Toolkit for Improving Quality of Care. 2012; Rockville, MD. http://www.ahrq.gov/professionals/systems/hospital/pressureulcertoolkit/index.html
9. National Pressure Ulcer Advisory Panel, European Pressure Ulcer Advisory Panel, Pan Pacific Pressure Injury Alliance. Prevention and Treatment of Pressure Ulcers: Quick Reference Guide. 2014.

[Response Ends]

1a.02. Provide evidence that the target population values the measured outcome, process, or structure and finds it meaningful.

Describe how and from whom input was obtained.

[Response Begins]

Current (2022) Submission

The 2019 European Pressure Ulcer Advisory Panel/National Pressure Injury Advisory Panel/Pan Pacific Pressure Injury Alliance (EPUAP/NPIAP/PPPIA) guidelines were developed in collaboration with the engagement of patients, caregivers, informal caregivers, and other stakeholders. The survey responses regarding care goals, priorities, and education needs from 1,233 patients and families were incorporated into the guideline development process.

References:

1. European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel and Pan Pacific Pressure Injury Alliance. Prevention and Treatment of Pressure Ulcers/Injuries: Clinical Practice Guideline. The International Guideline. Emily Haesler (Ed.). EPUAP/NPIAP/PPPIA: 2019. https://www.internationalguideline.com/

[Response Ends]

1a.03. Provide empirical data demonstrating the relationship between the outcome (or PRO) and at least one healthcare structure, process, intervention, or service.

[Response Begins]

Current (2022) Submission

There are two recent evidence-based guidelines that outline prevention of pressure ulcers: The European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel and Pan Pacific Pressure Injury Alliance (Prevention and Treatment of Pressure Ulcers/Injuries: Clinical Practice Guideline. The International Guideline. EPUAP/NPIAP/PPPIA:2019) and The American College of Physicians (ACP) (Risk Assessment and Prevention of Pressure Ulcers: A Clinical Practice Guideline From the American College of Physicians. ACP: 2015). These guidelines outline evidence-based recommendations for prevention of pressure injuries through risk assessment, assessment of skin and tissue, preventive skin care, reducing progression through treatment of pressure injuries including nutrition, repositioning and early mobilization, static mattresses and overlays, early and accurate pressure injury classification and other evidenced based treatment modalities.

Selected guideline recommendations with the highest level of evidence from the EPUAP/NPIAP/PPPIA guidelines include:

Skin Status as a Risk Factor for Pressure Injuries

"1.2 Consider individuals with a Category/Stage I pressure injury to be at risk of developing a Category/Stage II or greater pressure injury" (Recommendation: strong positive; strength of evidence: A)

Twenty-four prognostic studies included factors associated with skin status in multivariable analysis of risk factors. Six prognostic studies provided evidence that Category/Stage I pressure injuries are a prognostic factor for Category/Stage II or greater pressure injuries and no studies found this factor to be non-significant. Evidence from two high quality Level 1 (Nixon et al., 2006; Smith et al., 2017) studies and one high quality (Reed et al., 2003) and three low quality (Allman et al., 1995; Demarre et al., 2015; Nixon, Cranny, & Bond, 2007) Level 3 studies supported the recommendation. Odds ratio of experiencing a Category/Stage II or greater pressure injury after experiencing a Category/Stage I pressure injury ranged from 1.95 to 7.02.

1. Allman, R., Goode, P., Patrick, M., Burst, N., & Bartolucci, A. (1995). Pressure ulcer risk factors among hospitalized patients with activity limitation. *J Am Med Assoc*, 273(11): 865-70.
2. Demarre, L., Verhaeghe, S., Van Hecke, A., Clays, E., Grypdonck, M., & Beeckman, D. (2015). Factors predicting the development of pressure ulcers in an at-risk population who receive standardized preventive care: Secondary analyses of a multicentre randomised controlled trial. *J Adv Nurs*, 71(2): 391-403.
3. Nixon, J., Cranny, G., & Bond, S. (2007). Skin alterations of intact skin and risk factors associated with pressure ulcer development in surgical patients: A cohort study. *Int J Nurs Stud*, 44(5): 655-63.
4. Nixon, J., Nelson, E.A., Cranny, G., Iglesias, C.P., Hawkins, K., Cullum, N.A., Phillips, A., Spilsbury, K., Torgerson, D.J., Mason, S, & PRESSURE Trial Group. (2006). Pressure relieving support surfaces: a randomised evaluation. *Health Technol Assess*, 10(22): iii-x, 1.
5. Reed, R., Hepburn, K., Adelson, R., Center, B., & McKnight, P. (2003). Low serum albumin levels, confusion, and fecal incontinence: Are these risk factors for pressure ulcers in mobility-impaired hospitalized adults? *Gerontology*, 49(4): 255-59.63.
6. Smith, I.L., Brown, S., McGinnis, E., Briggs, M., Coleman, S., Dealey, C., Muir, D., Nelson, E.A., Stevenson, R., Stubbs, N., Wilson, L., Brown, J.M., & Nixon, J. (2017) Exploring the role of pain as an early predictor of category 2 pressure ulcers: A prospective cohort study. *BMJ Open*, 7(1): e013623.62.

Conducting Skin and Tissue Assessment

"2.2 Inspect the skin of individuals at risk of pressure injuries to identify presence of erythema" (recommendation: strong positive; strength of evidence: A)

Ongoing skin assessment is necessary to detect early signs of pressure injury. Evidence from three Level 1 studies, one Level 2 study and a Level 3 study indicates that the presence of non-blanching erythema, a Category/Stage I pressure injury is predictive of development of a Category/Stage II or greater pressure injury. Across five prognostic studies, the risk of developing a more severe pressure injury was between three and five times' higher once a Category/Stage I pressure injury had been identified (odds ratio [OR] ranged from 3.1 to 7.98) (Level 1, 2 and 3 evidence). (Smith et al., 2017; Nixon, Cranny, & Bond, 2007; Reed et al., 2003; Allman et al., 1995; Demarre et al., 1995). Evidence from three Level 3 studies (Compton et al., 2008; Marchette et al., 1991; Schnelle et al., 1997) indicates that the presence of reddened skin other than blanchable erythema is associated with Stage/Category II pressure injury development. In a large (n = 698) prognostic study in acute care, critical care and non-surgical care, presence of erythema was associated with a more than two-fold increase in the risk of pressure injuries of Category/Stage II or greater. Identifying presence of erythema alerts health professionals to the need for further assessment and potential development of a pressure injury prevention and/or treatment plan. Identification of erythema is a component of a skin inspection.

1. Allman, R., Goode, P., Patrick, M., Burst, N., & Bartolucci, A. (1995). Pressure ulcer risk factors among hospitalized patients with activity limitation. *J Am Med Assoc*, 273(11): 865-70.
2. Compton, F., Hoffmann, F., Hortig, T., Strauss, M., Frey, J., Zidek, W., & Schafer, J.H. (2008). Pressure ulcer predictors in ICU patients: Nursing skin assessment versus objective parameters. *J Wound Care*, 17(10): 417.
3. Demarre, L., Verhaeghe, S., Van Hecke, A., Clays, E., Grypdonck, M., & Beeckman, D. (2015). Factors predicting the development of pressure ulcers in an at-risk population who receive standardized preventive care: Secondary analyses of a multicentre randomised controlled trial. *J Adv Nurs*, 71(2): 391-403.

4. Marchette, L., Arnell, I., & Redick, E. (1991). Skin ulcers of elderly surgical patients in critical care units. *Dimens Crit Care Nurs*, 10(6): 321-29.
5. Nixon, J., Cranny, G., & Bond, S. (2007). Skin alterations of intact skin and risk factors associated with pressure ulcer development in surgical patients: A cohort study. *Int J Nurs Stud*, 44(5): 655-63.
6. Reed, R., Hepburn, K., Adelson, R., Center, B., & McKnight, P. (2003). Low serum albumin levels, confusion, and fecal incontinence: Are these risk factors for pressure ulcers in mobility-impaired hospitalized adults? *Gerontology,* 49(4): 255-59.
7. Schnelle, J.F., Adamson G.M., Cruise, P.A., Al-Samarrai, N., Sarbaugh F.C., Uman, G., & Ouslander, J.G. (1997). Skin disorders and moisture in incontinent nursing home residents: Intervention implications. *J Am Geriatr Soc*, 45(10): 1182-1188.
8. Smith, I.L., Brown, S., McGinnis, E., Briggs, M., Coleman, S., Dealey, C., Muir, D., Nelson, E.A., Stevenson, R., Stubbs, N., Wilson, L., Brown, J.M., & Nixon, J. (2017). Exploring the role of pain as an early predictor of category 2 pressure ulcers: A prospective cohort study. *BMJ Open*, 7(1): e013623.

Quality Improvement Initiatives

"20.5 At an organizational level, develop an implement a structured, tailored and multi-faceted quality improvement program to reduce the incidence of pressure injuries" (recommendation: strong positive; strength of evidence: A)

Evidence from two high (Beeckman et al., 2013; Chaboyer et al., 2016) and two moderate (Tayyib, Coyer, & Lewis, 2015; Rantz et al., 2012) quality Level 1 studies indicated that a multi-faceted quality improvement program is associated with reductions in facility-acquired pressure injuries. This was supported by 17 Level 2 studies (Antonio & Conrad, 2013; Baldellu & Paciella, 2008; Sving et al., 2014; Tippet, 2009; McInerney, 2008; Bales & Duvendack, 2011; Bales & Padwojski, 2009; Boesch et al., 2012; Anderson et al., 2015;Crawford, Corbett, & Zuniga, 2014; Horn et al., 2010; Mallah, Nassar, & Kurdahi Badr, 2014; Milne et al., 2009; Rantz et al., 2009; Sebastian-Viana et al., 2016; Beinlich & Meehan, 2014; Fisher, Grosh, & Felty, 2016) of high, moderate and low quality; five Level 3 studies (Padula et al., 2016; Burston et al., 2015; Van Leen et al., 2014; Olsho et al., 2014; Stifter et al., 2015) of moderate and low quality and 11 Level 4 studies (Richardson et al., 2017; Smith et al., 2017; Asimus, Maclellan, & Li, 2011; Lewis et al., 2017; Hall & Ryan, 2015; Peterson et al., 2015; Thomas, 2008; Tzeng, Grandy, & Yin, 2013; Young et al., 2014; Baier et al., 2009; Baier et al., 2008) of high, moderate and low quality. The studies were conducted in a range of facilities including acute medical-surgical hospitals, critical/intensive care facilities, skilled nursing facilities, community care and pediatric hospitals. The studies were also delivered in a range of geographic locations including the US, Europe, the Middle East and the Pan-Pacific. The interventions in all studies included a range of initiatives that were tailored to the facility and often increased as the quality improvement program continued. Reported effectiveness varies and is likely contributed to by the baseline pressure injury incidence and factors discussed throughout this chapter.
Qualitative studies indicated that health professionals (Tayyib, Coyer, & Lewis, 2016; Roberts et al., 2016; Chaboyer & Gillespie, 2014) and individuals and their informal caregivers (Roberts et al., 2017) find quality improvement programs to be acceptable.

Chaboyer et al. (2016) evaluated a multi-faceted program in eight hospitals in Australia. The bundle included promoting patient engagement in pressure injury prevention, nurse education and promotional material. Although there was no statistically significant difference in pressure injury rates at the patient level compared to standard care (6.1% versus 10.5%, $p > 0.05$), there was a significant 52% reduction in hospital-acquired pressure injuries associated with the multi-faceted pressure injury bundle (Level 1). Beeckman et al.'s (2013) bundle included a wide range of components at the professional and organizational level aimed at reducing pressure injuries in nursing homes in Belgium (n = 11 facilities with n = 646 residents). Over the course of the four-month study, Category/Stage I and greater pressure injury rates decreased from 14.6% to 7.1%. Although nursing knowledge about pressure injuries did not change, the comprehensive bundle demonstrated a positive impact on the attitude of health professionals toward pressure injury prevention (Level 1). Also set in nursing homes, Rantz et al. (2012) found that introduction to US facilities of a comprehensive bundle that included education, clinical resources and mentoring was associated with a reduction in pressure injury incidence over a two year period (odds ratio [OR] 1.23, 95% confidence interval 1.00 to 1.51). Multi-faceted bundles have also been implemented with success in critical care settings. Tayyib et al. (2015) included evidence-based guidelines, education, risk

assessment and routine repositioning in a bundle for Saudi Arabian critical care units. There was a 70% lower likelihood of pressure injuries compared to prior to the bundle's introduction (7.14% versus 32.86%, $p < 0.001$)48 (Level 1).

1. Anderson, M., Finch Guthrie, P., Kraft, W., Reicks, P., Skay, C., & Beal, A.L. (2015). Universal Pressure Ulcer Prevention Bundle With WOC Nurse Support. *Journal of Wound, Ostomy & Continence Nursing*, 42(3): 217-225.
2. Antonio, T., & Conrad, K. (2013). Clinical and economic improvements in pressure injury care at Ballarat Health Services. *Wound Practice & Research*, 21(1): 4-10.
3. Asimus, M., Maclellan, L., & Li, P.I. (2011). Pressure ulcer prevention in Australia: the role of the nurse practitioner in changing practice and saving lives. *International Wound Journal*, 8(5): 508-513.
4. Baier, R.R., Butterfield, K., Harris, Y., & Gravenstein, S. (2008). Aiming for star performance: the relationship between setting targets and improved nursing home quality of care. *Journal of the American Medical Directors Association,* 9(8): 594-598.
5. Baier, R., Butterfield, K., Patry, G., Harris, Y., & Gravenstein, S. (2009). Identifying star performers: the relationship between ambitious targets and nursing home quality improvement. *Journal of the American Geriatrics Society*, 57(8): 1498-1503.
6. Baldelli, P., & Paciella, M. (2008). Creation and implementation of a pressure ulcer prevention bundle improves patient outcomes. *American Journal of Medical Quality*, 23(2): 136-142.
7. Bales, I., & Duvendack, T. (2011). Reaching for the moon: achieving zero pressure ulcer prevalence, an update. *Journal of Wound Care*, 20(8): 374-377.
8. Bales, I., & Padwojski, A. (2009). Reaching for the moon: achieving zero pressure ulcer prevalence. *Journal of Wound Care*, 18(4): 137-144.
9. Beeckman, D., Clays, E., Van Hecke, A., Vanderwee, K., Schoonhoven, L., & Verhaeghe, S. (2013). A multi-faceted tailored strategy to implement an electronic clinical decision support system for pressure ulcer prevention in nursing homes: A two-armed randomized controlled trial. *International Journal of Nursing Studies*, 50(4): 475-486.
10. Beinlich, N., & Meehan, A. (2014). Resource nurse program: A nurse-initiated, evidence-based program to eliminate hospital-acquired pressure ulcers. *Journal of Wound, Ostomy and Continence Nursing*, 41(2): 136-141.
11. Boesch, R.P., Myers, C., Garrett, T., Nie, A., Thomas, N., Chima, A., McPhail, G.L., Ednick, M., Rutter, M.J., & Dressman, K. (2012). Prevention of tracheostomy-related pressure ulcers in children. *Pediatrics*, 129(3): e792-e797.
12. Burston, S., Chaboyer, W., Gillespie, B., & Carroll, R. (2015). The effect of a transforming care initiative on patient outcomes in acute surgical units: a time series study. *J Adv Nurs*, 71(2): 417-29.
13. Chaboyer, W., Bucknall, T., Webster, J., McInnes, E., Gillespie, B.M., Banks, M., Whitty, J.A., Thalib, L., Roberts, S., Tallott, M., Cullum, N., & Wallis, M. (2016). The effect of a patient centred care bundle intervention on pressure ulcer incidence (INTACT): A cluster randomised trial. *International Journal of Nursing Studies*, 64: 63-71.
14. Chaboyer, W., & Gillespie, B.M. (2014). Understanding nurses' views on a pressure ulcer prevention care bundle: A first step towards successful implementation. *Journal of Clinical Nursing*, 23: 3415-3423.
15. Crawford, B., Corbett, N., & Zuniga, A. (2014). Reducing hospital-acquired pressure ulcers: A quality improvement project across 21 hospitals. *Journal of Nursing Care Quality*, 29(4): 303-310.
16. Fisher, K., Grosh, A., & Felty, V. (2016). Using nurse-led rounds to improve quality measures related to HAPUs. *Nursing,* 46(11): 63-68.
17. Hall, S., & Ryan, E. (2015). How a mattress selection matrix helped to sustain pressure ulcer prevention and also cut costs. *Wounds UK,* 11(3): 16-21.
18. Horn, S.D., Sharkey, S.S., Hudak, S., Gassaway, J., James, R., & Spector, W. (2010). Pressure ulcer prevention in long-term-care facilities: a pilot study implementing standardized nurse aide documentation and feedback reports. *Advances in Skin & Wou nd Care*, 23(3): 120-131.
19. Lewis, H., Hughes, D., Madell, D., Coomarasamy, C., Villa, L., & Hayward, B. (2017). Estimated reduction in expenditure on hospital-acquired pressure injuries after an intervention for early identification and treatment. *New Zealand Medical Journal*, 130(1461): 42-46.
20. Mallah, Z., Nassar, N., & Kurdahi Badr, L. (2014). The Effectiveness of a Pressure Ulcer Intervention Program on the Prevalence of Hospital Acquired Pressure Ulcers: Controlled Before and After Study. *Applied Nursing Research*.
21. Mathiesen, A.S.M., Norgaard, K., Andersen, M.F.B., Moller, K.M., & Ehlers, L.H. (2013). Are labour-intensive efforts to prevent pressure ulcers cost-effective? *Journal of Medical Economics*, 16(10): 1238-1245.

22. McInerney, J.A. (2008). Reducing hospital-acquired pressure ulcer prevalence through a focused prevention program. *Advances in Skin & Wound Care*, 21(2): 75-78.36.
23. Milne, C.T., Trigilia, D., Houle, T.L., Delong, S., & Rosenblum, D. (2009). Reducing pressure ulcer prevalence rates in the long-term acute care setting. *Ostomy Wound Management*, 55(4): 50-59.
24. Olsho, L.E.W., Spector, W.D., Williams, C.S., Rhodes, W., Fink, R.V., Limcangco, R., & Hurd, D. (2014). Evaluation of AHRQ's on-time pressure ulcer prevention program: A facilitator-assisted clinical decision support intervention for nursing homes. *Medical Care*, 52(3): 258-266.
25. Padula, W.V., Gibbons, R.D., Valuck, R.J., Makic, M.B.F., Mishra, M.K., Pronovost, P.J., & Meltzer, D.O. (2016). Are evidence-based practices associated with effective prevention of hospital-acquired pressure ulcers in US Academic Medical Centers? *Medical Care*, 54(5): 512-518.
26. Peterson, J., Adlard, K., Walti, B.I., Hayakawa, J., McClean, E., & Feidner, S.C. (2015). Clinical Nurse Specialist Collaboration to Recognize, Prevent, and Treat Pediatric Pressure Ulcers. *Clinical Nurse Specialist*, 29(5): 276-82.
27. Rantz, M.J., Cheshire, D., Flesner, M., Petroski, G.F., Hicks, L., Alexander, G., Aud, M.A., Siem, C., Nguyen, K., Boland, C., & Thomas, S. (2009). Helping nursing homes "at risk" for quality problems: a statewide evaluation. *Geriatric Nursing*, 30(4): 238-249.
28. Rantz, M.J., Zwygart-Stauffacher, M., Hicks, L., Mehr, D., Flesner, M., Petroski, G.F., Madsen, R.W., & Scott-Cawiezell, J. (2012). Randomized multilevel intervention to improve outcomes of residents in nursing homes in need of improvement. *Journal of the American Medical Directors Association*, 13(1): 60-68.
29. Richardson, A., Peart, J., Wright, S.E., & McCullagh, I.J. (2017). Reducing the incidence of pressure ulcers in critical care units: A 4-year quality improvement. *International Journal for Quality in Health Care*, 29(3): 433-439.
30. Roberts, S., McInnes, E., Wallis, M., Bucknall, T., Banks, M., & Chaboyer, W. (2016). Nurses' perceptions of a pressure ulcer prevention care bundle: a qualitative descriptive study. *BMC Nursing*, 15: 64.
31. Roberts, S., Wallis, M., McInnes, E., Bucknall, T., Banks, M., Ball, L., & Chaboyer, W. (2017). Patients' Perceptions of a Pressure Ulcer Prevention Care Bundle in Hospital: A Qualitative Descriptive Study to Guide Evidence-Based Practice. *Worldviews Evid Based Nurs*.
32. Sebastian-Viana, T., Losa-Iglesias, M., Gonzalez-Ruiz, J.M., Lema-Lorenzo, I., Nunez-Crespo, F.J., Salvadores Fuentesm P., & team A. (2016). Reduction in the incidence of pressure ulcers upon implementation of a reminder system for health-care providers. *Applied Nursing Research*, 29: 107-12.
33. Smith, S.K., Ashby, S.E., Thomas, L., & Williams, F. (2017). Evaluation of a multifactorial approach to reduce the prevalence of pressure injuries in regional Australian acute inpatient care settings. *International Wound Journal*, 07: 07.
34. Stifter, J., Yao, Y., Lodhi, M.K., Lopez, K.D., Khokhar, A., Wilkie, D.J., & Keenan, G.M. (2015). Nurse Continuity and Hospital-Acquired Pressure Ulcers: A Comparative Analysis Using an Electronic Health Record "Big Data" Set. *Nursing research*, 64(5): 361-371.
35. Sving, E., Högman, M., Mamhidir, A.G., & Gunningberg, L. (2014). Getting evidence-ba sed pressure ulcer prevention into practice: A multi-faceted unit-tailored intervention in a hospital setting. *International Wound Journal*.
36. Tayyib, N., Coyer, F., & Lewis, P.A. (2015). A two-arm cluster randomized control trial to determine the effectiveness of a pressure ulcer prevention bundle for critically ill patients. *Journal of Nursing Scholarship*, 47(3): 237-47.
37. Tayyib, N., Coyer, F., & Lewis, P. (2016). Pressure injury prevention in a Saudi Arabian intensive care unit: Registered nurse attitudes toward prevention strategies and perceived facilitators and barriers to evidence implementation. *Journal of Wound, Ostomy, & Continence Nursing*, 43(4): 369-74.
38. Thomas, M.E. (2008). The providers' coordination of care: a model for collaboration across the continuum of care. *Profes sional Case Management,* 13(4): 220-227.
39. Tippet, A.W. (2009). Reducing the incidence of pressure ulcers in nursing home residents: a prospective 6-year evaluation. *Ostomy Wound Management*, 55(11): 52-58.
40. Tzeng, H-M., Grandy, G.A., & Yin, C-Y. (2013). Staff response time to call lights and unit-acquired pressure ulcer rates in adult in-patient acute care units. *Contemporary Nurse*, 45(2): 182-187.
41. Van Leen, M.W.F., Schols, J.M.G.A., Hovius, S.E.R., & Halfens, R.J.G. (2014). The effect of a simple 3-step pressure relieving strategy for preventing pressure ulcers: An explorative longitudinal study from 2002-2011. *Wounds*, 26(10): 285-292.

42. Whitty, J.A., McInnes, E., Bucknall, T., Webster, J., Gillespie, B.M., Banks, M., Thalib, L., Wallis, M., Cumsille, J., Roberts, S., & Chaboyer, W. (2017). The cost-effectiveness of a patient centred pressure ulcer prevention care bundle: Findings from the INTACT cluster randomised trial. *Int J Nurs Stud*, 75: 35-42.
43. Young, D.L., Borris-Hale, C., Falconio-West, M., Chakravarthy, D. (2014). A Single Long-Term Acute Care Hospital Experience with a Pressure Ulcer Prevention Program. *Rehabil Nurs*.

ACP guidelines 2015

Moderate-quality evidence showed that the use of advanced static mattresses or overlays was associated with a lower risk for pressure ulcers compared with standard hospital mattresses, and no brand was shown to be superior. Advanced static mattresses and overlays are also less expensive than alternating-air or low– air-loss mattresses and can be used as part of a multicomponent approach to pressure ulcer prevention.

- 2. ACP recommends that clinicians should choose advanced static mattresses or advanced static overlays in patients who are at an increased risk of developing pressure ulcers (recommendation: strong recommendation; quality of evidence: moderate-quality evidence)

It is widely accepted that the risk of developing a pressure injury can be reduced through best practices. Systematically measuring patients who develop new pressure injuries while in the hospital setting will provide hospitals with a reliable and timely measurement, to more reliably assess harm reduction efforts and modify their improvement efforts in near real-time. This eCQM will fill a gap in measurement and provide incentives for hospitals' quality improvement. Although several pressure injury measures are currently in use, there are no electronic health record (EHR)-based measures intended for use in acute care hospitals. In addition, the intent of this measure is to incentivize greater achievements in reducing harms and enhance hospital performance on patient safety outcomes.

1. Bly, D., Schallom, M., Sona, C., & Klinkenberg, D. (2016). A model of pressure, oxygenation, and perfusion risk factors for pressure ulcers in the intensive care unit. American Journal of Critical Care, 25(2), 156–154. https://doi.org/10.4037/ajcc2016840
2. Cremasco, M. F., Wenzel, F., Zanei, S. S. V., & Whitaker, I. Y. (2013). Pressure ulcers in the intensive care unit: The relationship between nursing workload, illness severity and pressure ulcer risk. Journal of Clinical Nursing, 22(15–16), 2183–2191. https://doi.org/10.1111/j.1365-2702.2012.04216.x
3. European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel and Pan Pacific Pressure Injury Alliance. Prevention and Treatment of Pressure Ulcers/Injuries: Clinical Practice Guideline. The International Guideline. Emily Haesler (Ed.). EPUAP/NPIAP/PPPIA: 2019. https://www.internationalguideline.com/
4. Qaseem, A., Mir, T. P., Starkey, M., Denberg, T. D., & Clinical Guidelines Committee of the American College of Physicians (2015). Risk assessment and prevention of pressure ulcers: a clinical practice guideline from the American College of Physicians. Annals of internal medicine, 162(5), 359–369. https://doi.org/10.7326/M14-1567
5. Tayyib, N., Coyer, F., & Lewis, P. (2016). Saudi Arabian adult intensive care unit pressure ulcer incidence and risk factors: A prospective cohort study. International Wound Journal, 13(5), 912–919. https://doi.org/10.1111/iwj.12406

Previous (2019) Submission

The Agency for Healthcare Research and Quality (AHRQ) identified hospital-acquired pressure injuries as a harm to patients that could be prevented, began collecting and reporting incident rates to measure the extent of the problem, and provided toolkits to providers around how to lower their rates.[1] It is widely accepted that the risk of developing a pressure injury can be reduced by best practices such as frequent repositioning, proper skin care, and specialized cushions or beds;[2,3] studies have also begun to assess the impact of nutritional interventions.[4] AHRQ published data that showed 3.1 million fewer incidents of hospital-acquired harm in 2011-2015 compared with 2010; 23% of this reduction was from a reduction in-hospital acquired pressure injuries.[1] A 3-year, intervention study found that implementation of a novel 7-step care-based process, acquisition of specialized equipment, and educational initiatives were associated with a significant decrease in incidence rate of pressure injuries.[5] A second study also showed a link between a hospital's

processes of care and the outcome of hospital-acquired pressure injury. Processes of care analyzed included risk/skin assessment, risk status at admission, and pressure injury prevention strategies (such as pressure relief). [3]

Early identification and effective facility-level prevention strategies are essential in health care systems for patients at risk for pressure injuries.[6] Further, studies suggest that variation in care delivered negatively impacts pressure injury rates.[7,8] Although the National Pressure Ulcer Advisory Panel (NPUAP) Board of Directors revised the pressure injury staging system in 2015, inaccurate staging of pressure injuries persists impacting the hospital care delivered to patients and influencing their pressure injury rates.[9]

References:

1. Agency for Healthcare Research and Quality. National Scorecard on Rates of Hospital-Acquired Conditions 2010 to 2015: Interim Data from National Efforts to Make Health Care Safer. 2016; https://www.ahrq.gov/professionals/quality-patient-safety/pfp/2015-interim.html?utm_source=AHRQ&utm_medium=PSLS&utm_term=&utm_content=14&utm_campaign=AHRQ_NSOHAC_2016. Accessed January 13, 2017.
2. Agency for Healthcare Research and Quality. Preventing Pressure Ulcers in Hospitals - A Toolkit for Improving Quality of Care. 2012; Rockville, MD. http://www.ahrq.gov/professionals/systems/hospital/pressureulcertoolkit/index.html
3. Gunningberg L, Donaldson, N., Aydin, C., Idvall, E. Exploring variation in pressure ulcer prevalence in Sweden and the USA: Benchmarking in action. 18. 10.1111/j.1365-2753.2011.01702.x. Journal of evaluation in clinical practice. 2011: 904-910.
4. Lozano-Montoya I, Vélez-Díaz-Pallarés M, Abraha I, et al. Nonpharmacologic Interventions to Prevent Pressure Ulcers in Older Patients: An Overview of Systematic Reviews (The Software Engine for the Assessment and optimization of drug and non-drug Therapy in Older persons [SENATOR] Definition of Optimal Evidence-Based Non-drug Therapies in Older People [ONTOP] Series). *J Am Med Dir Assoc.* 2016;17(4):370.e371-370.e310.
5. Lam C, Elkbuli A, Benson B, et al. Implementing a Novel Guideline to Prevent Hospital-Acquired Pressure Ulcers in a Trauma Population: A Patient-Safety Approach. J Am Coll Surg. 2018;226(6):1122-1127.
6. Lyder CH, Ayello EA. Pressure Ulcers: A Patient Safety Issue. In: Hughes RG, editor. Patient Safety and Quality: An Evidence-Based Handbook for Nurses. Rockville (MD): Agency for Healthcare Research and Quality (US); 2008 Apr. Chapter 12. Available from: https://www.ncbi.nlm.nih.gov/books/NBK2650/
7. Edsberg LE, Black JM, Goldberg M, McNichol L, Moore L, Sieggreen M. Revised National Pressure Ulcer Advisory Panel Pressure Injury Staging System: Revised Pressure Injury Staging System. *J Wound Ostomy Continence Nurs.* 2016;43(6):585-597.
8. Horn S, Buerhaus P, Bergstrom N, J. Smout R. RN Staffing Time and Outcomes of Long-Stay Nursing Home Residents. Am J Nurs. 2005; 105:58–70.
9. National Pressure Ulcer Advisory Panel and European Pressure Ulcer Advisory Panel. Prevention and Treatment of Pressure Ulcers: Clinical Practice Guideline. Washington, DC: National Pressure Ulcer Advisory Panel. 2014.

[Response Ends]

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

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]

Current (2022) Submission:

This safety eCQM captures the number of patients who experience harm in the form of a pressure injury, during their inpatient hospitalization. The incidence of pressure injuries in hospitalized patients has been estimated at 5.4 per 10,000 patient-days and the rate of hospital-acquired pressure injuries has been estimated at 8.4% (Li et al., 2020). While studies have identified a reduction in the incidence of pressure injuries from 1990 to 2017 (Siotos et al., 2022), other studies have

found that pressure injuries are consistently underreported, with lower-stage pressure ulcers the least likely to be reported (Chen et al., 2022). Over 50% of reported pressure injuries in hospitals were Stage 2 or higher (Li et al., 2020). Hospital-acquired pressure injuries are serious events and one of the most common patient harms. Pressure injuries commonly cause local infection, osteomyelitis, anemia, and sepsis (Brem et al., 2010), in addition to causing significant depression, pain, and discomfort to patients (Gunningberg et al., 2011). Hospital-acquired pressure injuries are associated with 1.5 to 2.0 times greater risk of 30, 60, and 90-day readmissions (Wassel et al, 2020). Any Stage 3, Stage 4, or unstageable pressure ulcer acquired after admission/presentation to a healthcare setting is considered a serious reportable event by the National Quality Forum (NQF) (National Quality Forum, 2011).

Systematically assessing patients who develop new pressure injuries while in the hospital setting will provide hospitals with a reliable and timely measurement, to more reliably assess harm reduction efforts and modify their improvement efforts in near realtime. The intent of this measure is to incentivize greater achievements in reducing harms and enhance hospital performance on patient safety outcomes. Stage 2 pressure injuries constitute a very real patient harm that should be monitored and addressed; however, the relative level of harm is less than with Stage 3, Stage 4, Unstageable pressure injuries and potentially DTI. (National Pressure Ulcer Advisory Panel, 2019). The revised measure specification allows a 24-hour time window for accurate and timely identification of stage 2, 3, 4, or unstageable pressure injury present on admission. The revised measure specification allows a 72-hour time window for accurate and timely identification of deep tissue injury (DTI) because early diagnosis of DTI allows prompt identification of possible causes, initiation of treatment, and implementation of preventive strategies. Up to 72 hours can lapse between the precipitating pressure event and the onset of purple or maroon skin, so a longer time window is needed to exclude cases when the precipitating event occurred before the patient's admission. (Wound Management and Prevention, 2018).

References:

1. Bly, D., Schallom, M., Sona, C., & Klinkenberg, D. (2016). A model of pressure, oxygenation, and perfusion risk factors for pressure ulcers in the intensive care unit. American Journal of Critical Care, 25(2), 156–154. https://doi.org/10.4037/ajcc2016840
2. Brem H, M. J., Nierman D, et al. (2010). High Cost of Stage IV Pressure Ulcers. American Journal of Surgery, 200(4), 473-477
3. National Quality Forum (2011). List of Serious Reportable Events (aka SRE or "Never Events"). https://www.qualityforum.org/Topics/SREs/List_of_SREs.aspx
4. Chen, Z., Gleason, L.J., & Sanghavi, P. (2022). Accuracy of pressure ulcer events in US nursing home ratings. Medical Care, 60(10), 775-783. https://doi.org/10.1097/MLR.0000000000001763
5. Cremasco, M. F., Wenzel, F., Zanei, S. S. V., & Whitaker, I. Y. (2013). Pressure ulcers in the intensive care unit: The relationship between nursing workload, illness severity and pressure ulcer risk. Journal of Clinical Nursing, 22(15–16), 2183–2191. https://doi.org/10.1111/j.1365-2702.2012.04216.x
6. Gunningberg, L., Donaldson, N., Aydin, C., Idvall, E. (2011). Exploring variation in pressure ulcer prevalence in Sweden and the USA: Benchmarking in action. 18. Journal of evaluation in clinical practice., 904 -910
7. Li, Z., Lin, F., Thalib, L., & Chaboyer, W. (2020). Global prevalence and incidence of pressure injuries in hospitalised adult patients: A systematic review and meta-analysis. International Journal of Nursing Studies, Vol. 105. https://doi.org/10.1016/j.ijnurstu.2020.103546
8. National Pressure Ulcer Advisory Panel (NPUAP), Submission of Open Comment August 23, 2019, retrieved May 12, 2020 from: https://cdn.ymaws.com/npiap.com/resource/resmgr/npuap_nqf_response_final_8.2.pdf
9. Rondinelli, J., Zuniga, S., Kipnis, P., Kawar, L. N., Liu, V., & Escobar, G. J. (2018). No TitlHospital-Acquired Pressure Injury: Risk-Adjusted Comparisons in an Integrated Healthcare Delivery System. Nurs Res, 67(1), 16–25.
10. Siotos, C., Bonett, A.M., et al. (2022) Burden of pressure injuries; Findings from the Global Burden of Disease Study. Eplasty, 22, e19.
11. Tayyib, N., Coyer, F., & Lewis, P. (2016). Saudi Arabian adult intensive care unit pressure ulcer incidence and risk factors: A prospective cohort study. International Wound Journal, 13(5), 912–919. https://doi.org/10.1111/iwj.12406
12. Wassel, C.L., Delhougne, G., Gayle, J.A., Dreyfus, J., & Larson, B. (2020) Risk of readmissions, mortality, and hospital-acquired conditions across hospital-acquired pressure injury (HAPI) stages in a US National Hospital Discharge database. Int Wound J., 17, 1924-1934. https://doi.org/10.1111/iwj.13482
13. Wound Management & Prevention: Volume 64 - Issue 11 - November 2018 ISSN 1943-2720 Index: Ostomy Wound Manage. 2018;64(11):30-41' DefinitionInpatient hospitalizations: Includes time in the eme

Previous (2019) Submission:

This safety eCQM captures the number of patients who experience harm in the form of a pressure injury, during their inpatient hospitalization. Hospital-acquired pressure injuries are serious events and one of the most common patient harms. Pressure injuries commonly cause local infection, osteomyelitis, anemia, and sepsis (Brem, et al., 2010), in addition to causing significant depression, pain, and discomfort to patients (Gunningberg et al., 2011). Pressure injury is considered a serious reportable event by the National Quality Forum (NQF) (Centers for Medicare and Medicaid Services, 2015). CMS also established non-payment for pressure injury (National Quality Forum, 2016), and the rate of pressure injuries is considered an indicator of the quality of nursing care a hospital provides (National Quality Forum, 2005).

It is widely accepted that the risk of developing a pressure injury can be reduced through best practices such as frequent repositioning, proper skin care, and specialized cushions or beds (Berlowitz, et al., 2012). Systematically measuring patients who develop new pressure injuries while in the hospital setting will provide hospitals with a reliable and timely measurement, to more reliably assess harm reduction efforts and modify their improvement efforts in near real-time. This eCQM will fill a gap in measurement and provide incentives for hospitals' quality improvement. Although several pressure injury measures are currently in use, there are no electronic health record (EHR)-based measures intended for use in acute care hospitals. In addition, the intent of this measure is to incentivize greater achievements in reducing harms and enhance hospital performance on patient safety outcomes.

References:

Brem H, M. J., Nierman D, et al. (2010). High Cost of Stage IV Pressure Ulcers. doi:10.1016/j.amjsurg.2009.12.021. American Journal of Surgery, 200(4), 473-477.

Gunningberg, L., Donaldson, N., Aydin, C., Idvall, E. (2011). Exploring variation in pressure ulcer prevalence in Sweden and the USA: Benchmarking in action. 18. 10.1111/j.1365-2753.2011.01702.x. Journal of evaluation in clinical practice., 904-910.

Centers for Medicare & Medicaid Services. (2015). Hospital-Acquired Conditions. Retrieved January 13, 2017, from https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/HospitalAcqCond/Hospital-Acquired_Conditions.html

National Quality Forum. (2016). List of SREs. Retrieved January 13, 2017, from http://www.qualityforum.org/Topics/SREs/List_of_SREs.aspx#sre4

National Quality Forum. (2005). National Voluntary Consensus Standards for Nursing-Sensitive Care: An Initial Performance Measure Set. Retrieved January 13, 2017, from http://www.qualityforum.org/Publications/2004/10/National_Voluntary_Consensus_Standards_for_Nursing-Sensitive_Care__An_Initial_Performance_Measure_Set.aspx.

Berlowitz, D. VanDeusen Lukas, C.; Parker, V.; Niederhauser, A.;, & Silver, J. L., C.; Ayello, E.; Zulkowski, K. (2012). Preventing Pressure Ulcers in Hospitals- A Toolkit for Improving Quality of Care.

[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]

Current (2022) Submission:

A total of 18 hospitals (test sites) with varying bed size, geographic location, teaching status, urbanicity, and EHR systems participated in measure testing. Using data from test sites' EHR systems over the full calendar year 2020, hospitals' performance rate in PI ranged from a low (min) of 0 to a high (max) of 2.02 per 100 qualified inpatient admissions. The

system-wide, weighted average measure rate equaled 1.06 per 100 qualified inpatient admissions. The standard deviation of measure performance rate across test sites was 0.56 per 100 qualified inpatient admissions. The interquartile range was 0.63 per 100 qualified inpatient admissions.

- Testing data came from test sites' EHR systems. Testing data in full calendar year 2020 (Jan 1, 2020 to Dec 31, 2020) were used. No partial year data were used.
- A total of 18 hospitals participated in measure testing.
- The number of **unique** patients included in measure denominator ranged from a low of 470 to a high of 30,650 across test sites.
- Measure denominator encounters ranged from a low of 553 to a high of 38,476 across test sites.

Table 24 below provides the high-level information on the measure testing sites and their performance rate in PI based on data from calendar year 2020.

Table 24. High-level Characteristics of Test Sites and Measure Performance Rate (Score) in CY2020

Hospital	Teaching Status	Urban/Rural	Bed Size	No. of Unique Patients	Denominator Count	Observed Measure Rate
1	Academic	Urban	>499	30,650	38,476	1.92%
2	Non-academic	Rural	25-99	2,996	3,708	1.02%
3	Academic	Urban	100-199	6,503	7,821	1.12%
4	Non-academic	Urban	25-99	1,087	1,202	1.05%
5	Non-academic	Urban	200-499	10,308	12,540	1.47%
6	Academic	Urban	200-499	10,878	14,576	0.76%
7	Academic	Urban	200-499	11,318	14,533	1.46%
8	Non-academic	Urban	100-199	4,196	5,782	3.31%
9	Non-academic	Urban	100-199	4,902	6,143	1.29%
10	Non-academic	Urban	200-499	5,468	6,587	1.01%
11	Non-academic	Rural	25-99	923	1,146	1.37%
12	Non-academic	Rural	25-99	1,289	1,623	1.05%
13	Non-academic	Rural	100-199	1,767	2,045	1.01%
14	Non-academic	Rural	25-99	1,002	1,252	1.72%
15	Non-academic	Rural	100-199	2,542	3,101	1.60%
16	Non-academic	Urban	200-499	4,518	5,623	1.36%
17	Non-academic	Rural	25-99	1,542	1,612	0.91%
18	Non-academic	Rural	25-99	470	553	1.69%

Notes: A total of 18 hospitals with two different EHR systems (Epic: Hospital 1 and Cerner: Hospitals 2-18) participated in measure testing. Data from test sites' EHR systems in CY2020 were used.

Previous (2019) Submission:

This eCQM was tested with 3 test sites (24 hospitals) in 3 states (located in Midwest, West, and Northeast). Hospitals varied in size (200+ beds, 15-500 beds, and 450-700 beds), EHR systems (Meditech, Cerner, Epic), teaching status (teaching and non-teaching hospitals), and location (urban, suburban, and rural). A detailed breakdown of the

characteristics of the measured facilities and the patient population can be found in the attached Measure Testing Form (Beta Datasets 1, 2, and 3).

The measure performance, including the denominator, numerator, and measure rate by hospital, follows.

Hospital Test Site 1 (Beta Dataset 1 per Testing Form)

- Number of Hospitals: 1
- Data collection period: 1/1/2017 - 12/31/2017
- Denominator: 7,573
- Numerator: 38
- Performance rate: 0.50%
- 95% confidence interval: 0.36%, 0.69%
- Standard Deviation: N/A (only one hospital)

Hospital Test Site 2 (Beta Dataset 2 per Testing Form)

- Number of Hospitals: 21
- Data collection period: 1/1/2017 - 12/31/2017
- Denominator: 100,238
- Numerator: 724
- Performance rate: 0.72%
- 95% confidence interval: 0.67%, 0.78%
- Standard Deviation: 0.47%

Hospital Test 3 (Beta Dataset 3 per Testing Form)

- Number of Hospitals: 2
- Data collection period: 1/1/2017 - 12/31/2017
- Denominator: 56, 330
- Numerator: 414
- Performance rate: 0.73%
- 95% confidence interval: 0.67%, 0.81%
- Standard Deviation: 0.06%

Overall Performance

- Number of Hospitals: 24
- Performance rate: 0.72%
- 95% confidence interval: 0.68%, 0.76%
- Standard deviation: 0.45%
- Range: 0.0% to 1.46%

[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]

Current (2022) Submission:

While performance data is reported in 1b.02, we highlight that prior studies confirm that significant variation in rates of hospital acquired pressure injuries exists between hospitals within multi-hospital systems (Rondinelli et al., 2018) and across research sites in North America (Li et al., 2020).

References:

1. Li, Z., Lin, F., Thalib, L., & Chaboyer, W. (2020). Global prevalence and incidence of pressure injuries in hospitalised adult patients: A systematic review and meta-analysis. International Journal of Nursing Studies, Vol. 105. https://doi.org/10.1016/j.ijnurstu.2020.103546

Rondinelli, J., Zuniga, S., Kipnis, P., Kawar, L. N., Liu, V., & Escobar, G. J. (2018). Hospital-Acquired Pressure Injury: Risk-Adjusted Comparisons in an Integrated Healthcare Delivery System. Nurs Res, 67(1), 16–25. https://doi.org/10.1097/NNR.0000000000000258

Previous (2019) Submission:

N/A

[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]

Current (2022) Submission:

A total of 18 hospitals (test sites) with varying bed size, geographic location, teaching status, urbanicity, and EHR systems participated in measure testing. Using data from test sites' EHR systems over the full calendar year 2020, hospitals' performance rate in PI ranged from a low (min) of 0 to a high (max) of 2.02 per 100 qualified inpatient admissions. The system-wide, weighted average measure rate equaled 1.06 per 100 qualified inpatient admissions. The standard deviation of measure performance rate across test sites was 0.56 per 100 qualified inpatient admissions. **Table 25** below provides information on measure denominator population, stratified by sex, age bins, race/ethnicity, and primary source of payment.

Table 25. Measure Denominator Population Characteristics

Measure Denominator Population Characteristics	EHR System: Epic	EHR System: Epic	EHR System: Cerner	EHR System: Cerner
*	n	%	n	%
Number of denominator inpatient encounters	38,476	100%	89,847	100%
Number of unique patients	30,650	100%	70,883	100%
Age Mean (Std.Dev)	50.9 (19.2)		56.1 (21.1)	
Age bins	*	*	*	*
18-35	10,477	34%	19,500	28%

Measure Denominator Population Characteristics	EHR System: Epic	EHR System: Epic	EHR System: Cerner	EHR System: Cerner
36-64	11,888	39%	24,736	35%
65+	8,291	27%	26,718	38%
Sex	*	*	*	*
Male	9,562	31%	28,095	40%
Female	21,088	69%	42,788	60%
Race	*	*	*	*
White	18,428	60%	53,044	75%
Black or African American	5,681	19%	2,795	4%
Other	4,966	16%	13,363	19%
Unknown	1,575	5%	1,681	2%
Ethnicity	*	*	*	*
Hispanic or Latino	4,008	13%	24,717	35%
Non-Hispanic	24,848	81%	44,735	63%
Unknown	1,794	6%	1,431	2%
(Primary) Payer	*	*	*	*
Medicare	6,730	22%	18,446	26%
Medicaid	4,315	14%	24,455	35%
Private Insurance	19,192	63%	24,247	34%
Self-pay or Uninsured	528	2%	1,337	2%
Other	56	0%	3,458	5%

Notes: std.dev = standard deviation. Of the 18 test sites, one uses Epic and 17 use Cerner as their EHR system. *Cells intentionally left blank.

Across all test sites and within the measure denominator population, male patients had higher chance of experiencing hospital acquired (HA) PI than female patients and patients aged 65 or above were more likely to experience HA-PI than those 64 or younger. Non-Hispanic African Americans had a moderately higher chance of developing HA-PI and Medicare beneficiaries were more likely than Medicaid beneficiaries or commercially insured patients to experience PI during hospitalization. **Table 26** below provides information on measure performance rate, stratified by sex, age bins, race/ethnicity, and primary source of payment.

Table 26. Measure Performance Rate (Score) - Overall and Stratified

Measure denominator population	EHR System: Epic	EHR System: Epic	EHR System: Cerner	EHR System: Cerner
Rate of PI per 100 denominator encounters	**Mean**	**Std.Err**	**Mean**	**Std.Err**
Overall	1.66	0.07	0.80	0.03

Measure denominator population	EHR System: Epic	EHR System: Epic	EHR System: Cerner	EHR System: Cerner
Sub-groups	*	*	*	*
Age bins	*	*	*	*
18-35	0.23	0.04	0.08	0.02
36-64	1.47	0.10	0.61	0.04
65+	3.40	0.17	1.39	0.06
Sex	*	*	*	*
Male	2.76	0.14	0.96	0.05
Female	1.08	0.07	0.67	0.04
Race	*	*	*	*
White	1.59	0.08	0.85	0.04
Black or African American	2.32	0.17	0.67	0.13
Other	1.13	0.14	0.61	0.06
Unknown	1.37	0.27	0.54	0.17
Ethnicity	*	*	*	*
Hispanic or Latino	1.18	0.15	0.56	0.04
Non-Hispanic	1.75	0.07	0.93	0.04
Unknown	1.41	0.26	0.65	0.20
(Primary) Payer	*	*	*	*
Medicare	3.25	0.18	1.40	0.07
Medicaid	1.25	0.15	0.45	0.04
Private Insurance	1.11	0.07	0.71	0.05
Self-pay or Uninsured	1.08	0.41	0.14	0.10
Other	1.67	1.67	0.43	0.10

Notes: PI = pressure injury; std.err = standard errors. *Cells intentionally left blank.

Previous (2019) Submission:

Data below are from initial development testing; this eCQM is not yet implemented. The measure performance was stratified for disparities by age, race, ethnicity, and payer source.

Hospital Test Site 1 (Beta Dataset 1 per Testing Form)

- Number of hospitals: 1
- Data collection period: 1/1/2017 - 12/31/2017
- Denominator (admissions): 7,573

Hospital Test Site 2 (Beta Dataset 2 per Testing Form)

- Number of hospitals: 21
- Data collection period: 1/1/2017 - 12/31/2017

- Denominator (admissions): 100,238

Hospital Test Site 3 (Beta Dataset 3 per Testing Form)

- Number of hospitals: 2
- Data collection period: 1/1/2017 - 12/31/2017
- Denominator (admissions): 56,330

Category//Denominator//Numerator//Measure Rate (95% Confidence Interval)

Across Sites (n=164,141, 24 hospitals)

Age//Denominator//Numerator//Measure Rate (95% Confidence Interval)

18-64//104,332//401//0.38% (0.3%, 0.4%)

65+//59,809//775//1.30% (1.2%, 1.4%)

Gender//Denominator//Numerator//Measure Rate (95% Confidence Interval)

Male//61,636//664//1.08% (1.0%, 1.2%)

Female//102,503//512//0.50% (0.5%, 0.5%)

Unknown//2//0//0.00% (0.0%, 0.7%)

Race//Denominator//Numerator// Measure Rate (95% Confidence Interval)

Black or African American//7,195//51//0.71% (0.5%, 0.9%)

White//133,894//974//0.73% (0.7%, 0.8%)

Other//21,795//142//0.65% (0.5%, 0.8%)

Unknown//1,257//9//0.72% (0.3%, 1.4%)

Ethnicity//Denominator//Numerator//Measure Rate (95% Confidence Interval)

Hispanic or Latino//18,030//89//0.49% (0.4%, 0.6%)

Non-Hispanic//142,251//1,057//0.74% (0.7%, 0.8%)

Unknown//3,860//30//0.78% (0.5%, 1.1%)

(Primary) Payer//Denominator//Numerator// Measure Rate (95% Confidence Interval)

Medicare//64,913//806//1.24% (1.2%, 1.3%)

Medicaid//12,280//96//0.78% (0.6%, 1.0%)

Private Insurance//75,895//236//0.31% (0.3%, 0.4%)

Self-pay or Uninsured//5,999//9//0.15% (0.1%, 0.3%)

Other (such as other government plans)//4,475//27//0.60% (0.4%, 0.9%)

Unknown//579//2//0.35% (0.0%, 1.2%)

It is important to note these results are derived from a small dataset that is not generalizable to the entire population, and the datasets include many characteristics that are 'unknown' in the EHR, which limits the usability of the results.

[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]

Current (2022) Submission:

N/A

Previous (2019) Submission:

N/A

[Response Ends]

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability

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]

Hospital Harm - Pressure Injury

[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]

This electronic clinical quality measure (eCQM) assesses the proportion of inpatient hospitalizations for patients ages 18 years and older at the start of the encounter who suffer the harm of developing a new stage 2, stage 3, stage 4, deep tissue, or unstageable pressure injury.

[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]

Integumentary: Pressure Injury

[Response Ends]

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

[Response Begins]

Safety: Complications

[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)

[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]

Final measure specifications for implementation will be made publicly available on CMS' appropriate quality website, once finalized through the NQF endorsement and CMS rulemaking processes.

[Response Ends]

sp.10. Indicate whether Health Quality Measure Format (HQMF) specifications are attached.

Attach the zipped output from the eCQM authoring tool (MAT) - if the MAT was not used, contact staff. (Use the specification fields in this online form for the plain-language description of the specifications).

[Response Begins]

HQMF specifications are attached.

[Response Ends]

Attachment: 3498e_PI-v0-1-050-QDM-5-6_For NQF.zip

sp.11. Attach the simulated testing attachment.

All eCQMs require a simulated testing attachment to confirm that the HTML output from Bonnie testing (or testing of some other simulated data set) includes 100% coverage of measured patient population testing, with pass/fail test cases for each sub-population. This can be submitted in the form of a screenshot.

[Response Begins]

Testing is attached

[Response Ends]

Attachment: 3498e_3498e_Bonnie v5.1.1_Measure View - CMS826v0_For NQF-508.pdf

sp.12. 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]

Available in attached Excel or csv file

[Response Ends]

Attachment: 3498e_3498e_Pressure Injury Value Set Directory v2022_For NQF-508.xlsx

For the question below: state the outcome being measured. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.13. 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]

Inpatient hospitalizations for patients with a new deep tissue pressure injury (DTI) or stage 2, 3, 4, or unstageable pressure injury, as evidenced by any of the following:

- A diagnosis of DTI with the DTI not present on admission;
- A diagnosis of stage 2, 3, 4 or unstageable pressure injury with the pressure injury diagnosis not present on admission;
- A DTI found on exam greater than 72 hours after the start of the encounter; or
- A stage 2, 3, 4 or unstageable pressure injury found on exam greater than 24 hours after the start of the encounter.

[Response Ends]

For the question below: describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.14. 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]

This is an eCQM, and therefore uses electronic health record data to calculate the measure score. The time period for data collection is during an inpatient hospitalization, beginning at hospital arrival including time in the emergency department or observation when these encounters are within an hour of the inpatient admission.

All data elements necessary to calculate this numerator are defined within value sets available in the Value Set Authority Center (VSAC) and listed below:

- Deep tissue pressure injuries found on physical exam are represented by the value set Pressure Injury Deep Tissue (2.16.840.1.113762.1.4.1147.112)
- Deep tissue pressure injury diagnoses not present on admission are represented by the value set Pressure Injury Deep Tissue Diagnoses (2.16.840.1.113762.1.4.1147.194)
- Stage 2, 3, 4 or unstageable pressure injuries found on exam are represented by the value set Pressure Injury Stage 2, 3, 4 or Unstageable (2.16.840.1.113762.1.4.1147.113)
- Stage 2, 3, 4 or unstageable pressure injury diagnoses not present on admission are represented by the value set Pressure Injury Stage 2, 3, 4, or Unstageable Diagnoses (2.16.840.1.113762.1.4.1147.196)
- The not present on admission indicators are represented by the value set Not Present On Admission or Documentation Insufficient to Determine (2.16.840.1.113762.1.4.1147.198)

To access the value sets for the measure, please visit the Value Set Authority Center (VSAC), sponsored by the National Library of Medicine, at https://vsac.nlm.nih.gov/.

[Response Ends]

For the question below: state the target population for the outcome. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.15. State the denominator.

Brief, narrative description of the target population being measured.

[Response Begins]

Inpatient hospitalizations where the patient is 18 years of age or older at the start of the encounter.

[Response Ends]

For the question below: describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in sp.22.

sp.16. 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]

This measure includes all inpatient hospitalizations for patients aged 18 years and older at the time of admission, and all payers. Inpatient hospitalizations include time in the emergency department and observation when the transition between these encounters (if they exist) and the inpatient encounter are within an hour or less of each other.

Measurement period is one year. This measure is at the hospital-by-admission level.

Inpatient Encounters are represented using the value set of Encounter Inpatient (2.16.840.1.113883.3.666.5.307).

Emergency Department visits are represented using the value set of Emergency Department Visit (2.16.840.1.113883.3.117.1.7.1.292).

Observation encounters are represented using the value set of Observation Services (2.16.840.1.113762.1.4.1111.143).

To access the value sets for the measure, please visit the Value Set Authority Center, sponsored by the National Library of Medicine, at https://vsac.nlm.nih.gov/.

[Response Ends]

sp.17. Describe the denominator exclusions.

Brief narrative description of exclusions from the target population.

[Response Begins]

- Inpatient hospitalizations for patients with a DTI or stage 2, 3, 4 or unstageable pressure injury diagnosis present on admission.
- Inpatient hospitalizations for patients with a DTI found on exam within 72 hours of the start of the encounter.
- Inpatient hospitalizations for patients with a stage 2, 3, 4, or unstageable pressure injury found on exam within 24 hours of the start of the encounter.
- Inpatient hospitalizations for patients with diagnosis of a COVID-19 infection during the encounter.

[Response Ends]

sp.18. 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]

To qualify for the denominator exclusions: DTI or stage 2, 3, 4 or unstageable pressure injury diagnoses must be present on admission to the qualifying inpatient hospitalization. DTIs found on exam must be within 72 hours of the start of the encounter. Stage 2, 3, 4, or unstageable pressure injuries found on exam must be within 24 hours of the start of the encounter. A diagnosis of COVID-19 infection must be during the qualifying inpatient hospitalization.

All data elements necessary to calculate this numerator are defined within value sets available in the Value Set Authority Center (VSAC) and listed below:

- Deep tissue pressure injuries found on physical exam are represented by the value set Pressure Injury Deep Tissue (2.16.840.1.113762.1.4.1147.112)
- Deep tissue pressure injury diagnoses not present on admission are represented by the value set Pressure Injury Deep Tissue Diagnoses (2.16.840.1.113762.1.4.1147.194)
- Stage 2, 3, 4 or unstageable pressure injuries found on exam are represented by the value set Pressure Injury Stage 2, 3, 4 or Unstageable (2.16.840.1.113762.1.4.1147.113)
- Stage 2, 3, 4 or unstageable pressure injury diagnoses not present on admission are represented by the value set Pressure Injury Stage 2, 3, 4, or Unstageable Diagnoses (2.16.840.1.113762.1.4.1147.196)
- The present on admission indicators are represented by the value set Present on Admission or Clinically Undetermined (2.16.840.1.113762.1.4.1147.197)
- COVID-19 diagnoses are represented by the value set COVID-19 (2.16.840.1.113762.1.4.1248.140)

To access the value sets for the measure, please visit the Value Set Authority Center (VSAC), sponsored by the National Library of Medicine, at https://vsac.nlm.nih.gov/.

[Response Ends]

sp.19. 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]

N/A; this measure is not stratified.

[Response Ends]

sp.20. Is this measure adjusted for socioeconomic status (SES)?

[Response Begins]

No

[Response Ends]

sp.21. 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.22. Select the most relevant type of score.

Attachment: If available, please provide a sample report.

[Response Begins]

Rate/proportion

[Response Ends]

sp.23. 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 = Lower score

[Response Ends]

sp.24. 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]

Target population:

Inpatient admission encounters, all payer, where individuals are aged 18 years or older at the start of the encounter and are discharged within the measurement period.

To create the denominator:

1. If the inpatient admission was during the measurement period, go to Step 2. If not, do not include in measure population.

2. Determine the patient's age in years. The patient's age is equal to the encounter start date minus the birth date. If the patient is 18 years or older, include in the measure population. If less than 18 years old, do not include in the measure population.

3. Apply denominator exclusions to remove encounters from the denominator:

- Remove encounters for patients with a DTI or stage 2, 3, 4 or unstageable pressure injury diagnosis present on admission.
- Remove encounters for patients with a DTI found on exam within 72 hours after the start of the encounter.
- Remove encounters for patients with a stage 2, 3, 4, or unstageable pressure injury found on exam within 24 hours after the start of the encounter.
- Remove encounters for patients with a diagnosis of a COVID-19 infection during the encounter.

To create the numerator:

1. For each encounter identify if the patient develops the harm of a new deep tissue pressure injury (DTI) or stage 2, 3, 4, or unstageable pressure injury as evidenced by any of the following:

- A diagnosis of DTI with the DTI not present on admission.

- A diagnosis of stage 2, 3, 4 or unstageable pressure injury with the pressure injury diagnosis not present on admission.
- A DTI found on exam greater than 72 hours after the start of the encounter.
- A stage 2, 3, 4 or unstageable pressure injury found on exam greater than 24 hours after the start of the encounter.

2. Only the first numerator harm event (a new deep tissue pressure injury (DTI) or stage 2, 3, 4, or unstageable pressure injury) is counted in the numerator, and only one numerator harm event is counted per encounter.

3. To calculate the hospital-level measure result, divide the total numerator events by the total number of qualifying encounters (denominator).

Please see **Figure 1:** Hospital Harm-Pressure Injury Measure Flow Diagram, below:

Figure 1: Hospital Harm-Pressure Injury Measure Flow Diagram

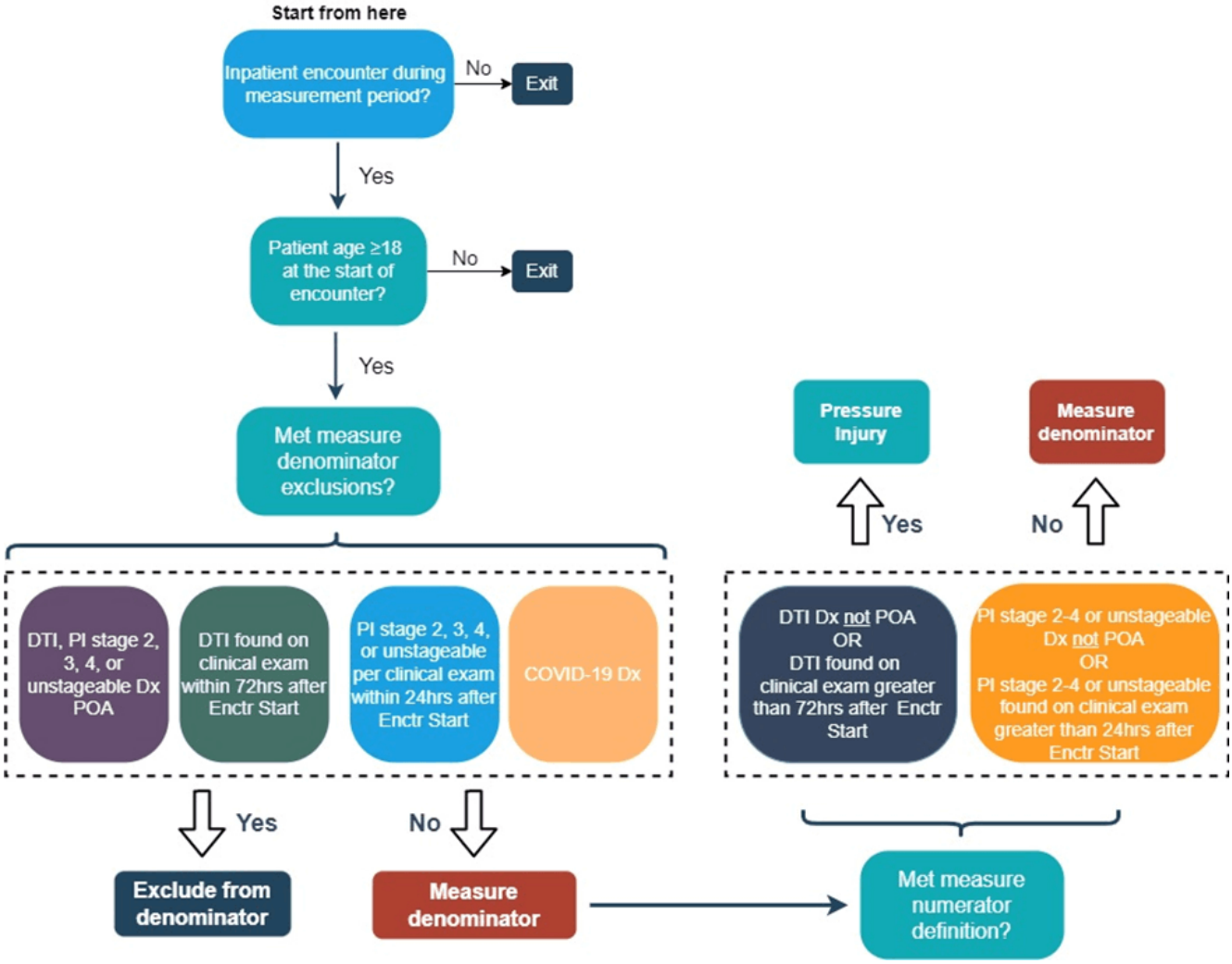


Figure 1 is a flow diagram illustrating the measure numerator and denominator for the pressure injury eCQM.

Note: enctr – encounter; DTI – deep pressure injury; Dx – Diagnosis; PI – pressure injury; POA – present on admission

[Response Ends]

sp.27. If measure testing is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.

Examples of samples used for testing:

- *Testing may be conducted on a sample of the accountable entities (e.g., hospital, physician). The analytic unit specified for the particular measure (e.g., physician, hospital, home health agency) determines the sampling strategy for scientific acceptability testing.*
- *The sample should represent the variety of entities whose performance will be measured. The 2010 Measure Testing Task Force recognized that the samples used for reliability and validity testing often have limited generalizability because measured entities volunteer to participate. Ideally, however, all types of entities whose performance will be measured should be included in reliability and validity testing.*
- *The sample should include adequate numbers of units of measurement and adequate numbers of patients to answer the specific reliability or validity question with the chosen statistical method.*
- *When possible, units of measurement and patients within units should be randomly selected.*

[Response Begins]

N/A; this measure does not use a sample.

[Response Ends]

sp.30. Select only the data sources for which the measure is specified.

[Response Begins]

Electronic Health Records

[Response Ends]

sp.31. 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]

Hospitals collect EHR data using certified electronic health record technology (CEHRT). The MAT output, which includes the human readable and XML artifacts of the clinical quality language (CQL) for the measure are contained in the eCQM specifications attached. No additional tools are used for data collection for eCQMs.

[Response Ends]

sp.32. Provide the data collection instrument.

[Response Begins]

No data collection instrument provided

[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
- 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 Scientific Acceptability sections. For example:

Current Submission:

Updated testing information here.

Previous (Year) Submission:

Testing from the previous submission here.

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

[Response Begins]

Electronic Health Records

[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]

A total of 18 hospitals (test sites) participated in measure testing with varying bed size, geographic location, teaching status, urbanicity, and EHR systems (Epic and Cerner). Data comprised the full calendar year 2020 (Jan 1, 2020 to Dec 31, 2020) from the site's EHR system. No partial year data were used. The number of **unique** patients included in measure denominator ranged from a low of 470 to a high of 30,650 across test sites. Measure denominator encounters ranged from a low of 553 to a high of 38,476 across test sites.

Table 1 below provides the high-level information on the data sets for data from calendar year 2020.

Table 1. High-level Characteristics of Data Set from Test Sites

Hospital	Teaching Status	Urban/Rural	Bed Size	No. of Unique Patients	No. of Denominator Eligible Encounters
1	Academic	Urban	>499	30,650	38,476
2	Non-academic	Rural	25-99	2,996	3,708
3	Academic	Urban	100-199	6,503	7,821
4	Non-academic	Urban	25-99	1,087	1,202
5	Non-academic	Urban	200-499	10,308	12,540
6	Academic	Urban	200-499	10,878	14,576
7	Academic	Urban	200-499	11,318	14,533
8	Non-academic	Urban	100-199	4,196	5,782
9	Non-academic	Urban	100-199	4,902	6,143
10	Non-academic	Urban	200-499	5,468	6,587
11	Non-academic	Rural	25-99	923	1,146
12	Non-academic	Rural	25-99	1,289	1,623
13	Non-academic	Rural	100-199	1,767	2,045
14	Non-academic	Rural	25-99	1,002	1,252
15	Non-academic	Rural	100-199	2,542	3,101
16	Non-academic	Urban	200-499	4,518	5,623
17	Non-academic	Rural	25-99	1,542	1,612
18	Non-academic	Rural	25-99	470	553

Notes: A total of 18 hospitals with two different EHR systems (Epic: Hospital 1 and Cerner: Hospitals 2-18) participated in measure testing. Data from test sites' EHR systems in CY2020 were used.

[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]

01-01-2020– 12-31-2020

[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]

A total of 18 hospitals (test sites) with varying bed size, geographic location, teaching status, urbanicity, and EHR systems participated in measure testing. A majority of test sites were in the Western U.S. Of the 18 test sites, 10 are in urban areas and four are teaching hospitals. Bed size ranged from a low of little more than 25 beds to a high of more than 499 beds.

Table 2. Hospital Test Site Characteristics

Health System	Hospital Test Site	EHR System	Census Region	Bed Size	Teaching Status	Urban/Rural
A	1	Epic	Midwest	> 499	Academic	Urban
B	2	Cerner	West	25-99	Non-academic	Rural
B	3	Cerner	West	100-199	Academic	Urban
B	4	Cerner	West	25-99	Non-academic	Urban
B	5	Cerner	West	200-499	Non-academic	Urban
B	6	Cerner	West	200-499	Academic	Urban
B	7	Cerner	West	200-499	Academic	Urban
B	8	Cerner	West	100-199	Non-academic	Urban

Health System	Hospital Test Site	EHR System	Census Region	Bed Size	Teaching Status	Urban/Rural
B	9	Cerner	West	100-199	Non-academic	Urban
B	10	Cerner	West	200-499	Non-academic	Urban
B	11	Cerner	West	25-99	Non-academic	Rural
B	12	Cerner	West	25-99	Non-academic	Rural
B	13	Cerner	West	100-199	Non-academic	Rural
B	14	Cerner	West	25-99	Non-academic	Rural
B	15	Cerner	West	100-199	Non-academic	Rural
B	16	Cerner	West	200-499	Non-academic	Urban
B	17	Cerner	West	25-99	Non-academic	Rural
B	18	Cerner	West	25-99	Non-academic	Rural

[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]

We collected data in calendar year 2020 (1/1/2020 and 12/31/2020) from 18 test sites. **Table 3** below provides information on measure denominator population, stratified by sex, age bins, race/ethnicity, and primary source of payment.

Table 3. Measure Denominator Population Characteristics

Measure Denominator Population Characteristics	EHR System: Epic	EHR System: Epic	EHR System: Cerner	*
*	n	%	n	%
Number of denominator inpatient encounters	38,476	100%	89,847	100%
Number of unique patients	30,650	100%	70,883	100%
Age Mean (Std.Dev)	50.9 (19.2)	*	56.1 (21.1)	*
Age bins	*	*	*	*
18-35	10,477	34%	19,500	28%
36-64	11,888	39%	24,736	35%
65+	8,291	27%	26,718	38%
Sex	*	*	*	*

Measure Denominator Population Characteristics	EHR System: Epic	EHR System: Epic	EHR System: Cerner	*
Male	9,562	31%	28,095	40%
Female	21,088	69%	42,788	60%
Race	*	*	*	*
White	18,428	60%	53,044	75%
Black or African American	5,681	19%	2,795	4%
Other	4,966	16%	13,363	19%
Unknown	1,575	5%	1,681	2%
Ethnicity	*	*	*	*
Hispanic or Latino	4,008	13%	24,717	35%
Non-Hispanic	24,848	81%	44,735	63%
Unknown	1,794	6%	1,431	2%
(Primary) Payer	*	*	*	*
Medicare	6,730	22%	18,446	26%
Medicaid	4,315	14%	24,455	35%
Private Insurance	19,192	63%	24,247	34%
Self-pay or Uninsured	528	2%	1,337	2%
Other	56	0%	3,458	5%

Notes: std.dev = standard deviation. Of the 18 test sites, one uses Epic and 17 use Cerner as their EHR system. Not all bins total to 100% due to rounding. *Cells intentionally left empty.

Tables 4 to **8** present characteristics of the measure denominator population for each of the 18 test sites. In 2020, 470 to 30,650 unique patients contributed from 553 to 38,476 denominator encounters, across sites. Note that while the measure is inpatient-based, the measure denominator includes emergency department visits and observation stays that eventually turned into inpatient hospitalizations.

Tables 4 to **8** reveal that the measure denominator population varied widely across sites, from 40% to 95% White, from 1% to 19% Black, from 5% to 84% Hispanic or Latino, from 11% to 55% elderly (65 or over), from 43% to 90% female, from 14% to 73% covered by Medicaid, and from 10% to 63% covered by private insurance.

Table 4. Measure Denominator Population Characteristics (Sites 1-4)

Measure Denominator Population Characteristics	Test Site 1	Test Site 1	Test Site 2	Test Site 2	Test Site 3	Test Site 3	Test Site 4	Test Site 4
*	n	%	n	%	n	%	n	%
Number of encounters	38,476	100%	3,708	100%	7,821	100%	1,202	100%
Number of unique patients	30,650	100%	2,996	100%	6,503	100%	1,087	100%

Measure Denominator Population Characteristics	Test Site 1	Test Site 1	Test Site 2	Test Site 2	Test Site 3	Test Site 3	Test Site 4	Test Site 4
Age Mean (Std.Dev)	50.9 (19.2)		62.2 (20.1)		50.3 (21.2)		57.6 (19.6)	
Age bins	*	*	*	*	*	*	*	*
18-35	10,477	34%	547	18%	2,605	40%	185	17%
36-64	11,888	39%	851	28%	2,127	33%	475	44%
65+	8,291	27%	1,601	53%	1,776	27%	427	39%
Sex	*	*	*	*	*	*	*	*
Male	9,562	31%	1,224	41%	2,032	31%	505	46%
Female	21,088	69%	1,772	59%	4,471	69%	582	54%
Race	*	*	*	*	*	*	*	*
White	18,428	60%	2,841	95%	5,708	88%	1,007	93%
Black or African American	5,681	19%	17	1%	335	5%	7	1%
Other	4,966	16%	119	4%	343	5%	69	6%
Unknown	1,575	5%	19	1%	117	2%	4	0%
Ethnicity	*	*	*	*	*	*	*	*
Hispanic or Latino	4,008	13%	155	5%	3,542	54%	749	69%
Non-Hispanic	24,848	81%	2,807	94%	2,817	43%	322	30%
Unknown	1,794	6%	34	1%	144	2%	16	1%
(Primary) Payer	*	*	*	*	*	*	*	*
Medicare	6,730	22%	1,511	50%	1,451	22%	313	29%
Medicaid	4,315	14%	636	21%	2,762	42%	455	42%
Private Insurance	19,192	63%	699	23%	1,747	27%	266	24%
Self-pay or Uninsured	528	2%	20	1%	99	2%	44	4%
Other	56	0%	141	5%	478	7%	11	1%
Unknown	38,476	100%	3,708	100%	7,821	100%	1,202	100%

Notes: std.dev = standard deviation. Not all bins total to 100% due to rounding. *Cells intentionally left empty.

Table 5. Measure Denominator Population Characteristics (Sites 5-8)

Measure Denominator Population Characteristics	Test Site 5	Test Site 5	Test Site 6	Test Site 6	Test Site 7	Test Site 7	Test Site 8	Test Site 8
*	n	%	n	%	n	%	n	%
Number of encounters	12,540	100%	14,576	100%	14,533	100%	5,782	100%
Number of unique patients	10,308	100%	10,878	100%	11,318	100%	4,196	100%
Age Mean (Std. Dev)	54.6 (20.4)	*	60.9 (21.0)	*	49.5 (20.0)	*	61.5 (19.4)	*
Age bins	*	*	*	*	*	*	*	*
18-35	2,918	28%	2,281	21%	4,201	37%	694	17%
36-64	3,829	37%	3,673	34%	4,378	39%	1,529	36%
65+	3,567	35%	4,932	45%	2,754	24%	1,976	47%
Sex	*	*	*	*	*	*	*	*
Male	4,131	40%	4,517	42%	4,115	36%	1,865	44%
Female	6,177	60%	6,361	58%	7,203	64%	2,331	56%
Race	*	*	*	*	*	*	*	*
White	7,357	71%	8,252	76%	7,512	66%	3,239	77%
Black or African American	720	7%	350	3%	644	6%	54	1%
Other	2,043	20%	1,858	17%	2,943	26%	830	20%
Unknown	188	2%	418	4%	219	2%	73	2%
Ethnicity	*	*	*	*	*	*	*	*
Hispanic or Latino	4,254	41%	2,366	22%	9,027	80%	789	19%
Non-Hispanic	5,891	57%	8,218	76%	2,117	19%	3,359	80%
Unknown	163	2%	294	3%	174	2%	48	1%
(Primary) Payer	*	*	*	*	*	*	*	*
Medicare	1,837	18%	3,807	35%	1,392	12%	1,348	32%
Medicaid	2,716	26%	3,272	30%	6,434	57%	701	17%
Private Insurance	5,127	50%	3,433	32%	2,790	25%	2,017	48%
Self-pay or Uninsured	124	1%	258	2%	376	3%	88	2%
Other	561	5%	688	6%	402	4%	123	3%
Unknown	12,540	100%	14,576	100%	14,533	100%	5,782	100%

Notes: std.dev = standard deviation. Not all bins total to 100% due to rounding. *Cells intentionally left empty.

Table 6. Measure Denominator Population Characteristics (Sites 9-12)

Measure Denominator Population Characteristics	Test Site 9	Test Site 9	Test Site 10	Test Site 10	Test Site 11	Test Site 11	Test Site 12	Test Site 12
	n	%	n	%	n	%	n	%
Number of encounters	6,143	100%	6,587	100%	1,146	100%	1,623	100%
Number of unique patients	4,902	100%	5,468	100%	923	100%	1,289	100%
Age Mean (Std.Dev)	53.6 (21.6)	*	60.9 (19.0)	*	61.3 (20.4)	*	64.0 (17.0)	*
Age bins	*	*	*	*	*	*	*	*
18-35	1,487	30%	878	16%	204	22%	118	9%
36-64	1,701	35%	2,006	37%	254	28%	466	36%
65+	1,723	35%	2,591	47%	466	50%	706	55%
Sex	*	*	*	*	*	*	*	*
Male	2,153	44%	2,572	47%	365	40%	595	46%
Female	2,749	56%	2,896	53%	558	60%	694	54%
Race	*	*	*	*	*	*	*	*
White	1,968	40%	4,390	80%	875	95%	1,082	84%
Black or African American	102	2%	291	5%	0	0%	11	1%
Other	2,627	54%	568	10%	38	4%	174	13%
Unknown	205	4%	219	4%	10	1%	22	2%
Ethnicity	*	*	*	*	*	*	*	*
Hispanic or Latino	311	6%	273	5%	50	5%	99	8%
Non-Hispanic	4,493	92%	4,954	91%	864	94%	1,176	91%
Unknown	98	2%	241	4%	9	1%	14	1%
(Primary) Payer	*	*	*	*	*	*	*	*
Medicare	930	19%	1,097	20%	352	38%	671	52%
Medicaid	1,363	28%	1,404	26%	212	23%	330	26%

Measure Denominator Population Characteristics	Test Site 9	Test Site 9	Test Site 10	Test Site 10	Test Site 11	Test Site 11	Test Site 12	Test Site 12
Private Insurance	2,220	45%	2,729	50%	310	34%	222	17%
Self-pay or Uninsured	55	1%	102	2%	6	1%	18	1%
Other	364	7%	162	3%	50	5%	54	4%
Unknown	6,143	100%	6,587	100%	1,146	100%	1,623	100%

Notes: std.dev = standard deviation. Not all bins total to 100% due to rounding. *Cells intentionally left empty.

Table 7. Measure Denominator Population Characteristics (Sites 13-16)

Measure Denominator Population Characteristics	Test Site 13	Test Site 13	Test Site 14	Test Site 14	Test Site 15	Test Site 15	Test Site 16	Test Site 16
*	n	%	n	%	n	%	n	%
Number of encounters	2,045	100%	1,252	100%	3,101	100%	5,623	100%
Number of unique patients	1,767	100%	1,002	100%	2,542	100%	4,518	100%
Age Mean (Std.Dev)	64.6 (15.9)	*	58.1 (20.3)	*	54.9 (21.9)	*	58.1 (21.5)	*
Age bins	*	*	*	*	*	*	*	*
18-35	120	7%	223	22%	825	32%	1,158	26%
36-64	691	39%	359	36%	787	31%	1,483	33%
65+	958	54%	420	42%	930	37%	1,883	42%
Sex	*	*	*	*	*	*	*	*
Male	1,013	57%	393	39%	872	34%	1,776	39%
Female	754	43%	609	61%	1,670	66%	2,742	61%
Race	*	*	*	*	*	*	*	*
White	1,451	82%	821	82%	1,814	71%	3,774	84%
Black or African American	57	3%	34	3%	18	1%	144	3%
Other	199	11%	127	13%	650	26%	566	13%
Unknown	60	3%	20	2%	60	2%	34	1%
Ethnicity	*	*	*	*	*	*	*	*
Hispanic or Latino	169	10%	127	13%	497	20%	1,192	26%
Non-Hispanic	1,534	87%	867	87%	1,994	78%	3,273	72%

Measure Denominator Population Characteristics	Test Site 13	Test Site 13	Test Site 14	Test Site 14	Test Site 15	Test Site 15	Test Site 16	Test Site 16
Unknown	64	4%	8	1%	51	2%	53	1%
(Primary) Payer	*	*	*	*	*	*	*	*
Medicare	927	52%	452	45%	905	36%	1,524	34%
Medicaid	337	19%	384	38%	1,018	40%	1,471	33%
Private Insurance	424	24%	105	10%	519	20%	1,363	30%
Self-pay or Uninsured	7	0%	18	2%	35	1%	70	2%
Other	85	5%	49	5%	76	3%	187	4%
Unknown	2,045	100%	1,252	100%	3,101	100%	5,623	100%

Notes: std.dev = standard deviation. Not all bins total to 100% due to rounding. *Cells intentionally left empty.

Table 8. Measure Denominator Population Characteristics (Sites 17 and 18)

Measure Denominator Population Characteristics	Test Site 17	Test Site 17	Test Site 18	Test Site 18
*	n	%	n	%
Number of encounters	1,612	100%	553	100%
Number of unique patients	1,542	100%	470	100%
Age Mean (Std.Dev)	36.8 (18.9)	*	62.6 (18.1)	*
Age bins	*	*	*	*
18-35	1,087	70%	46	10%
36-64	279	18%	197	42%
65+	176	11%	227	48%
Sex	*	*	*	*
Male	151	10%	243	52%
Female	1,391	90%	227	48%
Race	*	*	*	*
White	1,264	82%	375	80%
Black or African American	6	0%	37	8%
Other	259	17%	52	11%
Unknown	13	1%	6	1%
Ethnicity	*	*	*	*
Hispanic or Latino	1,301	84%	72	15%
Non-Hispanic	221	14%	391	83%

Measure Denominator Population Characteristics	Test Site 17	Test Site 17	Test Site 18	Test Site 18
Unknown	20	1%	7	1%
(Primary) Payer	*	*	*	*
Medicare	142	9%	172	37%
Medicaid	1,120	73%	118	25%
Private Insurance	266	17%	134	29%
Self-pay or Uninsured	11	1%	9	2%
Other	4	0%	39	8%
Unknown	1,612	100%	553	100%

Notes: std.dev = standard deviation. Not all bins total to 100% due to rounding.

*Cells intentionally left empty.

[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]

Measure score level reliability and validity testing used data from the full denominator population.

Measure data element level validity testing, on the other hand, were based on subsamples drawn from the measure initial population using the approach of random sampling without replacement. These subsamples served as the foundation upon which clinical abstractors compared data exported from the EHR (eData) to data manually abstracted from patients' medical charts (mData, or "gold standard"). This process is commonly known as the parallel-form comparison. When drawing the subsamples, we held constant the distribution of patient characteristics exhibited in the initial population to the extent possible (e.g., % of male, % of white, % of black, etc. in the abstraction sample are comparable to those in the initial population to the extent possible).

[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]

We collected patient race, ethnicity, and primary source of payment and have shown how the measure denominator population differs within each of these dimensions and how measure performance rate varies within the subpopulation. **Table 9** below provides information on measure denominator population, stratified by sex, age bins, race/ethnicity, and primary source of payment.

Table 9. Measure Denominator Population Characteristics

Measure Denominator Population Characteristics	EHR System: Epic	EHR System: Epic	EHR System: Cerner	EHR System: Cerner
*	n	%	n	%
Number of denominator inpatient encounters	38,476	100%	89,847	100%
Number of unique patients	30,650	100%	70,883	100%
Age Mean (Std.Dev)	50.9 (19.2)	*	56.1 (21.1)	*
Age bins	*	*	*	*
18-35	10,477	34%	19,500	28%
36-64	11,888	39%	24,736	35%
65+	8,291	27%	26,718	38%
Sex	*	*	*	*
Male	9,562	31%	28,095	40%
Female	21,088	69%	42,788	60%
Race	*	*	*	*
White	18,428	60%	53,044	75%
Black or African American	5,681	19%	2,795	4%
Other	4,966	16%	13,363	19%
Unknown	1,575	5%	1,681	2%
Ethnicity	*	*	*	*
Hispanic or Latino	4,008	13%	24,717	35%
Non-Hispanic	24,848	81%	44,735	63%
Unknown	1,794	6%	1,431	2%
(Primary) Payer	*	*	*	*
Medicare	6,730	22%	18,446	26%
Medicaid	4,315	14%	24,455	35%
Private Insurance	19,192	63%	24,247	34%
Self-pay or Uninsured	528	2%	1,337	2%
Other	56	0%	3,458	5%

Notes: std.dev = standard deviation. Of the 18 test sites, one uses Epic and 17 use Cerner as their EHR system. Not all bins total to 100% due to rounding. *Cells intentionally left empty.

[Response Ends]

Note: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a.09 check patient or encounter-level data; in 2a.010 enter "see validity testing section of data elements"; and enter "N/A" for 2a.11 and 2a.12.

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

Choose one or both levels.

[Response Begins]

Accountable Entity Level (e.g., signal-to-noise analysis)

[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]

To assess the measure score level reliability, we used Adams' signal-to-noise ratio (SNR) and the intra-class correlation coefficient (ICC) via the split-half sample approach.

To implement an empirical approach based on Adams' SNR, consider that each hospital has a true measure performance rate $\hat{p}$ that follows a beta distribution. The true rate varies from hospital to hospital due to variation in hospital quality of care in general or variation in the extent to which hospitals exert efforts to prevent PI in particular. The observed measure rate $\hat{p}$, on the other hand, is binomially distributed (whether or not PI occurred) conditional on the true rate p. Observed rate $\hat{p}$ also varies and will vary in any given time period (e.g., calendar year 2020) either due to that the number of events occurring in a selected window is small or the random variation around the true rate p.

Based on the setup, the alpha and beta parameters underlying the beta distribution can be estimated and then used to calculate the hospital-to-hospital variance, which is frequently known as the signal. This signal records the proportion of variability across measured entities that are attributable to the real difference in quality of care. The hospital-specific, or within-hospital, variance can be calculated from the conventional method for any random binomial variable. Therefore, $o^2 hospital - to - hospital = \frac{\alpha\beta}{(\alpha+\beta+1)(\alpha+\beta)^2}$ where α and β are the estimated alpha and beta parameters within the testing data, and $\sigma^2 within - hospital = \frac{\hat{p}(1-\hat{p})}{n}$ where $\hat{p}$ is the observed measure rate for a given hospital and n is the denominator size for that hospital. Reliability, or SNR, is thus equal to $\frac{\sigma^2 hospital-to-hospital}{\sigma^2\ hospital-to-hospital+\ \sigma^2 within-hospital)}$ To motivate the empirical ICC based on the split-half sample approach, consider that hospital $h_i (i = 1, \dots, H)$ in subsample $T_i (i = 1, \dots, T)$ and each hospital subsample T_i is comprised of a possibly varying number of denominator encounters n_{ht}. We assume that the measure performance rate, y_{ht}, follows a simple two-level model: $Y_{ht} = \mu + \alpha_h + \varepsilon_{ht}$ where the hospital-level effects a_h are sampled from a normal distribution with mean 0 and variance σ_h^2 and the residual errors are independently and normally distributed with mean 0 and variance $\frac{(\sigma_e^2)}{n_{ht}{}^{1}}$. The subsamples here could come from different calendar periods or from randomly generated subsamples (e.g. split-halves) of all denominator encounters, stratified by hospital. Note that the specification of residual error variance assumes that, conditional on the hospital random effects a_h, the variance is *inversely* proportional to the sample size used to form the hospital-subsample estimate. Although such a model can be directly motivated by assuming that encounter-level data follow the standard two-level model for normally distributed data (frequently used in classical testing theory), and that encounter-level data from the same hospital and subsamples are then averaged to form the estimated hospital performance, the proposed model can apply more generally.

The two variance components σ_h^2 and σ_e^2 can be estimated by any statistical software that is capable of fitting maximum likelihood methods. By deriving the estimates of σ_h^2 and σ_e^2, we then compute a "plug-in" estimator of the ICC for performance indicator $CC_h = \frac{\sigma_h^2}{\sigma_h^2 + \frac{\sigma_e^2}{n}} = \frac{nR}{nR+1}$, where $R = \frac{\sigma_h^2}{\sigma_e^2}$. Note that ICC is a function only of the size of the denominator and the ratio of between-hospital to within-hospital variance. The higher the SNR or ICC the higher the statistical reliability of the measure, and the greater the amount of variation can be attributed to systematic differences in performance across hospitals (i.e., signals as opposed to noises).

We used the rubric established by Landis and Koch (1977) to interpret the estimated SNRs and ICCs:[2]

- 0 – 0.2: slight agreement
- 0.21 – 0.39: fair agreement
- 0.4 – 0.59: moderate agreement
- 0.6 – 0.79: substantial agreement
- 0.8 – 0.99: almost perfect agreement
- 1: perfect agreement

We note that assumptions underlying the calculation of SNR (or ICC) are potentially very strong when only 18 hospitals participated in measure testing. To gauge the impact of number of hospitals on the SNR estimation, we ran two simulation tests. In the first test, we randomly selected a subset of hospitals and estimated each hospital's SNR in that subsample. We used random sampling with replacement and hence, "small" hospitals (we define "small" hospitals as those with no more than 50 beds) could be selected in or selected out. In our testing these are hospitals 4, 11, 12, 14, 17. The second test was similar in spirit except that we always included "small" hospitals in the subsample. The number of hospitals included ranged from six to the full set of 18. In the section that follows, we present the scatterplot of the median SNR estimated within a given sample against the number of hospitals included in that subsample.

References

1. Dickens, William T. "Error components in grouped data: is it ever worth weighting?." *The Review of Economics and Statistics* (1990): 328-333.

2. Landis, J. Richard, and Gary G. Koch. "The measurement of observer agreement for categorical data." *biometrics* (1977): 159-174.

[Response Ends]

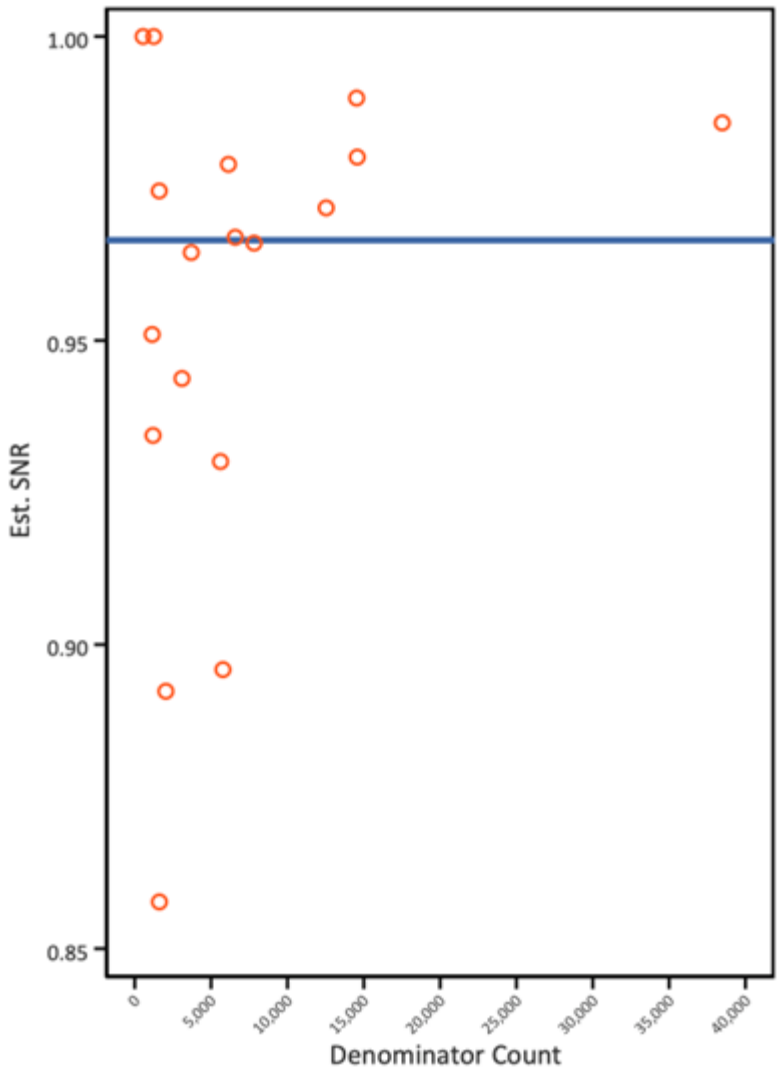
2a.11. For each level of reliability testing checked above, what were the statistical results from reliability testing?

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]

Using the encounter-level data from 18 test sites and in CY 2020, Adams' SNRs ranged from 0.86 to 1.00, with the mean and median equal to 0.96 and 0.97 respectively. **Exhibit 1** below shows the distribution of SNRs across test sites with different denominator size. HH PI demonstrates robust score-level reliability; however, Adams' formula generates a perfect score of 1.0 at two sites with zero numerator events, which does not reflect their latent true reliability.

Exhibit 1. Distribution of SNRs Across 18 Hospital Sites

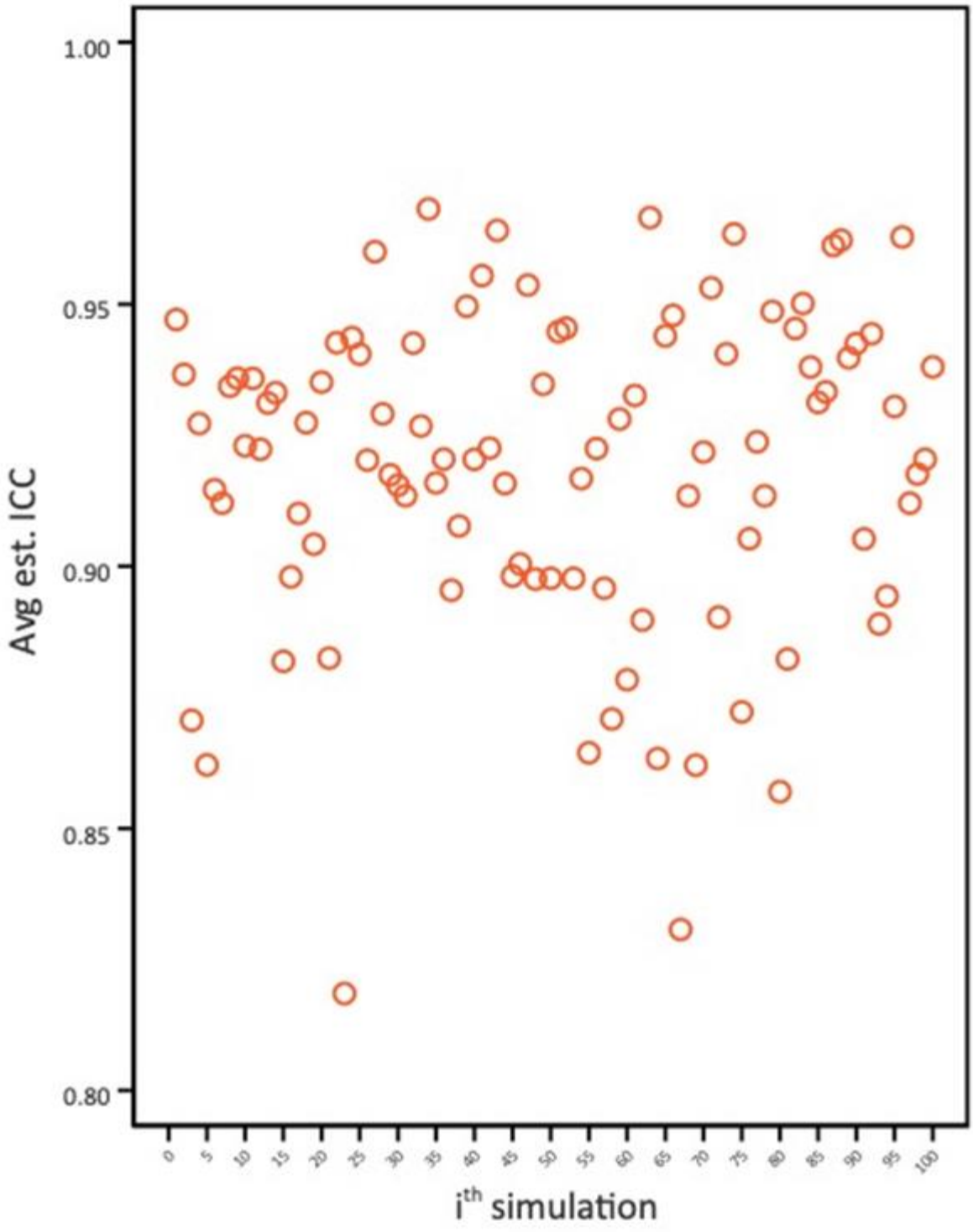


Note: Each red circle indicates an estimated SNR. The blue horizontal line denotes the median value of SNR from the distribution.

Empirical findings are supportive, but we underscore that interpretation requires caution. First, SNR quantifies the score-level (i.e., hospital-level) and not patient-level reliability and hence, estimation accuracy is dependent upon the number of hospitals used for the analysis. With only 18 hospitals in testing, it is possible that the alpha and beta parameters underlying the SNR are estimated with noise. Second, measure performance rates across sites are relatively low but exhibit wide dispersion. To gauge the impact of hospital counts on SNR estimation, we ran two simulation tests. In the first test, we randomly selected a subset of hospitals and estimated each hospital's SNR in that subsample. We used random sampling with replacement and hence, "small" hospitals (we define "small" hospitals as those with no more than 50 beds, and in our testing data these are hospitals 4, 11, 12, 14, 17, and 18) could be selected in or selected out. The second test was similar except that we always included "small" hospitals in the subsample. The number of hospitals included ranged from six to the full set of 18. Our testing results showed that the median SNR always exceeded 0.8.

Next, we estimated ICCs using the split-half sample approach. Since the ICC is another form of signal-to-noise ratio, we anticipate a similar conclusion in light of the findings above. To avoid any one-time estimate being driven by chance, we performed the sample split and ICC estimation 100 times. The 100 estimated ICCs showed mean values ranging from 0.79 to 0.97. **Exhibit 2** below shows the distribution of mean values of estimated ICC from the 100 simulation runs.

Exhibit 2. Distribution of Average Estimated ICC Via the Split-half Sample Approach



Note: Each red circle indicates the average value of ICC across 18 test sites from a given sample split. A total number of 100 different sample splits was performed.

[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]

As shown in the section above, HH PI demonstrates robust score-level reliability, evaluated by Adams' SNR and ICC via the split-half sample approach. Specifically, Adams' SNRs ranged from 0.86 to 1.00 across test sites, with the mean and median equal to 0.96 and 0.97 respectively. Analogously, the 100 estimated ICCs had a median of 0.99 and a mean ranging from 0.79 to 0.97. A total of 128,323 qualified inpatient encounters across 18 test sites contributed to the calculation of reliability estimates.

Overall, testing results clearly showed that HH PI, as currently specified, can distinguish the true performance in HA-PI from one hospital to another.

[Response Ends]

2b. Validity

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

[Response Begins]

Patient or Encounter-Level (data element validity must address ALL critical data elements)

Accountable Entity Level (e.g. hospitals, clinicians)

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]

To empirically assess the data element validity, we compared data exported from the EHR (eData) to data manually abstracted from patients' medical charts (mData) for a subsample of measure initial population. We then quantified the validity by calculating 1) the frequency of missingness for all data elements needed for measure implementation, 2) percentage of match/agreement in each data element between data sources (eData vs. mData), and 3) positive predicative value (PPV), sensitivity, negative predicative value (NPV), and specificity. The first two statistics tell us if measure performance would be severely biased if data missing is systematic and the latter statistics tells us if the measure is subject to false positives and false negatives. We note that information in patients' medical charts is typically deemed the "gold standard," and hence Cohen's Kappa is not relevant here. This process of data comparison is frequently known as the parallel-form comparison.

Expectedly, manual abstraction is labor intensive; therefore, reducing burden while maximizing test result validity (e.g., level of power and significance) is important. To that end, we calculated the minimum required sample size (MRSS) for the abstraction using PPV as the primary endpoint and approximated MRSS using the conventional one-sample proportion formula, while accounting for the intracluster correlation:[1] $n = \frac{\frac{z_{\alpha}^2}{2} \cdot p \cdot (1-p)}{moe^2} \times VIF$ where α denotes the type I error rate, moe denotes the margin of error, p is PPV, and VIF is the variance inflation factor that accounts for the intracluster correlation. We simulated a series of $moes$, target ps and the 95% confidence intervals associated with each p for different MRSS. Simulations indicated that with a moe of 6%, a target PPV of 0.9, a reasonable precision of

PPV bounded by 0.84 and 0.96, and a conventionally accepted minimum number of observations that can render the sampling distribution of p to be normal, MRSS approximated 125.

We therefore randomly sampled 125 measure denominator encounters from site 1 (Epic site) and another set of 125 cases across sites 2-18 (Cerner sites under one health system). We sampled abstraction cases across all 17 Cerner sites to enhance efficiency given their somewhat homogeneous clinical workflows and a shared central data warehouse. Manual chart review of patient medical records from any one site would thus be informative of records from other sites of the same health system. We additionally sampled 30 denominator-excluded encounters from the initial population for site 1 and 30 denominator-excluded encounters across sites 2-18 to assess whether excluded cases per eData truly met the clinical intent for exclusion.

In the midst of testing with site 1 we added COVID-19 as a new denominator exclusion criterion[2] in response to stakeholders' concerns about skin manifestations of patients with COVID-19 symptoms.

We assessed measure score level validity using convergent validity, a concept that refers to whether multiple measures of an underlying concept are correlated in the direction (positive or negative) suggested by theory. For this exercise, we collected test sites' patient safety outcomes from a set of related measures (e.g., healthcare associated infections and nursing care) on Hospital Care Compare (data.cms.gov) and estimated Spearman's rank correlation coefficients between HH PI and each of the related measures at the hospital level. Positive (or negative, pending context) correlations then provide support for the score level validity.

Notes:

[1] What we mean by intracluster correlation here is a notion that hospitals with the same EHR system may have seen patients who are more alike. In this case, information revealed by patient A will not be entirely independent from that revealed by patient B. On the contrary, two sets of information share similarities and exhibit strong correlation. Without accounting for such intracluster correlation, we run the risk of underestimating sample size needed to yield a desired level of power and significance for the test statistics.

[2] At the time of abstraction sample creation for site 1, COVID-19 was not a denominator exclusion and therefore not all 125 cases met the final denominator measure specification.

[Response Ends]

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

Examples may include correlations or t-test results.

[Response Begins]

Data Element Validity: Percent Agreement for Data Elements

Tables 10 to 13 below show the data element validity and percent agreement for data elements for the 17 Cerner sites combined and for the single Epic site, respectively. Of note, measure concept (e.g., patient had a stage 2 to 4 or unstageable pressure injury according to the clinical examination documented after 24 hours of encounter start) calculation was based on either the eData alone or the mData alone.

Table 10. Data Element Validity and Percent Agreement for Data Elements; Measure Initial Population (Epic Site)

*	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)
Data Element	**Cases per EHR**	**Cases per Abstraction**	**Percent Agreement**

*	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)
Patient had an inpatient encounter with discharge date between 1/1/20 and 12/31/20 (measurement period)	155	155	100%
Patient aged 18 or older at the start of encounter	155	155	100%
Patient had a stage 2-4 or unstageable pressure injury per the clinical exam after 24 hours of encounter start	57	56	98%
Patient had a stage 2-4 or unstageable pressure injury ICD-10-CM indicating not present-on-admission	7	7	100%
Patient had a deep tissue injury per the clinical examination after 72 hours of encounter start	10	11	99%
Patient had a deep tissue pressure injury ICD-10-CM indicating not present-on-admission	3	3	100%

Notes: *Cells intentionally left empty.

Table 11. Data Element Validity and Percent Agreement for Data Elements; Measure Initial Population (Cerner Sites)

*	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)
Data Element	Cases per EHR	Cases per Abstraction	Percent Agreement
Patient had an inpatient encounter with discharge date between 1/1/20 and 12/31/20 (measurement period)	155	155	100%
Patient aged 18 or older at the start of encounter	155	155	100%
Patient had a stage 2-4 or unstageable pressure injury per the clinical exam after 24 hours of encounter start	53	54	99%
Patient had a stage 2-4 or unstageable pressure injury ICD-10-CM indicating not present-on-admission	1	1	100%
Patient had a deep tissue injury per the clinical examination after 72 hours of encounter start	11	12	99%
Patient had a deep tissue pressure injury ICD-10-CM indicating not present-on-admission	2	2	100%

Notes: *Cells intentionally left empty.

Table 12. Data Element Validity and Percent Agreement for Data Elements; Measure Denominator Exclusion (Epic Site)

*	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)
Data Element	**Cases per EHR**	**Cases per Abstraction**	**Percent Agreement**
Patient had a stage 2- 4 or unstageable pressure injury per clinical exam within 24 hours of encounter start	24	24	100%
Patient had a stage 2- 4 or unstageable pressure injury ICD-10-CM indicating present-on-admission	15	15	100%
Patient had a deep tissue pressure injury per clinical examination within 72 hours of encounter start	1	1	100%
Patient had a deep tissue pressure injury ICD-10-CM indicating present-on-admission	1	1	100%
Patient had a ICD-10-CM indicating COVID-19 infection	19	19	100%

Notes: *Cells intentionally left empty.

Table 13. Data Element Validity and Percent Agreement for Data Elements; Measure Denominator Exclusion (Cerner Sites)

*	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)
Data Element	**Cases per EHR**	**Cases per Abstraction**	**Percent Agreement**
Patient had a stage 2- 4 or unstageable pressure injury per clinical exam within 24 hours of encounter start	5	6	99%
Patient had a stage 2- 4 or unstageable pressure injury ICD-10-CM indicating present-on-admission	5	5	100%
Patient had a deep tissue pressure injury per clinical examination within 72 hours of encounter start	3	2	99%
Patient had a deep tissue pressure injury ICD-10-CM indicating present-on-admission	1	1	100%
Patient had a ICD-10-CM indicating COVID-19 infection	21	21	100%

Notes: *Cells intentionally left empty.

Data Element Validity: PPV, NPV, Sensitivity and Specificity

Tables 14 to 16 present findings for PPV, sensitivity, NPV, and specificity, calculated for measure initial population, measure denominator exclusion, measure numerator negative, and measure numerator, respectively.

Table 14. Measure Data Element Validity (PPV, Sensitivity, NPV, Specificity); Sampled Cases (Site 1)

Element	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)
Element	**Total cases per EHR**	**Total cases per abstraction**	**PPV**	**Sensitivity**	**NPV**	**Specificity**
Initial population	**155**	**155**	**100%**	**100%**	**100%**	**100%**
Denominator exclusion	**47**	**47**	**100%**	**100%**	**100%**	**100%**
Numerator negative	**60**	**61**	**100%**	**98%**	**99%**	**100%**
Numerator	**48**	**47**	**98%**	**100%**	**100%**	**99%**

Note: PPV – positive predicative value; NPV – negative predicative value.

Table 15. Measure Data Element Validity (PPV, Sensitivity, NPV, Specificity); Sampled Cases (Sites 2-18)

Element	Total cases per EHR	Total cases per abstraction	PPV	Sensitivity	NPV	Specificity
Initial population	**155**	**155**	**100%**	**100%**	**100%**	**100%**
Denominator exclusion	**30**	**29**	**97%**	**100%**	**100%**	**99%**
Numerator negative	**62**	**62**	**100%**	**100%**	**100%**	**100%**
Numerator	**63**	**64**	**100%**	**98%**	**99%**	**100%**

Note: PPV – positive predicative value; NPV – negative predicative value.

Measure Score Validity

Table 16 below shows the Spearman rank correlation coefficients between test sites' PI scores and six infection-related measure scores, along with the number of sites used in the calculation and two-sided p-values.1

Table 16. Convergent Validity Between HH PI and Infection-Related Outcome Measures by Spearman's Rank Correlation

Hospital Compare Measures	Rho	Number of hospitals	Two-sided P-val
Central line-associated bloodstream infection (CLABSI)	**-0.342**	**11**	**0.304**
Catheter-associated urinary tract infection (CAUTI)	**-0.219**	**15**	**0.433**
Surgical Site Infection from colon surgery	**-0.170**	**11**	**0.617**
Methicillin-resistant Staphylococcus aureus (MRSA) bacteremia	**0.565**	**10**	**0.089**
Clostridium difficile (C. diff) infection	**0.152**	**17**	**0.559**

Note: Data on Hospital Compare measures came from the January 2022 version of Hospital Compare. Measure reporting period for the Hospital Compare measures was between 10/1/19 and 3/31/21. Reporting period for HH PI was between 1/1/2020 and 12/31/2020.

Similarly, **Table 17** below shows the correlation between hospital sites' PI scores and 12 quality measures reflecting patients' perspectives of hospital care.

Table 17. Convergent Validity Between HH PI and HCAHPS Measures by Spearman's Rank Correlation

Hospital Compare Measures	Rho	Number of hospitals	Two-sided P-val
Nurse communication (extent to which nurses communicated well) - star rating	-0.378	18	0.122
Staff responsiveness (extent to which patients received help as soon as they needed) - star rating	-0.678	18	0.002
Communication about medicine (extent to which staff explained about medicines before giving it to them) - star rating	-0.277	18	0.265
Discharge information (if they were given information about what to do during their recovery at home) - star rating	-0.436	18	0.070
Care transition (extent to which they agree that they understood their care when leaving the hospital) - star rating	0.062	18	0.808
Overall rating of hospital (0 to 10) - star rating	-0.142	18	0.575
Nurse communication (extent to which nurses communicated well) – linear mean score	-0.446	18	0.063
Staff responsiveness (extent to which patients received help as soon as they needed) - linear mean score	-0.649	18	0.004
Communication about medicine (extent to which staff explained about medicines before giving it to them) - linear mean score	-0.354	18	0.150
Discharge information (if they were given information about what to do during their recovery at home) - linear mean score	-0.442	18	0.066
Care transition (extent to which they agree that they understood their care when leaving the hospital) - linear mean score	-0.050	18	0.843
Overall rating of hospital (0 to 10) - linear mean score	-0.054	18	0.832

Note: HCAPHS - Hospital Consumer Assessment of Healthcare Providers and Systems. Data on hospital compare measures came from the January 2022 version of Hospital Compare. Measure reporting period for the hospital compare measures was between 7/1/20 and 3/31/21. Reporting period for HH PI was between 1/1/2020 and 12/31/2020.

Notes:

Not every test site provided information for all measures in the Hospital Compare data.

[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]

Data Element Validity

Testing results clearly show that all measure's data elements are consistently stored in the EHR and can be accurately exported for calculation. For example, the frequency of data missingness is virtually zero across test sites. Similarly, the

percentage of agreement in measure concepts between data sources is close to perfect. Results also show robust PPVs for every measure component, with the minimum PPV equal to 0.97 and a great majority of PPVs equal to 1.0. We also see near perfect sensitivity, NPV, and specificity across measure components and sites. Using measure numerator as the example, the minimum sensitivity of 0.98 suggests that for every 100 patients that truly developed a HAPI, the measure correctly classified 98 of them based on structured data in the EHR. The minimum NPV of 0.99 suggests that for every 100 patients deemed not to have developed a HAPI according to the EHR data, 99 of them were confirmed as true numerator negatives.

Overall, results offer clear evidence that HH PI, as currently specified, can detect true HAPIs with high precision and that the measure will have very low false positives in implementation.

Measure Score Validity

Construct validity for HH PI at the hospital level is moderate, and correlational directions are largely in line with our expectation. For example, the rate of PI correlates with several independently collected and NQF-endorsed measures of hospital harms, such as healthcare associated MRSA bloodstream and intestinal infections. Higher PI rates are inversely related to patients' experiences with hospital care, particularly nursing components such as nurse communications, staff responsiveness, and discharge information. Several of the strong correlations are consistent with findings in the literature on nursing care and supportive of measure score level validity. These results should be interpreted cautiously, however, as only 18 hospitals were included in the analysis. Further, not every site has complete measure information in Hospital Compare and the measure reporting period between our measure and the related measures is not fully aligned, with only six-month overlap.

[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]

Using the full denominator data, we first calculated the hospital-level measure performance rate and its 95% confidence interval for each of the 18 test sites. We then calculated the system-wide, weighted average measure performance rate across sites. Third, we compared each test site's performance in PI against the system-wide average and gauge if its performance deviates significantly from the weighted mean. We also estimated a linear regression model, relating the incidence of PI to a set of hospital-specific indicators (or hospital-specific fixed effects) with a generalized T-test.

[Response Ends]

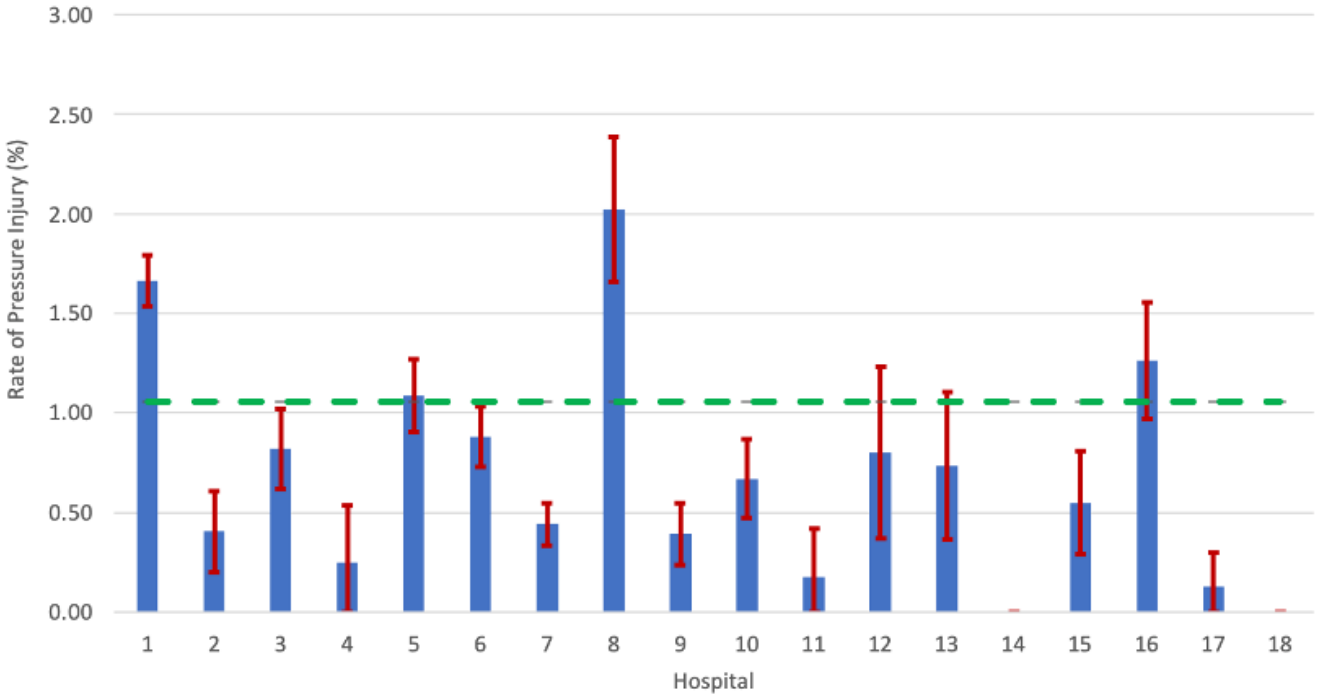
2b.06. Describe the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities.

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]

Exhibit 3 plots the distribution of observed hospital performance rate in PI and its 95% confidence interval. It further shows the system-wide, weighted average measure performance rate across the 18 test sites. Testing results clearly show that several hospitals' performance in PI are consistently below the system-wide, weighted average while a few are above that mean.

Exhibit 3. Distribution of Measure Performance Rate Across Test Sites; Data in CY2020



Notes: Vertical I-beam bars indicate the 95% confidence intervals, and the green dashed horizontal line indicates the system-wide, weighted average. Data from CY2020.

Table 18 shows the regression coefficients and their cluster-robust 95% confidence intervals. Regression results, complementing the histogram above, indicate that there exists noticeable and meaningful differences in PI rate across hospitals.

Table 18. Regression Coefficients and 95% Confidence Intervals

*	Coefficient	95% Confidence Interval
Test site 1	0.013***	[0.013,0.013]
Test site 2 (ref grp)	*	*
Test site 3	0.004***	[0.004,0.004]
Test site 4	-0.002***	[-0.002,-0.002]
Test site 5	0.007***	[0.007,0.007]
Test site 6	0.005***	[0.005,0.005]
Test site 7	0.000***	[0.000,0.000]
Test site 8	0.016***	[0.016,0.016]
Test site 9	-0.000***	[-0.000,-0.000]
Test site 10	0.003***	[0.003,0.003]
Test site 11	-0.002***	[-0.002,-0.002]
Test site 12	0.004***	[0.004,0.004]
Test site 13	0.003***	[0.003,0.003]
Test site 14	-0.004***	[-0.004,-0.004]

*	Coefficient	95% Confidence Interval
Test site 15	0.001***	[0.001,0.001]
Test site 16	0.009***	[0.009,0.009]
Test site 17	-0.003***	[-0.003,-0.003]
Test site 18	-0.004***	[-0.004,-0.004]
Observations	128,323	*

Notes: Regressions used all data points from the measure denominator population. Ref grp = reference group. Standard errors clustered at the level of hospital. *** $p < 0.01$.

*Cells intentionally left empty.

[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]

Testing data show that measure performance rates ranged from 0 to 2.02% (for every 100 qualified hospital admissions there are 44 inpatient encounters where patients suffered PI), indicating ample room for quality improvement in inpatient setting. Several hospitals' performance rates are consistently below the system-wide average while a few others are above that mean. Regression results, complementing the histogram, demonstrate that the measure can detect clinically meaningful differences in PI across hospitals.

[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]

We assessed the magnitude of data missingness during the process of parallel-form comparison and have included testing results in section 2b.03. As discussed in section 2b.02, samples that went into the abstraction were based on the approach of random sampling without replacement. When drawing the samples, we maintained constant the distribution of patient characteristics in the full measure population to the extent possible. Please refer to section 2b.02 for a detailed sampling method and the number of cases sampled from test hospitals.

During abstraction, we compared data exported from the EHR (eData) to data manually abstracted from patients' medical charts (mData) for every patient included in the abstraction sample. Given that information in patients' medical charts is typically deemed the "gold standard," this process helped us to identify the extent and distribution of missing data and assess in what direction performance results could be biased if data missing is systematic. We tabulated the frequency of data missingness (see **Tables 19 to 22** in section 2b.09) and calculated the percent agreement in data element validity (see **Tables 11 to 14** in section 2b.03 Data Element Validity).

[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]

Table 19. Data Element Missingness for Critical Data Elements; Measure Initial Population (Epic Site)

*	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)
Data Element	**Cases per EHR**	**Cases per Abstraction**	**Freq. of Missingness**
Patient had an inpatient encounter with discharge date between 1/1/20 and 12/31/20 (measurement period)	**155**	**155**	**0**
Patient aged 18 or older at the start of encounter	**155**	**155**	**0**
Patient had a stage 2-4 or unstageable pressure injury per the clinical exam after 24 hours of encounter start	**57**	**56**	**0**
Patient had a stage 2-4 or unstageable pressure injury ICD-10-CM indicating not present-on-admission	**7**	**7**	**0**
Patient had a deep tissue injury per the clinical examination after 72 hours of encounter start	**10**	**11**	**1**
Patient had a deep tissue pressure injury ICD-10-CM indicating not present-on-admission	**3**	**3**	**0**

*Cells intentionally left empty.

Table 20. Data Element Missingness for Critical Data Elements; Measure Initial Population (Cerner Sites)

*	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)
Data Element	**Cases per EHR**	**Cases per Abstraction**	**Freq. of Missingness**
Patient had an inpatient encounter with discharge date between 1/1/20 and 12/31/20 (measurement period)	**155**	**155**	**0**
Patient aged 18 or older at the start of encounter	**155**	**155**	**0**
Patient had a stage 2-4 or unstageable pressure injury per the clinical exam after 24 hours of encounter start	**53**	**54**	**1**

*	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)
Patient had a stage 2-4 or unstageable pressure injury ICD-10-CM indicating not present-on-admission	1	1	0
Patient had a deep tissue injury per the clinical examination after 72 hours of encounter start	11	12	1
Patient had a deep tissue pressure injury ICD-10-CM indicating not present-on-admission	2	2	0

*Cells intentionally left empty.

Table 21. Data Element Missingness for Critical Data Elements; Measure Denominator Exclusion (Epic Site)

*	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)	Site 1 (EHR: Epic)
Data Element	Cases per EHR	Cases per Abstraction	Freq. of Missingness
Patient had a stage 2-4 or unstageable pressure injury per clinical exam within 24 hours of encounter start	24	24	0
Patient had a stage 2-4 or unstageable pressure injury ICD-10-CM indicating present-on-admission	15	15	0
Patient had a deep tissue pressure injury per clinical examination within 72 hours of encounter start	1	1	0
Patient had a deep tissue pressure injury ICD-10-CM indicating present-on-admission	1	1	0
Patient had a ICD-10-CM indicating COVID-19 infection	19	19	0

*Cells intentionally left empty.

Table 22. Data Element Missingness for Critical Data Elements; Measure Denominator Exclusion (Cerner Sites)

*	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)
Data Element	Cases per EHR	Cases per Abstraction	Freq. of Missingness
Patient had a stage 2-4 or unstageable pressure injury per clinical exam within 24 hours of encounter start	5	6	1
Patient had a stage 2-4 or unstageable pressure injury ICD-10-CM indicating present-on-admission	5	5	0
Patient had a deep tissue pressure injury per clinical examination within 72 hours of encounter start	3	2	0

*	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)	Sites 2-18 (EHR: Cerner)
Patient had a deep tissue pressure injury ICD-10-CM indicating present-on-admission	**1**	**1**	**0**
Patient had a ICD-10-CM indicating COVID-19 infection	**21**	**21**	**0**

*Cells intentionally left empty.

[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]

Testing results clearly show that all measure's critical data elements are consistently stored in the EHR and can be accurately exported for calculation. For example, the frequency of data missingness is virtually zero across test sites.

[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]

Yes, the measure uses 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]

We used two methods for testing the empirical impact of measure denominator exclusions. First, using the full denominator data, we removed measure exclusion criterion one at a time from the logic and calculated the marginal and relative (%) impact on the prevalences of numerator and denominator, as well as the observed measure rate as a result. Second, through parallel-form comparison, we evaluated whether patients excluded from the denominator per the EHR truly met the clinical intent for exclusion.

The first method allowed us to gauge the marginal change in measure performance rate due to a particular exclusion criterion. It further helps us to understand which criterion had the largest bite in terms of preventing false positives. The second method, being part of the measure data element validity testing, supplied additional evidence for the likelihood of measure suffering false negatives.

[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]

Table 23 below shows the testing result for the first method.

Table 23. Impact of Denominator Exclusion Criteria on Denominator Count, Numerator Count, and Measure Rate; All Sites Combined

Pressure Injury	Denominator		Numerator		Rate per 100	
*	Count (N)	% change	Count (N)	% change	Dropping Exclsn	% change
Current specification	**128,323**		**1,355**		**1.06**	
Relax: PI Stage 2 - 4 or Unstageable Clinical Examination Within 24hrs of Encounter Start	**129,437**	**0.87**	**1,735**	**28.04**	**1.34**	**26.94**
Relax: PI Stage 2 - 4 or Unstageable Dx POA	**129,251**	**0.72**	**1,821**	**34.39**	**1.41**	**33.43**
Relax: PI Deep Tissue Injury Clinical Examination Within 72hrs of Encounter Start	**128,807**	**0.38**	**1,538**	**13.51**	**1.19**	**13.08**
Relax: PI Deep Tissue Injury Dx POA	**128,404**	**0.06**	**1,395**	**2.95**	**1.09**	**2.89**
Relax: COVID-19 Dx	**137,484**	**7.14**	**1,724**	**27.23**	**1.25**	**18.75**

Note: Dx – Diagnosis. Exclsn – exclusion. Hrs – hours. POA – Present-on-admission. PI – Pressure injury.

*Cells intentionally left empty.

Please refer to section 2b.03 for the test results for the second method.

[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]

cross five exclusion categories and 18 test sites, average impacts on the denominator, numerator, and measure rate ranged from no more than 1% to more than 7%, from less than 3% to over 30%, and from less than 3% to 33%, respectively.

1. **PI Stage 2 - 4 or Unstageable Clinical Examination Within 24hrs of Encounter Start**. This exclusion determines if patients have PI at the start of care using a time threshold suggested by the National Pressure Injury Advisory Panel (NPIAP) Clinical Practice Guidelines.[1] **Table 23** above indicates that removal increased the measure denominator by slightly less than 1% but the numerator by 28%. The resulting measure rate thus increased considerably after dropping this exclusion.
2. **PI Stage 2 - 4 or Unstageable Dx POA**. Differing from the above, this criterion leverages diagnosis information and its POA status to ascertain if the observed PI was present at the start of care and hence should not be

deemed as hospital acquired. **Table 23** above shows that removal increased the measure denominator by 0.7% across sites but increased the numerator by 34%. Clinical documentation is often the basis for adding relevant PI diagnosis codes to patients' medical records; therefore, it is not surprising that the relative impact of PI diagnoses is comparable to that of PI clinical documentation.

3. **Deep Tissue Injury Clinical Examination Within 72hrs of Encounter Start**. This criterion uses the suggested time threshold per NPIAP's Clinical Practice Guidelines to determine if DTI was present at the start of care. **Table 23** above indicates that, across the 18 test sites, removing this criterion from the measure logic increased the measure denominator by less than 0.1% but the numerator by 13%, leading to higher measure performance rates.
4. **PI Deep Tissue Injury Dx POA.** This criterion uses the structured diagnosis information and its POA status to determine if patients had DTI at the start of care. **Table 23** above indicates that removal increased the measure denominator across sites minimally and the numerator by roughly 3%. The relative impact of removing the DTI Dx POA exclusion appears to be much smaller than that of removing the DTI clinical examination exclusion, because DTI is not consistently documented in physician notes, which serve as the basis for professional coders to determine the POA status.
5. **COVID-19 Dx**. **Table 23** above shows that across test sites, 27% more encounters would have been flagged as developing a HAPI without this exclusion.

Overall, all exclusions (either clinical documentation or code-based) are necessary to reduce the measure's false positive rate and to prevent hospitals from being penalized by appropriate management of pre-existing or comorbid conditions, such as COVID-19. The COVID-19 exclusion is viewed as a temporary exclusion, given that COVID-19 patients in 2020 were treated by continuous prone positioning, with turning discouraged or prohibited due to its negative impact on oxygenation. In addition, COVID-related skin lesions in 2020 were not uniformly recognized and distinguished from pressure injuries. With improved antiviral therapies and greater natural and vaccine-induced immunity, continuous prone positioning (without turning) may no longer be necessary in 2022, and this exclusion will be reconsidered before the next endorsement cycle. None of the exclusions imposes a burden on providers by increasing the complexity of data collection or analysis, since all data exist in the EHR in structured fields.

References

1. *Guidelines - National Pressure Ulcer Advisory Panel*. (n.d.). Retrieved May 10, 2022, from https://npiap.com/page/Guidelines

[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]

There are certain characteristics including the patient's age, reason for hospitalization, clinical status on arrival to the hospital or comorbid conditions that may influence the risk of harm during a hospitalization. However, many harms should be avoidable, regardless of patient risk. The following factors are taken into consideration when determining whether risk adjustment is warranted: 1) if many patients could be at risk of the harm regardless of their age, clinical status, comorbidities, or reason for admission, 2) if the majority of occurrences of harm are linked to care that is within the control of providers, and 3) if there is evidence that the risk of harm can be largely ameliorated by best practices regardless of patients' inherent risk profile.

New-onset pressure injuries of stage 2 or greater are widely considered to be potentially avoidable with best practices. For example, the National Pressure Ulcer (now Injury) Advisory Panel convened a consensus conference in 2011 involving 24 stakeholder organizations. Unanimous consensus was achieved for the following statements: "most pressure ulcers are avoidable; not all pressure ulcers are avoidable; there are situations that render pressure ulcer development unavoidable, including hemodynamic instability that is worsened with physical movement and inability to maintain nutrition and hydration status and the presence of an advanced directive prohibiting artificial nutrition/hydration… and (even) if enough pressure was removed from the external body the skin cannot always survive." A subsequent conference on the same topic in 2014, informed by extensive review of the relevant clinical literature, concluded that "in the vast majority of cases, appropriate identification and mitigation of risk factors can prevent or minimize PU formation. However, in some cases, PUs are unavoidable because the magnitude and severity of risk are overwhelmingly high or preventive measures are either contraindicated or inadequate, given the magnitude and severity of risk." Unfortunately, a subsequent case control study of 475 participants was unable to distinguish avoidable from unavoidable pressure injuries using any set of patient risk factors or other characteristics, suggesting little benefit to risk-adjustment.[1, 2, 3]

Although certain patients may be particularly vulnerable to pressure injuries in certain settings (e.g., permanent or prolonged immobility), the most common causes are limited mobility during an acute illness, combined with friction or shear against sensitive skin. There are many actions hospitals can take to reduce patient harm risk, such as conducting a structured risk assessment to identify individuals at risk for pressure injury as soon as possible upon arrival and repeating at regular intervals, as well as proper skin care, nutrition, and careful repositioning of patients. As many of the causes can be mitigated through best care in hospital environments, our research indicates that risk adjustment is not currently warranted for this measure (after the denominator exclusions). We will continue to evaluate the appropriateness of risk adjustment in measure reevaluation.

References

[1]Pittman J, Beeson T, Dillon J, Yang Z, Mravec M, Malloy C, Cuddigan J. Hospital-Acquired Pressure Injuries and Acute Skin Failure in Critical Care: A Case-Control Study. J Wound Ostomy Continence Nurs. 2021 Jan-Feb 01;48(1):20-30. doi: 10.1097/WON.0000000000000734. PMID: 33427806.

[2] European Pressure Ulcer Advisory Panel, National Pressure Injury Advisory Panel and Pan Pacific Pressure Injury Alliance. Prevention and Treatment of Pressure Ulcers/Injuries: Clinical Practice Guideline. (2019) The International Guideline. Emily Haesler (Ed.).

[3] Wound Healing Society (2015): https://pubmed.ncbi.nlm.nih.gov/26683529/

[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]

N/A

[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]

Criterion 3. Feasibility

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

[Response Begins]

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score)

Coded by someone other than person obtaining original information (e.g., DRG, ICD-10 codes on claims)

[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]

ALL data elements are in defined fields in electronic health records (EHRs)

[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]

Current (2022) Submission

N/A. This is an eCQM that uses all data elements from defined fields in the electronic health record (EHR).

[Response Ends]

3.05. Complete and attach the NQF Feasibility Score Card.

[Response Begins]

Please see the attached NQF Feasibility Score Card for the measure.

[Response Ends]

Attachment: 3498e_3498e_PI NQF feasibility scorecard_vFinal External_For NQF-508.xlsx

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]

Current (2022) Submission:

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

Twenty hospitals participated in the evaluation of feasibility—three Epic and 17 Cerner users. All hospital sites confirmed that all data elements used in the measure are captured within the EHR in a structured and codified manner either using nationally accepted terminology standards or local system codes that could be easily mapped. The feasibility scorecard shows the site-level scores across all sites for each domain, which demonstrated high feasibility.

However, there were 2 Epic hospitals that opted to not participate in beta testing due to their documentation practices to support pressure injury staging. While these two sites were able to capture pressure injuries in a structured field, the final staging assessments were documented as free text clinical notes by the wound care specialist. Although, a workflow modification for these sites would enable more accurate capture of staging documentation, we compensate for this scenario in the measure specification through the use of diagnosis codes that incorporate the final staging information. Despite this workflow challenge, there are no concerns with the feasibility to capture the required data to support eCQM implementation.

[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]

Current (2022) Submission:

There are no fees associated with the use of this eCQM. Value sets are housed in the Value Set Authority Center (VSAC), which is provided by the National Library of Medicine (NLM), in coordination with the Office of the National Coordinator for Health Information Technology and the Centers for Medicare & Medicaid Services.

Viewing or downloading value sets requires a free Unified Medical Language System® (UMLS) Metathesaurus License, due to usage restrictions on some of the codes included in the value sets.

Individuals interested in accessing value set content can request a UMLS license at (https://www.nlm.nih.gov/databases/umls.html).

[Response Ends]

Criterion 4: Use and Usability

4a. Use

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]

Not in use

[Not in use Please Explain]

This eCQM is under initial endorsement review and is not currently used in any accountability program. This measure was submitted to the 2022 Measures Under Consideration (MUC) List and will be reviewed by the Measure Applications Partnership (MAP) during their 2022-2023 review cycle. CMS has sought MAP support for implementation in an accountability program (Hospital Inpatient Quality Reporting and Promoting Interoperability Programs) pending feedback received from the MAP, during NQF endorsement, and rulemaking.

[Response Ends]

4a.02. Check all planned uses.

[Response Begins]

Public reporting

Payment Program

[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]

Current (2022) Submission

This eCQM is under initial endorsement review and is not currently used in any accountability program. In December 2018, this eCQM was presented to the Measure Applications Partnership (MAP), who noted conditional support for rulemaking. This measure was subsequently reviewed by NQF during the Spring 2019 cycle, but withdrawn due to anticipated substantive changes (e.g., expanded value set that would improve capture of pressure injuries, incorporation of present on admission indicator for ICD-10-CM diagnoses, and denominator exclusion for pressure injuries present on admission). CMS subsequently made those substantive updates and re-tested the measure.

This measure was submitted to the 2022 Measures Under Consideration (MUC) List and will be reviewed by the Measure Applications Partnership (MAP) during their 2022-2023 review cycle. CMS has sought MAP support for implementation in an accountability program (Hospital Inpatient Quality Reporting and Promoting Interoperability Programs) pending feedback received from the MAP, during NQF endorsement, and rulemaking.

Previous (2019) Submission

N/A; this eCQM is under initial endorsement review and is not currently used in any accountability program. In December 2018, this eCQM was presented to the Measure Applications Partnership (MAP), who noted conditional support for rulemaking. Thus, CMS is considering implementation in an accountability program pending feedback received during NQF endorsement and rulemaking.

[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]

Current (2022) Submission

CMS is seeking MAP's recommendations and support for implementation in the Inpatient Quality Reporting and Promoting Interoperability for eligible hospitals programs.

Previous (2019) Submission

Following MAP's recommendations and support, we envision that this measure will be considered for accountability programs via future rulemaking.

[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]

Current (2022) Submission

N/A; this measure is being submitted as de novo and has not yet been implemented.

For eCQMs included in CMS reporting programs, implementation resources are provided through the CMS eCQI Resource Center and The ONC Project Tracking System (a collaboration platform hosted by the HHS's Office of National Coordinator for Health Information Technology (ONC) that provides users with a common place to transparently log, track, and discuss and clarify issues with eCQM implementation and logic interpretation). As part of the measure rollout, CMS (in collaboration with The Joint Commission) also provides an annual webinar series for measured entities to review the measure specification, logic, and answer implementation questions.

Previous (2019) Submission

N/A; this measure is being submitted as de novo as has not yet been implemented. Implementation is planned pending finalization of the NQF and CMS rulemaking processes.

[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]

Current (2022) Submission

N/A; this measure is being submitted as de novo and has not yet been implemented. Implementation is planned pending finalization of the NQF and CMS rulemaking processes.

Previous (2019) Submission

N/A; this measure is being submitted as de novo as has not yet been implemented. Implementation is planned pending finalization of the NQF and CMS rulemaking processes.

[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]

Current (2022) Submission

N/A; this measure is being submitted as de novo and has not yet been implemented.

CMS obtains feedback on all of its measures through various avenues including: (1) Measures Management System (MMS) posting with announcements to stakeholders, (2) NQF endorsement review, (3) Measures Application Partnership (MAP) review, (4) Proposed Rules published in the Federal Register, (5) ongoing feedback from the user community through the QualityNet portal, (6) ongoing review by a Technical Advisory Panel representing key stakeholders and clinical experts, which will continue to support the measure.

Additionally, for eCQMs included in CMS reporting programs, implementation resources are provided through the CMS eCQI Resource Center and The ONC Project Tracking System (a collaboration platform hosted by the HHS's Office of National Coordinator for Health Information Technology (ONC) that provides users with a common place to transparently log, track, and discuss and clarify issues with eCQM implementation). These implementation feedback are evaluated and, as appropriate, presented during the CMS Annual Update Change Review Process for measure refinements.

Previous (2019) Submission

N/A; this measure is being submitted as de novo as has not yet been implemented. Implementation is planned pending finalization of the NQF and CMS rulemaking processes.

[Response Ends]

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

[Response Begins]

Current (2022) Submission

N/A; this measure is being submitted as de novo and has not yet been implemented. Implementation is planned pending finalization of the NQF and CMS rulemaking processes.

However, for eCQMs included in CMS reporting programs, there are measure feedback loops provided through the CMS eCQI Resource Center and The ONC Project Tracking System (a collaboration platform hosted by the HHS's Office of National Coordinator for Health Information Technology (ONC) that provides users with a common place to transparently log, track, and discuss and clarify issues with eCQM implementation). Additionally, eCQMs go through an Annual Update Cycle, which includes the Change Review Process (a mechanism for public comment and suggested measure refinements).

Previous (2019) Submission

N/A; this measure is being submitted as de novo as has not yet been implemented. Implementation is planned pending finalization of the NQF and CMS rulemaking processes.

[Response Ends]

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

[Response Begins]

Current (2022) Submission

While this measure does not have usability information from measured entities, as it is being developed de novo and has not been implemented yet, our team sought input from multiple stakeholder groups throughout the measure development process. We believe in a transparent measure development process, and highly value the feedback received on the measure. During development, a technical expert panel (TEP) composed of a variety of stakeholders was engaged at various stages of development to obtain balanced, expert input. In addition to the TEP, feedback to evaluate the measure specifications was collected through rulemaking (FY2020 IPPS Proposed Rule inviting public comment on potential future inclusion of this eCQM in the Hospital Inpatient Quality Reporting and Promoting Interoperability Programs), during the Spring 2019 NQF 16-week public comment cycle, and posting to the Electronic Clinical Quality

Improvement (eCQI) Resource Center Collaboration Workspace. We also collected feedback from pilot sites following measure testing, using a post-pilot survey, to assess the measure's usability and its prospect of field implementation.

Previous (2019) Submission

While this measure does not have usability information from measured entities, as it is being developed de novo and has not been implemented yet, our team sought input from multiple stakeholder groups throughout the measure development process. We believe in a transparent measure development process, and highly value the feedback received on the measure. During development, a technical expert panel composed of a variety of stakeholders was engaged at various stages of development to obtain balanced, expert input. We also solicited and received feedback on the measure through an MMS Blueprint 44-day Public Input Period during development.

[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]

Current (2022) Submission

As noted above, input received from TEP members was instrumental to the development and specification of this measure. Feedback received during public comment (Rulemaking, NQF, and Measure Collaboration Workspace) was also explored during the measure testing process. For example, commenters indicated that the measure would be useful in improving quality of care with some changes to the measure definitions to capture missing key elements for reliable and valid capture of PI (Pressure Injuries). Since then, we have made changes to incorporate the latest research findings to produce a more accurate numerator using clinical data and the present on admission indicator, and to exclude encounters where a PI was found on admission. Following measure refinements and testing, expert stakeholders commented that the changes improved the applicability and usability of the measure to accurately identify PI as intended.

Previous (2019) Submission

As noted above, input received from TEP members was instrumental to the development and specification of this measure. Feedback received during public comment was also explored during the measure testing process.

[Response Ends]

4b. Usability

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]

Current (2022) Submission

This is a new eCQM and there is no time trend information available regarding facility performance improvement. This eCQM is not currently used in any quality improvement program, but a primary goal of the eCQM is to provide hospitals with performance information necessary to implement focused quality improvement efforts.

Previous (2019) Submission

This is a new eCQM and there is no time trend information available regarding facility performance improvement. This eCQM is not currently used in any quality improvement program, but a primary goal of the eCQM is to provide hospitals with performance information necessary to implement focused quality improvement efforts.

[Response Ends]

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

[Response Begins]

Current (2022) Submission

We did not identify any unintended consequences during eCQM development or testing. However, CMS is committed to monitoring this eCQM's use and assessing potential unintended consequences over time, such as the inappropriate shifting of care, and other negative unintended consequences for patients.

Previous (2019) Submission

We did not identify any unintended consequences during eCQM development or testing. However, CMS is committed to monitoring this eCQM's use and assessing potential unintended consequences over time, such as the inappropriate shifting of care, and other negative unintended consequences for patients.

[Response Ends]

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

[Response Begins]

Current (2022) Submission

No unexpected benefits were noted during eCQM development testing.

Previous (2019) Submission

No unexpected benefits were noted during eCQM development testing.

[Response Ends]

Criterion 5: Related and Competing Measures

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]

[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]

Current (2022) Submission

Related measures that are not currently NQF endorsed include:

- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05737-C-Long Term Care Hospital-LTCHQR),
- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05737-C-Long Term Care Hospital-LTCHQR),
- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05737-C-Long Term Care Hospital-LTCHC),
- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05740-C-Inpatient Rehabilitation Facility-IRFQR),
- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05740-C-Inpatient Rehabilitation Facility-IRFC),
- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05741-C-Skilled Nursing Facility-NHC),
- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05741-C-Skilled Nursing Facility-SNFQRP),
- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05852-C-Home Health-HHC),
- Changes in Skin Integrity Post-Acute Care: Pressure Ulcer/Injury (Steward: Centers for Medicare and Medicaid Services; CMS#05852-C-Home Health-HHQR),
- Percent of Residents or Patients with Pressure Ulcers That Are New or Worsened (Short-Stay) (Centers for Medicare and Medicaid Services; CMS #04056-C-Skilled Nursing Facility-NHC; NQF #0678)
- Percent of Residents or Patients with Pressure Ulcers That Are New or Worsened (Short-Stay) (Centers for Medicare and Medicaid Services; CMS #04056-C-Skilled Nursing Facility-NHQI; NQF #0678)
- Pressure Ulcer Rate (Steward: Agency for Healthcare Research and Quality; PSI 03)

No competing measures that are not currently NQF-endorsed.

[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]

Current (2022) Submission:

Harmonization between our measure and NQF #0531 is not necessary because the measures are not related (i.e., they do not have the same measure focus or the same target population). NQF #0531 is a composite measure of 10 hospital-

acquired complications (Patient Safety and Adverse Events Composite), and the only component that overlaps with the proposed measure is PSI 03 (Pressure Ulcer Rate). PSI 03 is not an endorsed measure, and it has a narrower focus than the proposed measure because it does not include stage 2 pressure injuries in the outcome, has additional exclusions in the cohort definition, and uses ICD-10-CM codes via claims as a data source. The proposed PI measure includes stage 2 pressure injuries in the outcome and identifies pressure injuries using direct extraction of structured data from the EHR.

Although both NQF #0531 and the proposed PI measure have a target population of hospitalized adults, their specific denominators are quite different. NQF #0531 is the CMS claims-based version of PSI 90, so its denominator is limited to adult (fee-for service) Medicare beneficiaries, whereas the proposed measure applies to adults of all ages and payers.

Previous (2019) Submission:

While there are several measures that target the reduction of hospital-acquired pressure injuries in use in various patient populations, there are no eCQMs intended for use to compare quality across acute care hospitals. The measures NQF# 0679 and #0678 target a different patient population and use chart review data from the following sources: Minimum Data Set (MDS); Long Term Care Hospitals Continuity Assessment Record and Evaluation (LTCH-CARE) Data set; and the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI) Data set. Additionally, NQF# 0678 measure includes worsening pressure injuries and NQF# 0679's population consists of only high-risk patients defined as those who are impaired in bed mobility, comatose, or suffering malnutrition. The new Hospital Harm -Pressure Injury eCQM identifies pressure injuries using direct extraction of structured data from the EHR and will provide hospitals with reliable and timely measurement of their pressure injury rates. As these measures do not apply to the same measured entities, it should not impact data collection burden.

While there are several measures that target the reduction of hospital-acquired pressure injuries in use in various patient populations, there are no eCQMs intended for use to compare quality across acute care hospitals. The measures NQF# 0679 and #0678 target a different patient population and use chart review data from the following sources: Minimum Data Set (MDS); Long Term Care Hospitals Continuity Assessment Record and Evaluation (LTCH-CARE) Data set; and the Inpatient Rehabilitation Facility Patient Assessment Instrument (IRF-PAI) Data set. Additionally, NQF# 0678 measure includes worsening pressure injuries and NQF# 0679's population consists of only high-risk patients defined as those who are impaired in bed mobility, comatose, or suffering malnutrition. The new Hospital Harm -Pressure Injury eCQM identifies pressure injuries using direct extraction of structured data from the EHR and will provide hospitals with reliable and timely measurement of their pressure injury rates. As these measures do not apply to the same measured entities, it should not impact data collection burden.

[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]

Current (2022) Submission:

N/A as there are currently no competing measures.

Previous (2019) Submission:

Hospital-acquired pressure injuries are currently measured and publicly reported in the Hospital-Acquired Condition Reduction Program (HACRP) as a component of the Patient Safety Indicator (PSI) 90 measure (PSI-03). PSI-03 does not include stage 2 pressure injuries in the outcome, has additional exclusions to the cohort, and uses ICD codes via claims as a data source. Hospital Harm - Pressure Injury Measure is an eCQM (EHR data-only), which stakeholders and TEP have noted as a more desirable data source with more face validity for measuring pressure injuries. Hospital-acquired pressure

injuries are currently measured and publicly reported in the Hospital-Acquired Condition Reduction Program (HACRP) as a component of the Patient Safety Indicator (PSI) 90 measure (PSI-03). PSI-03 does not include stage 2 pressure injuries in the outcome, has additional exclusions to the cohort, and uses ICD codes via claims as a data source. Hospital Harm - Pressure Injury Measure is an eCQM (EHR data-only), which stakeholders and TEP have noted as a more desirable data source with more face validity for measuring pressure injuries.

[Response Ends]