NATIONAL QUALITY FORUM

MEASURE WORKSHEET

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

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

Purple text represents the responses from measure developers.

Red text denotes developer information that has changed since the last measure evaluation review.

Brief Measure Information

NQF #: 0383

Corresponding Measures:

De.2. Measure Title: Oncology: Medical and Radiation - Plan of Care for Pain

Co.1.1. Measure Steward: American Society of Clinical Oncology

De.3. Brief Description of Measure: Percentage of visits for patients, regardless of age, with a diagnosis of cancer currently receiving chemotherapy or radiation therapy who report having pain with a documented plan of care to address pain.

1b.1. Developer Rationale: Proper pain management is critical to achieving pain control. Pain has a severe impact on a patient's quality of life (1). Additionally, cancer pain is associated with numerous psychosocial responses (2-3). One third of patients describe cancer pain as intolerable aspect of cancer (4). Adequate pain treatment results in clinically relevant improvement in health-related quality of life (5). This is reflected in the most recent NCCN guidelines which stated that unrelieved pain denies [patients] comfort and greatly affects their activities, motivation, interactions with family and friends, and overall quality of life (6). Moreover, the importance of assessing pain in cancer patients is included in European guidelines, which go as far to say that despite published guidelines and education programs on the assessment and treatment of cancer related pain, unrelieved pain continues to be a substantial concern in patients worldwide (7). Given that it is projected that there will be over 15 million cancer patients in 2020 worldwide, this only increased the importance of addressing address patient pain (8).

This measure aims to improve attention to pain management and requires a plan of care for cancer patients be documented who report having pain to allow for individualized treatment based on clinical circumstances and patient wishes and focuses on early documentation of a pain plan of care.

Citations:

1. IASP. 2008-2009 Global Year Against Cancer Pain 2008. Available at: https://www.iasp-pain.org/GlobalYear/Cancer Pain. Accessed February 10, 2015.

2. Kroenke K, Theobald D, Wu J, et al. The association of depression and pain with health-related quality of life, disability, and health care use in cancer patients. J Pain Symptom Manage 2010;40:327e341.

3. Porter LS, Keefe FJ. Psychosocial issues in cancer pain. Curr Pain Headache Rep 2011;15:263e270.

4. Breivik H, Cherny N, Collett B, et al. Cancer-related pain: a pan-European survey of prevalence, treatment, and patient attitudes. Ann Oncol 2009;20:1420e1433.

5. Puetzler J, Feldmann RE Jr, Brascher AK, Gerhardt A, Benrath J. Improvements in health-related quality of life by comprehensive cancer pain therapy: a pilot study with breast cancer outpatients under palliative chemotherapy. Oncol Res Treat 2014;37:456e462.

6. National Comprehensive Cancer Network (NCCN). Clinical Practice Guidelines in Oncology: Adult Cancer Pain. Version 2, 2017. Available at: http://www.nccn.org.

7. Management of Cancer Pain: ESMO Clinical Practice Guidelines. C. I. Ripamonti, D. Santini, E. Maranzano, M. Berti, F. Roila . Ann Oncol 2012; 23 (Suppl 7): vii39-vii154.

8. Frankish H. 15 million new cancer cases per year by 2020, says WHO. Lancet 2003; 361: 1278.

S.4. Numerator Statement: Patient visits that include a documented plan of care* to address pain.

*A documented plan of care may include: use of non-opioid analgesics, opioids, psychological support, patient and/or family education, referral to a pain clinic, or reassessment of pain at an appropriate time interval.

S.6. Denominator Statement: All visits for patients, regardless of age, with a diagnosis of cancer currently receiving chemotherapy or radiation therapy who report having pain

S.8. Denominator Exclusions: None

De.1. Measure Type: Process

S.17. Data Source: Paper Medical Records, Registry Data

S.20. Level of Analysis: Clinician : Group/Practice

IF Endorsement Maintenance – Original Endorsement Date: Jul 31, 2008 **Most Recent Endorsement Date:** Aug 09, 2012

IF this measure is included in a composite, NQF Composite#/title:

IF this measure is paired/grouped, NQF#/title:

2100:Paired Measure 0383 and 0384

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? 2100:Paired Measure 0383 and 0384

This measure is paired with NQF #0384 Oncology: Medical and Radiation - Pain Intensity Quantified, which assesses whether there is documentation of a clinical assessment for the presence or absence of pain using a standardized tool. These measures together represent a stepwise approach to attenuating pain that commonly results from cancer therapy. This measure requires the initial and ongoing assessment and quantification of pain which are required to formulate the most appropriate plan with the intent of improving patient outcomes.

Preliminary Analysis: Maintenance of Endorsement

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

Criteria 1: Importance to Measure and Report

1a. Evidence

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

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

The developer provides the following evidence for this measure:

- **Systematic Review of the evidence specific to this measure?** ☐ **Yes** ☒ **No**
- **Quality, Quantity and Consistency of evidence provided?** ☐ **Yes** ☒ **No**
- **Evidence graded?** ☐ **Yes** ☒ **No**

Summary of prior review in 2012:

- In 2012, the Committee was concerned that including any report of pain, even mild, may dilute the impact of this measure. However, the Steering Committee stated that simply noting that the patient was experiencing mild pain and the need to follow up on it would be sufficient to meet this measure, alleviating concerns.

Changes to evidence from last review

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

☒ **The developer provided updated evidence for this measure:**

Updates:

- The NCCN Clinical Practice Guidelines in Oncology, Adult Cancer Pain, includes management of pain in both opidod-naïve and opioid tolerant patients. [**Level of Evidence: Category 2A**: Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate.]
    - The NCCN guideline does not include an overview of the body of evidence used for the recommendations specific to the overall management of pain. However, the guideline does provide an in-depth discussion on the evidence, benefits and harms of specific therapies and interventions (e.g., aspirin, opioids, strategies for specific cancer pain syndromes, non-pharmacologic).

Guidance from the Evidence Algorithm

Process measure about documentation submitted with evidence about treatment and management of pain (something other than what is being measured) (Box 3) → Clinical practice guidelines submitted do not support measure focus (Box 7) → Are there, OR could there be, performance measures of a related health outcome, OR evidence-based intermediate clinical outcome or process? (Box 10) → Yes → No exception → Insufficient

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

RATIONALE: Process measure about *documentation* submitted with evidence about *treatment and management of pain*

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

Maintenance measures – increased emphasis on gap and variation

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

- The developer provided data from 2015-2017 for this measure from the CMS Physician Quality Reporting System (PQRS) and the Merit-based Incentive Payment System (MIPS):
    - 2015: 83.43%; 88 eligible professionals
    - 2016: 89.11%; 106 eligible professionals
    - 2017: 75.24%; 244 eligible professionals

Disparities

- According to the developer, although this measure is included in the MIPS program, this program has not yet made disparities data available to analyze and report.
- The developer provided a summary of data from literature to demonstrate that patients with cancer continue to receive disparate treatment for pain.

Questions for the Committee:

- Do the performance rates from CMS PQRS and MIPS demonstrate an opportunity for further improvement in the documentation of a pain care plan in the first two office visits for cancer patients with moderate or severe pain receiving chemotherapy or radiation?
- Do the performance rates warrant a national performance measure?
- Are you aware of evidence that disparities exist related to documenting a pain plan of care in the first two office visits?

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

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

1a. Evidence to Support Measure Focus: For all measures (structure, process, outcome, patient-reported structure/process), empirical data are required. How does the evidence relate to the specific structure, process, or outcome being measured? Does it apply directly or is it tangential? How does the structure, process, or outcome relate to desired outcomes? For maintenance measures –are you aware of any new studies/information that changes the evidence base for this measure that has not been cited in the submission? For measures derived from a patient report: Measures derived from a patient report must demonstrate that the target population values the measured outcome, process, or structure."

- This is an extremely important measure for patients. We had a long discussion about the value of merely quantifying pain, but this measure goes a step beyond and measures whether there is a plan of care in place to address pain. This is important to patients.
- No new evidence
- The measure was intended to make certain that any patient in pain had pain addressed. more recent evidence discusses the risk/benefit ratio of pain regimens. low level of evidence supports intervention and the measure.
- Process measure that is a necessary first step towards improvemenet
- This is a process measure that assesses the clinician intervention for reported pain. This measure can improve the awareness of clinician to pain and help facilitate interventions for addressing pain.

Problem is that most providers want to prescribe pain meds and that many other interventions (yoga, meditation, acupuncture) is not covered by insurance

- Outcomes appear to be consistant
- Evidence does not match the metric which is being evaluated on plan creation. There is no evidence shown to support the different pain levels and plan interventions documented
- This application includes updated information concerning the status of medical evidence. This measure is a guidelines-driven measure that is based on NCCN practice guidelines and there is level 2A evidence for this recommendation which is based on lower level evidence but uniform consensus among the expert panel that the intervention is appropriate. The guideline includes an in-depth discussion on evidence, benefits, and risks of specific interventions. Although the level of evidence is rated as insufficient, I believe that expert consensus supports the use of this measure to assess performance. I also believe that this is an important measure to patients with respect to quality of life.
- Insufficient, evidence does not directly support measure
- Yes - this is important to patients
- The developer provided more evidence but it does not seem to provide significant additional information.

1b. Performance Gap: Was current performance data on the measure provided? How does it demonstrate a gap in care (variability or overall less than optimal performance) to warrant a national performance measure? Disparities: Was data on the measure by population subgroups provided? How does it demonstrate disparities in the care?

- The data are limited, but do seem to show a performance gap.
- There is a gap. No disparity info
- performance ranges from 75-89%. interestingly has decreased from 2016 to 2017, likley due to trends towards decreased opioid prescribing. disparity data not provided
- performance continues to be variable from year to year
- PQRS and MIPS measures indicate that 11-25% of patients don't receive pain intervention. Given that this measure is dependent upon the pain assessment measure and that the facilities that measure the pain may have better pathways for addressing pain, the gravity of the situation is significant and needs to be addressed.
- Performance Data on subgroups provided
- Performance data was submitted and there is a gap, no performance data on disparities and if the MIPS data submission matches other programs
- The developer provided data from the PQRS reporting systems and the MIPS program from 2015-2017. The data provided demonstrated a decrease in performance between 2016 and 2017 as the number of eligible professionals increased (from 89.11% to 75.24%.) The developer noted that disparity data has not yet been made available. The developer provided a review of the recent literature to suggest that disparities in care continue to exist. I believe that data presented suggests a gap in care and supports the importance to measure and report for this measure.
- Moderate, summary data from literature demonstrates gap in presence of plans of care for pain
- Some gap in performance, no increase yearly measure
- I think there is an opportunity for improvement in the documentation of of a pain care plan.

Criteria 2: Scientific Acceptability of Measure Properties

2a. Reliability: Specifications and Testing

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

Reliability

2a1. Specifications requires the measure, as specified, to produce consistent (reliable) and credible (valid) results about the quality of care when implemented. For maintenance measures – no change in emphasis – specifications should be evaluated the same as with new measures.

2a2. Reliability testing demonstrates if the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise enough to distinguish differences in performance across providers. For maintenance measures – less emphasis if no new testing data provided.

Validity

2b2. Validity testing should demonstrate the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For maintenance measures – less emphasis if no new testing data provided.

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

Composite measures only:

2d. Empirical analysis to support composite construction. Empirical analysis should demonstrate that the component measures add value to the composite and that the aggregation and weighting rules are consistent with the quality construct.

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

Evaluators: NQF Staff

Questions for the Committee regarding reliability:

- Do you have any concerns that the measure can be consistently implemented (i.e., are measure specifications adequate)?
- The staff are satisfied with the reliability testing for the measure. Does the Committee think there is a need to discuss and/or vote on reliability?

Questions for the Committee regarding validity:

- Do you have any concerns regarding the validity of the measure (e.g., exclusions, risk-adjustment approach, etc.)?
- The staff are satisfied with the validity analyses for the measure. Does the Committee think there is a need to discuss and/or vote on validity?

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

Type of measure:

☒ **Process** ☐ **Process: Appropriate Use** ☐ **Structure** ☐ **Efficiency** ☐ **Cost/Resource Use**

☐ **Outcome** ☐ **Outcome: PRO-PM** ☐ **Outcome: Intermediate Clinical Outcome** ☐ **Composite**

Data Source:

☐ **Claims** ☐ **Electronic Health Data** ☐ **Electronic Health Records** ☐ **Management Data**
☐ **Assessment Data** ☒ **Paper Medical Records** ☐ **Instrument-Based Data** ☒ **Registry Data**
☐ **Enrollment Data** ☐ **Other**

Level of Analysis:

☒ **Clinician: Group/Practice** ☐ **Clinician: Individual** ☐ **Facility** ☐ **Health Plan**
☐ **Population: Community, County or City** ☐ **Population: Regional and State**
☐ **Integrated Delivery System** ☐ **Other**

Measure is:

☐ **New** ☒ **Previously endorsed (**NOTE: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.)

RELIABILITY: SPECIFICATIONS

1. **Are submitted specifications precise, unambiguous, and complete so that they can be consistently implemented?** ☒ **Yes** ☐ **No**

   Submission document: "MIF_xxxx" document, items S.1-S.22

   NOTE*: NQF staff will conduct a separate, more technical, check of eCQM specifications, value sets, logic, and feasibility, so no need to consider these in your evaluation.*

2. **Briefly summarize any concerns about the measure specifications.**

RELIABILITY: TESTING

Submission document: "MIF_xxxx" document for specifications, testing attachment questions 1.1-1.4 and section 2a2

3. **Reliability testing level** ☒ **Measure score** ☒ **Data element** ☐ **Neither**
4. **Reliability testing was conducted with the data source and level of analysis indicated for this measure** ☒ **Yes** ☐ **No**
5. If score-level and/or data element reliability testing was NOT conducted or if the methods used were NOT appropriate, was **empirical VALIDITY testing** of patient-level data conducted?

   ☐ **Yes** ☐ **No**

6. **Assess the method(s) used for reliability testing**

   Submission document: Testing attachment, section 2a2.2

   Reliability of the computed measure score was measured as the ratio of signal to noise and testing was performed by using a beta-binomial model

7. **Assess the results of reliability testing**

   Submission document: Testing attachment, section 2a2.3

   A reliability equal to zero implies that all the variability in a measure is attributable to measurement error. A reliability equal to one implies that all the variability is attributable to real differences in physician performance. Testing results indicated that the reliability above the minimum level of quality reporting events for 366 providers (based on TIN identifiers) was 0.98.

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

   Submission document: Testing attachment, section 2a2.2

   ☒ **Yes**

   ☐ **No**

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

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

   Submission document: Testing attachment, section 2a2.2

   ☐ **Yes**

   ☐ **No**

   ☒ **Not applicable** (data element testing was not performed)

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

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

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

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

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

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

    - Reliability score = 0.98

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

    Submission document: Testing attachment, section 2b2.

    N/A

13. **Please describe any concerns you have regarding the ability to identify meaningful differences in performance.**

    Submission document: Testing attachment, section 2b4.

    Data source is not clear. Measure is specified at group level; however, this data refers to facilities.

14. **Please describe any concerns you have regarding comparability of results if multiple data sources or methods are specified.**
    Submission document: Testing attachment, section 2b5.
    N/A

15. **Please describe any concerns you have regarding missing data.**

    Submission document: Testing attachment, section 2b6.

    Testing and analysis of missing data not done.

16. **Risk Adjustment**

    16a. **Risk-adjustment method** ☒ **None** ☐ **Statistical model** ☐ **Stratification**

    16b. **If not risk-adjusted, is this supported by either a conceptual rationale or empirical analyses?**

☐ Yes ☐ No ☒ Not applicable

16c. **Social risk adjustment:**

16c.1 Are social risk factors included in risk model? ☐ Yes ☐ No ☒ Not applicable

16c.2 Conceptual rationale for social risk factors included? ☐ Yes ☐ No

16c.3 Is there a conceptual relationship between potential social risk factor variables and the measure focus? ☐ Yes ☐ No

16d.**Risk adjustment summary:**

16d.1 All of the risk-adjustment variables present at the start of care? ☐ Yes ☐ No

16d.2 If factors not present at the start of care, do you agree with the rationale provided for inclusion? ☐ Yes ☐ No

16d.3 Is the risk adjustment approach appropriately developed and assessed? ☐ Yes ☐ No

16d.4 Do analyses indicate acceptable results (e.g., acceptable discrimination and calibration) ☐ Yes ☐ No

16d.5.Appropriate risk-adjustment strategy included in the measure? ☐ Yes ☐ No

16e**. Assess the risk-adjustment approach**

VALIDITY: TESTING

17. **Validity testing level: ☒ Measure score ☐ Data element ☐ Both**

18. **Method of establishing validity of the measure score:**

☐ Face validity

☒ Empirical validity testing of the measure score

☐ N/A (score-level testing not conducted)

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

Submission document: Testing attachment, section 2b2.2

- The developer performed a correlation analysis with measure: Oncology: Medical and Radiation – Pain Intensity Quantified (PQRS #143/NQF #0384) due to the the similarities in patient population and domain. This method can demonstrate an association between patients with a diagnosis of cancer receiving chemotherapy or radiation therapy in which pain intensity is quantified (NQF # 0384) and those with a diagnosis of cancer receiving chemotherapy or radiation therapy who report having pain with a documented plan of care to address pain (PQRS #144/NQF #0383).
    - Number of providers included in analysis based on TIN identifiers = 183

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

Submission document: Testing attachment, section 2b2.3

- The developer reports a coefficient correlation of 0.69 (P-value = > 0.001).

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

Submission document: Testing attachment, section 2b1.

☒ Yes

☐ No

☐ Not applicable (score-level testing was not performed)

22. **Was the method described and appropriate for assessing the accuracy of ALL critical data elements?** *NOTE that data element validation from the literature is acceptable.*

Submission document*: Testing attachment, section 2b1.*

☐ **Yes**

☐ **No**

☒ **Not applicable** (data element testing was not performed)

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

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

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

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

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

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

- The developer provided additional information on January 14, 2019 regarding the number of provider and patients include din the correlation analysis. This information will be added to the measure testing attachment at a later date.
    - Number of shared providers included in analysis based on NPI and TIN identifiers = 111
    - Number of PQRS #143 (Pain Intensity Quantified) qualifying events included in analysis = 26,635
    - Number of PQRS #144 (Plan of Care) qualifying events included in analysis = 10,785

ADDITIONAL RECOMMENDATIONS

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

The Standing Committee should discuss the testing method used to demonstrate validity and determine if it is appropriate and if the testing results are sufficient

Committee Pre-evaluation Comments:
Criteria 2: Scientific Acceptability of Measure Properties (including all 2a, 2b, and 2c)

2a1. Reliability-Specifications: Which data elements, if any, are not clearly defined? Which codes with descriptors, if any, are not provided? Which steps, if any, in the logic or calculation algorithm or other specifications (e.g., risk/case-mix adjustment, survey/sampling instructions) are not clear? What concerns do you have about the likelihood that this measure can be consistently implemented?

- No concerns
- Variability in who reports pain and how the plan is documented
- no concerns
- no concerns
- No reliability Threats
- Information can change about pain management, which could make the measure not consistently implemented.

- Data elements regarding what constitutes a plan are very vague and could be interpreted multiple ways at each center, and by data collectors.
- The measure specifics are well-described and complete and would allow the reliable evaluation of this measure. The measure was revised in 2019 to include two different populations (chemotherapy patients and radiation therapy patients both undergoing active therapy and experiencing pain.) I believe that this measure can be reliably implemented.
- No concerns
- Problematic as different scales and very subjective
- No concerns that the measure can be consistently implemented.

2a2. Reliability - Testing: Do you have any concerns about the reliability of the measure?

- No
- Testing seemed fine
- no concerns
- no
- No concern
- No concerns
- Yes
- The reliability score is submitted as a single score based on the split population. The study was assessed for ICD 10 codes updates. For this submission, data from 2015-2017 was provided by CMS. This included 366 unique sites and 108,765 patients representing a very large data set. The overall reliability score was 0.98 which suggests a high degree of reliability. I have no concerns regarding the reliability of this measure.
- No concerns
- yes
- No concerns

2b1. Validity -Testing: Do you have any concerns with the testing results?

- No
- Seems okay
- no concerns
- no
- No validity issues
- No concerns
- Yes
- For validity testing, CMS provided data from 371 physicians reporting on this measure through the registry option for CMS's PQRS in 2016. Of those, 251 physicians had all the required data elements and met the minimum number of quality reporting events (10) for a total of 44,795 quality events. For validity testing with this submission, this measure was correlated with another variable, pain intensity qualified. The data was positively correlated with a coefficient of correlation of 0.69 and a P-value of > 0.001. This methodology appears to be appropriate. I have no concerns regarding the validity of these testing results.
- No concerns
- yes
- No concerns

2b4-7. Threats to Validity (Statistically Significant Differences, Multiple Data Sources, Missing Data)2b4. Meaningful Differences: How do analyses indicate this measure identifies meaningful differences about quality? 2b5. Comparability of performance scores: If multiple sets of specifications: Do analyses indicate they produce comparable results? 2b6. Missing data/no response: Does missing data constitute a threat to the validity of this measure?

- No
- No issues
- none appreciated
- missing data/failure in chart abstraction is the biggest risk
- Missing data will be significant for this measure.
- No threat to the validity of this measure
- Measure may not be able to show meaningful differences cannot be shown due to this being a process measure only
- There are no exclusions to the testing. There was no risk assessment or stratification. Socio-demographic variables were not captured in this data set. There was no missing data. The data presented suggest that the mean performance was 78.6% and most fall outside of the 95% confidence interval suggesting statistically significant differences between facilities.
- No concerns.
- no
- No concerns

2b2-3. Other Threats to Validity (Exclusions, Risk Adjustment)2b2. Exclusions: Are the exclusions consistent with the evidence? Are any patients or patient groups inappropriately excluded from the measure?2b3. Risk Adjustment: If outcome (intermediate, health, or PRO-based) or resource use performance measure: Is there a conceptual relationship between potential social risk factor variables and the measure focus? How well do social risk factor variables that were available and analyzed align with the conceptual description provided? Are all of the risk-adjustment variables present at the start of care (if not, do you agree with the rationale provided)? Was the risk adjustment (case-mix adjustment) appropriately developed and tested? Do analyses indicate acceptable results? Is an appropriate risk-adjustment strategy included in the measure?

- No concerns
- No risk adjustment or risk stratification
- none appreciated
- no risk adjustment
- No risk adjustment is available.
- Yes - appropriate strategy included and includes ASCO related information
- N/A
- There are no exclusions to the testing. There was no risk assessment or stratification. Socio-demographic variables were not captured in this data set. There was no missing data. The data presented suggest that the mean performance was 78.6% and most fall outside of the 95% confidence interval suggesting statistically significant differences between facilities.
- No Concerns
- not sure
- No concerns

Criterion 3. Feasibility

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

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

- Data generated by and used by healthcare personnel during the provision of care.
- All data elements are in defined fields in electronic clinical data
- All the data elements needed for this measure are collected through electronic data or through the use of keyword searches.

Questions for the Committee:

- Are the required data elements routinely generated and used during care delivery?
- Are the required data elements available in electronic form, e.g., EHR or other electronic sources?
- Does using keyword searches to collect some of the data elements needed for this measure impact its feasibility?

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

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 3: Feasibility

3. Feasibility: Which of the required data elements are not routinely generated and used during care delivery? Which of the required data elements are not available in electronic form (e.g., EHR or other electronic sources)? What are your concerns about how the data collection strategy can be put into operational use?

- No concerns
- It's definitely feasible
- it might be fairly difficult to capture exchanges including mild pain or limited pain or transient pain given lack of fields in EHR
- This is going to require chart abstraction for the vast majority of charts currently which may be cumbersome
- The pain intervention is not directly documented and needs to be searched in the records. What exactly is considered intervention is also unclear, e.g. does providing education is considered intervention.
- Certain data only targets late stage patients
- Elements are documented during routine care however they are either documented in a narrative note, an order (i.e. pain medication, referral), or in an electronic way depending on EHR build. There is no standard element built into most EHR platforms. This metric requires manual audit.
- The data elements are routinely generated during the course of patient care. The treatment plan is not always included in an electronic form but is part of the medical record. I am not concerned about the ability to capture this data and to use the data operationally.
- No concerns about feasibility, data elements routinely collected in EHRs
- Concern that it may not be Broadly measurable in EMRs
- I do not have any input on the feasibility of implementing this measure.

Criterion 4: Usability and Use

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

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

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

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

Current uses of the measure

Publicly reported? ☒ **Yes** ☐ **No**

Current use in an accountability program? ☒ **Yes** ☐ **No** ☐ **UNCLEAR**

OR

Planned use in an accountability program? ☐ **Yes** ☐ **No**

Accountability program details

- This measure is currently used in the Merit Based Incentive Payment System (MIPS) program and the American Society of Clinical Oncology´s Quality Oncology Practice Initiative (QOPI®).
- NQF 0383 is in the PPS-Exempt Cancer Hospital Quality Reporting (PCHQR) Program, measure PCH-15 and is publicly reported on data.medicare.gov: https://data.medicare.gov/Hospital-Compare/Oncology-Care-Measures-PPS-Exempt-Cancer-Hospital/g3fg-3myg/data

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

Feedback on the measure by those being measured or others

- QOPI staff send survey data to all participating practices at the end of each round twice a year. Feedback is also obtained by emailing the QOPI Help Desk.

Additional Feedback:

- No additional feedback has been received by QOPI staff or ASCO on this measure.

Questions for the Committee:

- How have 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**

RATIONALE:

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

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

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

- This measure has shown improvement for patients that have been by either physicians participating in QOPI and/or physicians participating in MIPs. While this measure has shown improvement, pain management continues to be a real issue for cancer patients.

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

Unexpected findings (positive or negative) during implementation

- None reported

Potential harms

- None reported

Questions for the 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**

RATIONALE:

Committee Pre-evaluation Comments:
Criteria 4: Usability and Use

4a1. Use - Accountability and Transparency: How is the measure being publicly reported? Are the performance results disclosed and available outside of the organizations or practices whose performance is measured? For maintenance measures - which accountability applications is the measure being used for? For new measures - if not in use at the time of initial endorsement, is a credible plan for implementation provided?4a2. Use - Feedback on the measure: Have those being measured been given performance results or data, as well as assistance with interpreting the measure results and data? Have those being measured or other users been given an opportunity to provide feedback on the measure performance or implementation? Has this feedback has been considered when changes are incorporated into the measure?

- Used in public reporting
- It's in use
- currently in use and reported
- part of MIPS
- reported through MIPS and PQRS.
- Feedback has been considered and incorporated
- Measure is used in multiple reporting programs
- The measure is already in use in the PQRS and the MIP programs and has been in use for a while. There have been changes to the measure over time reflecting input from external users.
- Currently used in MIPS
- No concern
- I believe the results can lead to further the goal of high-quality, efficient healthcare.

4b1. Usability – Improvement: How can the performance results be used to further the goal of high-quality, efficient healthcare? If not in use for performance improvement at the time of initial endorsement, is a credible rationale provided that describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations?4b2. Usability – Benefits vs. harms: Describe any actual unintended consequences and note how you think the benefits of the measure outweigh them.

- Some improvement has been noted
- Unintended consequences unlikely
- no concerns
- reasonable usability, possible unintended consequences of increasing opioid usage, but in cancer patients this may be an acceptable tradeoff
- focus on patients with advanced cancer and guidance on the type of intervention and documentation of intervention can help with adoption and usability of the measure.
- None
- Unintended consequence could be the overtreating of patients with lower level of pain. Since the metric doesnt account for patients with acceptable vs unacceptable pain levels, the feedback from clinicians is that the measure is not as usable to improve performance.
- This measure is based on the documentation of a plan of care based on the associated documentation of pain. The measure describes alternative pain interventions that are all appropriate. The risks of addressing pain are associated with abuse and/or misuse of narcotics but the measure and the clinical guidelines upon which it is based suggest that there are many appropriate alternatives to analgesics. The risks of documenting appropriate therapeutic interventions are reasonable when compared to the benefits of addressing pain and quality of life in the patient.
- Improvements can certainly further the goal of high quality health care through improvements in the QOL of cancer survivors undergoing treatment
- Highly likely to benefit patients QOL. Better pain control and important question to patients.
- I believe the benefits outweigh any negative consequences.

Criterion 5: Related and Competing Measures

Related or competing measures

- 0524 : Pain Interventions Implemented During Short Term Episodes Of Care
- 1628 : Patients with Advanced Cancer Screened for Pain at Outpatient Visits

Harmonization

- Measure 0524 focuses on steps to monitor and mitigate pain were implemented. 0383 is similar in concept seeking a plan of care to address pain. A plan of care is further defined as include: use of opioids, non-opioid analgesics, psychological support, patient and/or family education, referral to a pain clinic, or reassessment of pain at an appropriate time interval. The measure also focuses on a broader population who are in home health. 0383 addresses the patients with a cancer diagnosis and does not limit itself to home health population. Measure 1628 targets only patients with Stage IV cancer. 0383 looks at any stage of cancer for purposes of managing pain for whom chemotherapy or radiation may be appropriate.

Committee Pre-evaluation Comments: Criterion 5: Related and Competing Measures

5. Related and Competing: Are there any related and competing measures? If so, are any specifications that are not harmonized? Are there any additional steps needed for the measures to be harmonized?

- This measure is more meaningful than 384 and e384 because it goes beyond simply quantifying pain and measures whether a plan of care is in place. We should discuss whether we need both measures. If not, this measure (383) is more meaningful for patients.
- Yes. Harmonization seems possible
- related to measures 0524 and 1628
- seems reasonably harmonized
- 0524 and 1628
- No additional steps at this time, however continued feedback for pain care is important
- the specifications between performance programs should be evaluated to look for opportunities to harmonize
- There are several competing measures and associated measures that must be harmonized.
- 0524 and 1628 harmonization appropriate
- There is a parallel one submitted from ASCO
- I do not have any concerns with related measures, 0524 & 1628.

Public and Member Comments

Comments and Member Support/Non-Support Submitted as of: 02/14/2020

- No comments received

ADDITIONAL RECOMMENDATIONS

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

Developer Submission

Additional evaluations and submission materials attachments…

1. Evidence and Performance Gap – Importance to Measure and Report

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

1a. Evidence to Support the Measure Focus – See attached Evidence Submission Form

383_NQF_EvidenceAttachment_1.23.17.docx,NQF_evidence_attachment_0383_110819_FINAL-637091624174224446.docx

1a.1 For Maintenance of Endorsement: Is there new evidence about the measure since the last update/submission?

Do not remove any existing information. If there have been any changes to evidence, the Committee will consider the new evidence. Please use the most current version of the evidence attachment (v7.1). Please use red font to indicate updated evidence.

No

1a. Evidence (subcriterion 1a)

Measure Number (*if previously endorsed*)**:** 0383

Measure Title: Oncology: Plan of Care for Pain – Medical Oncology and Radiation Oncology

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: Click here to enter composite measure #/ title

Date of Submission: 11/1/2019

1a.1.This is a measure of: (*should be consistent with type of measure entered in De.1*)

Outcome

☐ Outcome: Click here to name the health outcome

☐ Patient-reported outcome (PRO): Click here to name the PRO

PROs include HRQoL/functional status, symptom/symptom burden, experience with care, health-related behaviors. (A PRO-based performance measure is not a survey instrument. Data may be collected using a survey instrument to construct a PRO measure.)

☐ Intermediate clinical outcome (*e.g., lab value*): Click here to name the intermediate outcome

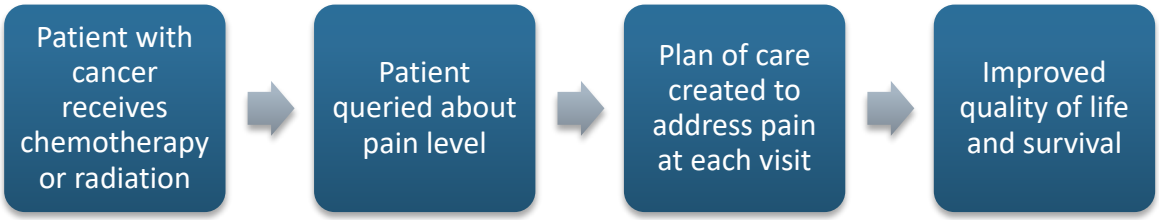
☒ Process: Plan of care to address pain at each visit for patients receiving chemotherapy or radiation therapy

☐ Appropriate use measure: Click here to name what is being measured

☐ Structure: Click here to name the structure

☐ Composite: Click here to name what is being measured

1a.2 LOGIC MODEL Diagram or 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.



Proper pain management is critical to achieving pain control. Pain has a severe impact on a patient's quality of life (1). Additionally, cancer pain is associated with numerous psychosocial responses (2-3). One third of patients describe cancer pain as intolerable aspect of cancer (4). Adequate pain treatment results in clinically relevant improvement in health-related quality of life (5). This is reflected in the most recent National Comprehensive Cancer Network (NCCN) guidelines, which stated that "unrelieved pain denies patients comfort and greatly affects their activities, motivation, interactions with family and friends, and overall quality of life" (6). Moreover, the importance of assessing pain in cancer patients is included in European guidelines, which go as far to say that despite published guidelines and education programs on the assessment and treatment of cancer related pain, unrelieved pain continues to be a substantial concern in patients worldwide (7). Given that it is projected that there will be over 15 million cancer patients in 2020 worldwide, this only increased the importance of addressing address patient pain (8).

1. IASP. 2008-2009 Global Year Against Cancer Pain 2008. Available at: https://www.iasp-pain.org/GlobalYear/CancerPain. Accessed October 17, 2019.

2. Kroenke K, Theobald D, Wu J, et al. The association of depression and pain with health-related quality of life, disability, and health care use in cancer patients. J Pain Symptom Manage 2010;40:327e341.

3. Porter LS, Keefe FJ. Psychosocial issues in cancer pain. Curr Pain Headache Rep 2011;15:263e270.

4. Breivik H, Cherny N, Collett B, et al. Cancer-related pain: a pan-European survey of prevalence, treatment, and patient attitudes. Ann Oncol 2009;20:1420e1433.

5. Puetzler J, Feldmann RE Jr, Brascher AK, Gerhardt A, Benrath J. Improvements in health-related quality of life by comprehensive cancer pain therapy: a pilot study with breast cancer outpatients under palliative chemotherapy. Oncol Res Treat 2014;37:456e462.

6. Swarm RA, Paice JA, Anghelescu DL, et al. NCCN Guidelines Panel. NCCN Clinical Practice Guidelines in Oncology – Adult Cancer Pain. Version 3. 2019. June 24, 2019.
https://www.nccn.org/professionals/physician_gls/pdf/pain.pdf

7. Management of Cancer Pain: ESMO Clinical Practice Guidelines. C. I. Ripamonti, D. Santini, E. Maranzano, M. Berti, F. Roila . Ann Oncol 2012; 23 (Suppl 7): vii39-vii154.

8. Frankish H. 15 million new cancer cases per year by 2020, says WHO. Lancet 2003; 361: 1278.

1a.3 Value and Meaningfulness: IF this measure is derived from patient report, provide evidence that the target population values the measured ***outcome, process, or structure*** and finds it meaningful. (Describe how and from whom their input was obtained.)

****RESPOND TO ONLY ONE SECTION BELOW -EITHER 1a.2, 1a.3 or 1a.4) ****

1a.2 FOR OUTCOME MEASURES including PATIENT REPORTED OUTCOMES - Provide empirical data demonstrating the relationship between the outcome (or PRO) to at least one healthcare structure, process, intervention, or service.

1a.3. SYSTEMATIC REVIEW(SR) OF THE EVIDENCE (for INTERMEDIATE OUTCOME, PROCESS, OR STRUCTURE PERFORMANCE MEASURES, INCLUDING THOSE THAT ARE INSTRUMENT-BASED) If the evidence is not based on a systematic review go to section 1a.4) If you wish to include more than one systematic review, add additional tables.

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

☒ Clinical Practice Guideline recommendation (with evidence review)

☐ US Preventive Services Task Force Recommendation

☐ Other systematic review and grading of the body of evidence (*e.g., Cochrane Collaboration, AHRQ Evidence Practice Center*)

☐ Other

Source of Systematic Review: • **Title** • **Author** • **Date** • **Citation, including page number** • **URL**	**Swarm RA, Paice JA, Anghelescu DL, et al. NCCN Guidelines Panel. NCCN Clinical Practice Guidelines in Oncology – Adult Cancer Pain. Version 3. 2019. June 24, 2019.** **https://www.nccn.org/professionals/physician_gls/pdf/pain.pdf**
Quote the guideline or recommendation verbatim about the process, structure or intermediate outcome being measured. If not a guideline, summarize the conclusions from the SR.	Management of pain in opioid-naïve patients • Select the most appropriate medication based on the pain diagnosis, comorbid conditions, and potential drug interactions. • Analgesic regimen may include an opioid, acetaminophen, nonsteroidal anti-inflammatory drugs (NSAIDs), and/or

	adjuvant analgesics. • Anticipate and treat analgesic adverse effects, including opioid-induced constipation. • Provide psychosocial support • Provide patient and family/caregiver education • Optimize integrative interventions • Reevaluate pain at each contact and as needed to meet patient-specific goals for comfort, function and safety. Management of pain in opioid-tolerant patients • Select the most appropriate medication based on the pain diagnosis, comorbid conditions, and potential drug interactions • Analgesic regimen may include an opioid, acetaminophen, NSAIDs, and/or adjuvant analgesics. • Anticipate and treat analgesic adverse effects, including opioid-induced constipation. • Provide psychosocial support. • Provide patient and family/caregiver education. • Optimize integrative interventions. • Reevaluate pain at each contact and as needed to meet patient-specific goals for comfort, function and safety. Goals of treatment If achieved: • Continue routine follow-up. • Re-evaluate need for opioids and reduce if appropriate. If not achieved: • See Universal Screening and Assessment. • Consider pain management specialty consultation. • Consider interventional strategies or other treatments. • Consider palliative care consultation.
Grade assigned to the **evidence** associated with the recommendation with the definition of the grade	Category 2A; Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate
Provide all other grades and definitions from the evidence grading system	NCCN Categories of Evidence and Consensus: • Category 1: Based upon high-level evidence, there is uniform NCCN consensus that the intervention is appropriate. • Category 2A: Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate. • Category 2B: Based upon lower-level evidence, there is NCCN consensus that the intervention is appropriate. • Category 3: Based upon any level of evidence, there is major NCCN disagreement that the intervention is appropriate.

Grade assigned to the **recommendation** with definition of the grade	Category 2A; Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate
Provide all other grades and definitions from the recommendation grading system	NCCN Categories of Evidence and Consensus: • Category 1: Based upon high-level evidence, there is uniform NCCN consensus that the intervention is appropriate. • Category 2A: Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate. • Category 2B: Based upon lower-level evidence, there is NCCN consensus that the intervention is appropriate. • Category 3: Based upon any level of evidence, there is major NCCN disagreement that the intervention is appropriate.
Body of evidence: • Quantity – how many studies? • Quality – what type of studies?	The NCCN guideline does not include an overview of the body of evidence used for the recommendations specific to the overall management of pain. However, the guideline does provide an in-depth discussion on the evidence, benefits and harms of specific therapies and interventions (e.g., aspirin, opioids, strategies for specific cancer pain syndromes, non-pharmacologic). This analysis includes the following summary (MS-29): • In most patients, cancer pain can be successfully managed with appropriate techniques and safe drugs. The overall approach to pain management encompassed in these guidelines is multimodal and comprehensive. It is based on routine pain assessments, utilizes both pharmacologic and nonpharmacologic interventions, and requires ongoing reevaluation of the patient. The NCCN Adult Cancer Pain Guidelines Panel advises that cancer pain can be well managed in the vast majority of patients if the algorithms presented are systematically applied, carefully monitored, and tailored to the needs of the individual patient.
Estimates of benefit and consistency across studies	See Body of Evidence section.
What harms were identified?	See Body of Evidence section.
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	Updated guidelines continue to support this measure.

1a.4 OTHER SOURCE OF EVIDENCE

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

1a.4.1 Briefly SYNTHESIZE the evidence that supports the measure. A list of references without a summary is not acceptable.

1a.4.2 What process was used to identify the evidence?

1a.4.3. Provide the citation(s) for the evidence.

1b. Performance Gap

1b.1. Briefly explain the rationale for this measure *(e.g., how the measure will improve the quality of care, the benefits or improvements in quality envisioned by use of this measure)*

If a COMPOSITE (e.g., combination of component measure scores, all-or-none, any-or-none), SKIP this question and answer the composite questions.

Proper pain management is critical to achieving pain control. Pain has a severe impact on a patient's quality of life (1). Additionally, cancer pain is associated with numerous psychosocial responses (2-3). One third of patients describe cancer pain as intolerable aspect of cancer (4). Adequate pain treatment results in clinically relevant improvement in health-related quality of life (5). This is reflected in the most recent NCCN guidelines which stated that unrelieved pain denies [patients] comfort and greatly affects their activities, motivation, interactions with family and friends, and overall quality of life (6). Moreover, the importance of assessing pain in cancer patients is included in European guidelines, which go as far to say that despite published guidelines and education programs on the assessment and treatment of cancer related pain, unrelieved pain continues to be a substantial concern in patients worldwide (7). Given that it is projected that there will be over 15 million cancer patients in 2020 worldwide, this only increased the importance of addressing address patient pain (8).

This measure aims to improve attention to pain management and requires a plan of care for cancer patients be documented who report having pain to allow for individualized treatment based on clinical circumstances and patient wishes and focuses on early documentation of a pain plan of care.

Citations:

1. IASP. 2008-2009 Global Year Against Cancer Pain 2008. Available at: https://www.iasp-pain.org/GlobalYear/Cancer Pain. Accessed February 10, 2015.

2. Kroenke K, Theobald D, Wu J, et al. The association of depression and pain with health-related quality of life, disability, and health care use in cancer patients. J Pain Symptom Manage 2010;40:327e341.

3. Porter LS, Keefe FJ. Psychosocial issues in cancer pain. Curr Pain Headache Rep 2011;15:263e270.

4. Breivik H, Cherny N, Collett B, et al. Cancer-related pain: a pan-European survey of prevalence, treatment, and patient attitudes. Ann Oncol 2009;20:1420e1433.

5. Puetzler J, Feldmann RE Jr, Brascher AK, Gerhardt A, Benrath J. Improvements in health-related quality of life by comprehensive cancer pain therapy: a pilot study with breast cancer outpatients under palliative chemotherapy. Oncol Res Treat 2014;37:456e462.

6. National Comprehensive Cancer Network (NCCN). Clinical Practice Guidelines in Oncology: Adult Cancer Pain. Version 2, 2017. Available at: http://www.nccn.org.

7. Management of Cancer Pain: ESMO Clinical Practice Guidelines. C. I. Ripamonti, D. Santini, E. Maranzano, M. Berti, F. Roila . Ann Oncol 2012; 23 (Suppl 7): vii39-vii154.

8. Frankish H. 15 million new cancer cases per year by 2020, says WHO. Lancet 2003; 361: 1278.

1b.2. Provide performance scores on the measure as specified (current and over time) at the specified level of analysis. *(This is required for maintenance of endorsement. Include mean, std dev, min, max, interquartile range, 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 (4b1) under Usability and Use.*

Data from 2015-2017 for this measure from the CMS Physician Quality Reporting System (PQRS) and the Merit-based Incentive Payment System (MIPS) demonstrate a continued opportunity for improvement.

2015:

88 eligible groups

4,643 patients

Mean: 83.43%

Confidence Interval for mean: (0.78, 0.89)

Minimum: 0.00%

Maximum: 100.00%

25th percentile: 77.60%

75th percentile: 100.00%

2016:

106 eligible groups

15,268 patients

Mean: 89.11%

Confidence Interval for mean: (0.85, 0.93)

Minimum: 25.55%

Maximum: 100.00%

25th percentile: 89.72%

75th percentile: 100.00%

2017:

244 eligible groups

88,854 patients

Mean: 75.24%

Confidence Interval for mean: (0.71, 0.80)

Minimum: 0.00%

Maximum: 100.00%

25th percentile: 64.25%

75th percentile: 100.00%

1b.3. If no or limited performance data on the measure as specified is reported in 1b2, 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.

1b.4. 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. *(This is required for maintenance of endorsement. Describe the data source including number of measured entities; number of patients; dates of data; if a sample, characteristics of the entities included.) 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 (4b1) under Usability and Use.

While this measure is included in the MIPS program, this program has not yet made disparities data available for us to analyze and report.

1b.5. If no or limited data on disparities from the measure as specified is reported in 1b.4, 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 1b.4

Studies and analyses of existing data demonstrate that patients with cancer continue to receive disparate treatment for pain. These differences demonstrate a continued need to evaluate the extent to which patients are adequately treated to address their pain. Examples of these disparities are highlight below.

Patients with cancer treated in centers with primarily minority populations have been shown to be three times more likely to have inadequately controlled pain than Caucasian, more affluent patients (1-3).

A study of 116 women in two programs with the aim of advocating, assisting, and supporting women with cancer in an urban area of northern California (4) found that being of low socioeconomic status, being Latino, and having a mastectomy followed by chemotherapy were important indicators for increased symptoms and poor pain management.

A recent study examined data from 4707 cancer survivors who reported experiencing pain from their cancer (5). A multilevel, socioecological, conceptual framework was used to generate a list of 15 barriers to pain management, representing patient, provider, and system levels. The study included separate multivariable logistic regressions for each barrier identified sociodemographic and health-related inequalities in cancer pain management, controlling for years since diagnosis, disease stage, and cancer treatment. The study found that two-thirds of survivors reported at least 1 barrier to pain management. While patient-related barriers were most common, the greatest disparities were noted in provider- and system-level barriers. Inequalities by race/ethnicity, education, age, and physical and mental health comorbidities were observed. Additionally, researchers found that survivors who were nonwhite, less educated, older, and/or burdened by comorbidities were most adversely affected.

Another study examined patterns of disparities in cancer pain by evaluating differences by race/ethnicity in the odds of reporting pain and in pain severity, controlling for key patient-level covariates (6). This study used data from a nationally representative cohort of colorectal and lung cancer patients. The study included 5761 individuals (14% black, 7% Hispanic/Latino, 6% Asian or Pacific Islander, and 3% multiracial), among whom 48% reported pain. The adjusted odds of reporting differed only for multiracial patients, who were more likely to report pain than whites (odds ratio: 1.54; P = 0.036). However, among those with pain, severity was higher for black patients (ß = 6.6; P = 0.001) and multiracial patients (ß = 4.5; P = 0.036) relative to white patients. Lower educational attainment, depressed affect, and lower levels of wealth also were associated with higher pain severity. The researchers concluded that sociodemographic status, health status, and depression were associated with severity but did not explain the disparity. Interventions to address these disparities will need to focus on reported severity and patient-level factors.

One study set out to see what factors are associated with unmet needs for symptom management in patients with lung and colorectal cancer (7). This study found that 15% (791 of 5,422) of patients had at least one unmet need for symptom management. Adjusting for sociodemographic and clinical factors, African American race, being uninsured or poor, having early-stage lung cancer, and the presence of moderate to severe symptoms were associated with unmet need (all P < .05). Furthermore, patients who rated their physician ´s communication score < 80 (on a 0 to 100 scale) had adjusted rates of an unmet need for symptom management that were more than twice as high as patients who rated their physicians with a perfect communication score (23.1% v 10.0%; P < .001). The researchers concluded that a significant minority of patients with newly diagnosed lung and colorectal cancer report unmet needs for symptom management.

Interventions to improve symptom management should consider the importance of physician communication to the patient ´s experience of disease.

The Outpatient Pain Clinics at Memorial Sloan Kettering Cancer Center participated in developing a pain registry to gain insight on the referral and management of cancer pain as related to demographic information, cancer history, prescription records, and interventional pain procedures stored in the institutional database. Five cohorts (subsets of one another) were defined and compared to describe demographics and differences in management and outcomes by age, race, sex, and cancer type. Clinic patients were compared with the entire institution to determine factors associated with better pain relief and reduced side effects. Researchers found that a small percentage of patients were referred to a pain specialist. A total of 1,043 patients completed 3,544 surveys. Compared with the institution, there were higher proportions of patients age 51 to 60 years, nonwhites, and patients with thoracic, abdominal, and head and neck cancers. Medical management-controlled pain with three drug categories in 40% of visits. Short-acting opioids were the only category that statistically provided good pain relief with fewer side effects. Pain scores were improved with increasing opioid dose. Management differed by sex, age, and race; women consistently had lower doses of opioids, poorer pain control, more side effects, and were prescribed a greater variety of medications.

McNeill JA, Reynolds J, Ney ML. Unequal quality of cancer pain management: disparity in perceived control and proposed solutions. Oncol Nurs Forum. 2007 Nov;34(6):1121-8. citing:

1. Anderson KO, Mendoza TR, Valero V, Richman SP, Russell C, Hurley J, et al. Minority cancer patients and their providers: Pain management attitudes and practice. Cancer. 2000; 88, 1929–1938.

2. Cleeland C, Gonin R, Hatfield A, Edmonson J, Blum R, Stewart J, et al. Pain and its treatment in outpatients with metastatic cancer. New England Journal of Medicine. 1994; 330, 592–596.

3. Vallerand A, Hasenau S, Templin T, Collins-Bohler D. Disparities between black and white patients with cancer pain: The effect of perception of control over pain. Pain Medicine. 2005; 6, 242–250.

4. Eversley R, Estrin D, Dibble S, Wardlaw L, Pedrosa M, Favila-Penney W. Post-treatment symptoms among ethnic minority breast cancer survivors. Oncology Nursing Forum. 2005; 32, 250–256.

Additional Citations:

5. Stein KD, Alcaraz KI, Kamson C, Fallon EA, Smith TG. Sociodemographic inequalities in barriers to cancer pain management: a report from the American Cancer Society ´s Study of Cancer Survivors-II (SCS-II). Psychooncology. 2016 Oct;25(10):1212-1221. doi: 10.1002/pon.4218. Epub 2016 Aug 12.

6. Martinez KA, Snyder CF, Malin JL, Dy SM. Is race/ethnicity related to the presence or severity of pain in colorectal and lung cancer? Pain Symptom Manage. 2014 Dec;48(6):1050-9. doi: 10.1016/j.jpainsymman.2014.02.005. Epub 2014 Apr 18.

7. Walling AM, Keating NL, Kahn KL, Dy S, Mack JW, Malin J, Arora NK, Adams JL, Antonio AL, Tisnado D.Lower Patient Ratings of Physician Communication Are Associated With Unmet Need for Symptom Management in Patients With Lung and Colorectal Cancer. J Oncol Pract. 2016 Jun;12(6):e654-69. doi: 10.1200/JOP.2015.005538. Epub 2016 May 24.

8. VT Malhotra, P Glare, KS Tan, J Wills, A Gulati, V Puttanniah, J Hung, K Cubert, C Inturrisi; The Tri-Institutional Pain Registry— Analysis of Outpatient Pain Management at a Specialized Cancer Center, Pain Medicine, Volume 18, Issue 12, 1 December 2017, Pages 2474–2484, https://doi.org/10.1093/pm/pnx136.

…

2.3 <u>For maintenance of endorsement</u>

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated

even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

2. Reliability and Validity—Scientific Acceptability of Measure Properties

Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented. ***Measures must be judged to meet the sub criteria for both reliability and validity to pass this criterion and be evaluated against the remaining criteria.***

2a.1. Specifications The measure is well defined and precisely specified so it can be implemented consistently within and across organizations and allows for comparability. eMeasures should be specified in the Health Quality Measures Format (HQMF) and the Quality Data Model (QDM).

De.5. Subject/Topic Area *(check all the areas that apply):*

Cancer

De.6. Non-Condition Specific*(check all the areas that apply):*

Person-and Family-Centered Care

De.7. Target Population Category *(Check all the populations for which the measure is specified and tested if any):*

Elderly

S.1. Measure-specific Web Page *(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.)*

The updated specifications for this measure are included with this form.

S.2a. If this is an eMeasure, HQMF specifications must be attached. Attach the zipped output from the eMeasure 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)

This is not an eMeasure **Attachment:** NQF_evidence_attachment_0383_110819_FINAL.docx

S.2b. Data Dictionary, Code Table, or Value Sets *(and risk model codes and coefficients when applicable) must be attached. (Excel or csv file in the suggested format preferred - if not, contact staff)*

Attachment **Attachment:** 0383_NQF_PlanofCarePain_CodeSet_07312019.xlsx

S.2c. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

No, this is not an instrument-based measure **Attachment:** NQFScaleInstrument.pdf

S.2d. Is this an instrument-based measure (i.e., data collected via instruments, surveys, tools, questionnaires, scales, etc.)? Attach copy of instrument if available.

Not an instrument-based measure

S.3.1. For maintenance of endorsement: Are there changes to the specifications since the last updates/submission. If yes, update the specifications for S1-2 and S4-22 and explain reasons for the changes in S3.2.

Yes

S.3.2. For maintenance of endorsement, please briefly describe any important changes to the measure specifications since last measure update and explain the reasons.

The measure title has been modified from Oncology: Care Plan for Pain – Medical Oncology and Radiation Oncology to Oncology: Medical and Radiation – Care Plan for Pain to align with the title of the paired measure NQF 0384.

Beginning with 2019 implementation, the measure was revised to have two populations: 1) All patient visits for patients with a diagnosis of cancer currently receiving chemotherapy OR 2) All patient visits for patients with a diagnosis of cancer currently receiving radiation therapy. This change was made to more clearly delineate the denominator requirements to promote accurate implementation. Based on feedback we heard regarding how users have implemented the measure, there was an inconsistent approach to applying the measure criteria. Therefore, we decided to split this measure out into two populations, based on the type of treatment the patient is receiving. Though the measure is split into two, the measure still requires only one performance rate for reporting.

The ICD9 diagnosis codes have been removed and the ICD10 codes remain.

S.4. Numerator Statement *(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.*

IF an OUTCOME MEASURE, state the outcome being measured. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Patient visits that include a documented plan of care* to address pain.

*A documented plan of care may include: use of non-opioid analgesics, opioids, psychological support, patient and/or family education, referral to a pain clinic, or reassessment of pain at an appropriate time interval.

S.5. Numerator Details *(All information required to identify and calculate the cases from the target population with the target process, condition, event, or outcome such as definitions, 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 S.2b)*

IF an OUTCOME MEASURE, describe how the observed outcome is identified/counted. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Patient visits that included a documented plan of care to address pain.

Time Period for Data Collection: At each visit within the measurement period for patients with a diagnosis of cancer and in which pain is present.

Guidance: A documented outline of care for a positive pain assessment is required. May include: use of non-opioid analgesics, opioids, psychological support, patient and/or family education, referral to a pain clinic, or reassessment of pain at an appropriate time interval.

S.6. Denominator Statement *(Brief, narrative description of the target population being measured)*

All visits for patients, regardless of age, with a diagnosis of cancer currently receiving chemotherapy or radiation therapy who report having pain

S.7. Denominator Details *(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 S.2b.)*

IF an OUTCOME MEASURE, describe how the target population is identified. Calculation of the risk-adjusted outcome should be described in the calculation algorithm (S.14).

Time Period for Data Collection: 12 consecutive months

Denominator Criteria (Eligible Cases):

For all eligible patient encounters when pain severity quantified and pain is present (e.g., CPT II: 1125F is submitted in the numerator for NQF 0384) for patients regardless of age, with a diagnosis of cancer currently receiving chemotherapy or radiation therapy.

Guidance: This measure is an episode-of-care measure; the level of analysis for this measure is every visit for patients with a diagnosis of cancer who are also currently receiving chemotherapy or radiation therapy and a positive pain assessment during the measurement period. For patients receiving radiation therapy, pain intensity should be quantified at each radiation treatment management encounter. For patients receiving chemotherapy, pain intensity should be quantified at each face-to-face encounter with the physician while the patient is currently receiving chemotherapy.

All visits for patients, regardless of age

AND

Diagnosis of cancer

AND

Patient encounter during the performance period

AND

Patient reported pain was present

AND

Radiation treatment management encounter

OR

Face-to-face encounter with the physician while the patient is currently receiving chemotherapy

S.8. Denominator Exclusions *(Brief narrative description of exclusions from the target population)*

None

S.9. Denominator Exclusion Details *(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 S.2b.)*

N/A, no denominator exclusion

S.10. Stratification Information *(Provide all information required to stratify the measure results, if necessary, including 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 with at S.2b.)*

N/A, no risk stratification

S.11. Risk Adjustment Type (Select type. Provide specifications for risk stratification in measure testing attachment)

No risk adjustment or risk stratification

If other:

S.12. Type of score:

Rate/proportion

If other:

S.13. Interpretation of Score *(Classifies interpretation of score according to whether better quality is associated with a higher score, a lower score, a score falling within a defined interval, or a passing score)*

Better quality = Higher score

S.14. Calculation Algorithm/Measure Logic (*Diagram or describe the calculation of the measure score as an ordered sequence of steps including identifying the target population; exclusions; cases meeting the target process, condition, event, or outcome; time period for data, aggregating data; risk adjustment; etc.)*

This measure is comprised of two populations but is intended to result in one reporting rate. The reporting rate is the aggregate of Population 1 and Population 2, resulting in a single performance rate. For the purposes of this measure, the single performance rate can be calculated as follows:

Performance Rate = (Numerator 1 + Numerator 2)/ (Denominator 1 + Denominator 2)

Calculation algorithm for Population 1: Patient visits for patients with a diagnosis of cancer currently receiving chemotherapy

1. Find the patients who meet the initial population (i.e., the general group of patients that a set of performance measures is designed to address).

2. From the patients within the initial population criteria, find the patients who qualify for the denominator (i.e., the specific group of patients for inclusion in a specific performance measure based on defined criteria). Note: in some cases, the initial population and denominator are identical.

3. From the patients within the denominator, find the patients who meet the numerator criteria (i.e., the group of patients in the denominator for whom a process or outcome of care occurs). Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator

If the patient does not meet the numerator, this case represents a quality failure.

Calculation algorithm for Population 2: Patient visits for patients with a diagnosis of cancer currently receiving radiation therapy

1. Find the patients who meet the initial population (i.e., the general group of patients that a set of performance measures is designed to address).

2. From the patients within the initial population criteria, find the patients who qualify for the denominator (i.e., the specific group of patients for inclusion in a specific performance measure based on defined criteria). Note: in some cases, the initial population and denominator are identical.

3. From the patients within the denominator, find the patients who meet the numerator criteria (ie, the group of patients in the denominator for whom a process or outcome of care occurs). Validate that the number of patients in the numerator is less than or equal to the number of patients in the denominator.

If the patient does not meet the numerator, this case represents a quality failure.

S.15. Sampling *(If measure is based on a sample, provide instructions for obtaining the sample and guidance on minimum sample size.)*

IF an instrument-based performance measure (e.g., PRO-PM), identify whether (and how) proxy responses are allowed.

Measure is not based on a sample

S.16. Survey/Patient-reported data *(If measure is based on a survey or instrument, provide instructions for data collection and guidance on minimum response rate.)*

Specify calculation of response rates to be reported with performance measure results.

N/A, measure is not based on a survey or instrment

S.17. Data Source *(Check ONLY the sources for which the measure is SPECIFIED AND TESTED).*

If other, please describe in S.18.

Paper Medical Records, Registry Data

S.18. Data Source or Collection Instrument *(Identify the specific data source/data collection instrument (e.g. name of database, clinical registry, collection instrument, etc., and describe how data are collected.)*

IF instrument-based, identify the specific instrument(s) and standard methods, modes, and languages of administration.

N/A, measure is not instrument-based

S.19. Data Source or Collection Instrument *(available at measure-specific Web page URL identified in S.1 OR in attached appendix at A.1)*

No data collection instrument provided

S.20. Level of Analysis *(Check ONLY the levels of analysis for which the measure is SPECIFIED AND TESTED)*

Clinician : Group/Practice

S.21. Care Setting *(Check ONLY the settings for which the measure is SPECIFIED AND TESTED)*

Outpatient Services

If other:

S.22. COMPOSITE Performance Measure - Additional Specifications *(Use this section as needed for aggregation and weighting rules, or calculation of individual performance measures if not individually endorsed.)*

N/A

2. Validity – See attached Measure Testing Submission Form

0383_NQF_testing_attachment_073019_Final.docx

2.1 For maintenance of endorsement

Reliability testing: If testing of reliability of the measure score was not presented in prior submission(s), has reliability testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.2 For maintenance of endorsement

Has additional empirical validity testing of the measure score been conducted? If yes, please provide results in the Testing attachment. Please use the most current version of the testing attachment (v7.1). Include information on all testing conducted (prior testing as well as any new testing); use red font to indicate updated testing.

Yes

2.3 For maintenance of endorsement

Risk adjustment: For outcome, resource use, cost, and some process measures, risk-adjustment that includes social risk factors is not prohibited at present. Please update sections 1.8, 2a2, 2b1,2b4.3 and 2b5 in the Testing attachment and S.140 and S.11 in the online submission form. NOTE: These sections must be updated even if social risk factors are not included in the risk-adjustment strategy. You MUST use the most current version of the Testing Attachment (v7.1) -- older versions of the form will not have all required questions.

No - This measure is not risk-adjusted

Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (*if previously endorsed*)**:** 0383
Measure Title: Oncology: Medical and Radiation: Plan of Care for Pain
Date of Submission: 7/31/2019

Type of Measure:

☐ **Outcome (*including PRO-PM*)**	☐ **Composite – *STOP – use composite testing form***
☐ Intermediate Clinical Outcome	☐ Cost/resource
☒ Process *(including Appropriate Use)*	☐ Efficiency
☐ Structure	

1.1. What type of data was used for testing? (*Check all the sources of data identified in the measure specifications and data used for testing the measure. Testing must be provided for all the sources of data specified and intended for measure implementation.* ***If different data sources are used for the numerator and denominator, indicate N [numerator] or D [denominator] after the checkbox.***)

Measure Specified to Use Data From: (***must be consistent with data sources entered in S.17***)	**Measure Tested with Data From:**
☒ abstracted from paper record	☒ abstracted from paper record
☐ claims	☐ claims
☒ registry	☒ registry
☐ abstracted from electronic health record	☐ abstracted from electronic health record
☐ eMeasure (HQMF) implemented in EHRs	☐ eMeasure (HQMF) implemented in EHRs
☐ other: Click here to describe	☐ other: Click here to describe

1.2. 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*).

2019 Submission

The data source is 2015-2017 Registry data from the PQRS/MIPS program, provided by the Center for Medicare & Medicaid Services (CMS).

2015 Submission

Data reported are from the Fall 2011 PCPI Testing Project (reflecting data submitted August-November 2011), which are consistent with the measure specifications

1.3. What are the dates of the data used in testing? 1/1/2015-12/31/2017

1.4. What levels of analysis were tested? (*testing must be provided for all the levels specified and intended for measure implementation, e.g., individual clinician, hospital, health plan*)

Measure Specified to Measure Performance of: (***must be consistent with levels entered in item S.20***)	**Measure Tested at Level of:**
☐ individual clinician	☐ individual clinician
☒ group/practice	☒ group/practice
☐ hospital/facility/agency	☐ hospital/facility/agency
☐ health plan	☐ health plan

☐ other: Click here to describe	☐ other: Click here to describe

1.5. How many and which measured entities were 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*)

2019 Submission

Testing to identify reliability and statistically significant and meaningful differences in performance analysis was conducted using 2015-2016 PQRS/2017 MIPS performance from a registry data set provided from CMS. Practices were identified by unique number of TINs. The 2015-2017 data set was from 366 unique TINs including 108,765 patients. Additional descriptive characteristics of the measured entities, such as size and location type, are unknown. Entities submitted data for inclusion in this data set according to the eligibility and reporting requirements for PQRS/MIPS during the 2015-2017 program years. We were unable to determine from our rolled-up data sample the number of clinicians who reported to PQRS/MIPS as an individual or a group; therefore, this measure should be considered for endorsement at the group/practice level, with a potential group size as n of 1 or group of 1.

For validity testing we received data from 371 physicians reporting on this measure through the registry option for CMS's PQRS in 2016. Of those, 251 physicians had all the required data elements and met the minimum number of quality reporting events (10) for a total of 44,795 quality events. For this measure, 73 percent of physicians are included in the analysis, and the average number of quality reporting events are 179 for the remaining 44,795 events. The range of quality reporting events for 251 physicians included is from 10 to 3,811. The average number of quality reporting events for the remaining 27 percent of physicians that aren't included is 4.

2015 Submission

Data abstracted from patient records were used to calculate inter-rater reliability for the measure.

PCPI Testing Project

Five practice sites representing various types, locations and sizes were identified to participate in testing the PCPI/ASCO/ASTRO-developed measures.

- Site A: hospital, multi-practice sites in urban, rural and suburban settings; 21 physicians; average 9600 oncology/prostate cancer patient visits per month for MD/NP assessment, chemotherapy; submitted PQRS claims for one measure and utilized a full-fledged EHR.
- Site B: physician owned private practice, suburban setting; 4 physicians; average 48 oncology/prostate cancer patients seen per day; submitted PQRS claims for one measure and utilized paper medical records.
- Site C: physician owned private practice, urban setting; 41 physicians; average 2500 oncology/prostate cancer patients seen per month; submitted PQRS claims for two measures and utilized a full-fledged EHR.
- Site D: academic, suburban setting; 9 physicians; average 240 oncology/prostate cancer patients seen per month; submitted PQRS claims for one measure and utilized paper and EHR.
- Site E: academic, urban setting; 14 physicians; average 250 oncology/prostate cancer patients seen per month; collected PQRS data on 3 measures and utilized a full-fledged EHR.

- The measurement period (data collected from patients seen) was 1/1/2010 through 12/31/2010.
- Chart abstraction was performed between 8/8/2011 and 11/3/2011.

1.6. How many and which patients were included in the testing and analysis (by level of analysis and data source)? (*identify the number and descriptive characteristics of patients included in the analysis (e.g., age, sex, race, diagnosis); if a sample was used, describe how patients were selected for inclusion in the sample*)

2019 Submission

There were 108,765 patients included in reliability testing and statistically significant and clinically/practically meaningful differences in performance measure scores analysis. These were the eligible patients that were associated with the TIN.

There were 44,795 patients included in validity testing and analysis. These were the patients that were associated with physicians who had 10 or more patients eligible for this measure.

2015 Submission

360 patient visits were reviewed

1.7. 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 reported below.

2019 Submission

CMS PQRS/MIPS data from 2015-2017 was used for reliability testing and statistically significant and clinically/practically meaningful differences in performance measure scores analysis. For validity testing, a subset of PQRS 2016 data from CMS was used.

2015 Submission

Five practice sites representing various types, locations and sizes were identified to participate in testing the PCPI/ASCO/ASTRO-developed measures.

- Site A: hospital, multi-practice sites in urban, rural and suburban settings; 21 physicians; average 9600 oncology/prostate cancer patient visits per month for MD/NP assessment, chemotherapy; submitted PQRS claims for one measure and utilized a full-fledged EHR.
- Site B: physician owned private practice, suburban setting; 4 physicians; average 48 oncology/prostate cancer patients seen per day; submitted PQRS claims for one measure and utilized paper medical records.
- Site C: physician owned private practice, urban setting; 41 physicians; average 2500 oncology/prostate cancer patients seen per month; submitted PQRS claims for two measures and utilized a full-fledged EHR.
- Site D: academic, suburban setting; 9 physicians; average 240 oncology/prostate cancer patients seen per month; submitted PQRS claims for one measure and utilized paper and EHR.
- Site E: academic, urban setting; 14 physicians; average 250 oncology/prostate cancer patients seen per month; collected PQRS data on 3 measures and utilized a full-fledged EHR.

1.8 What were 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.

2019 Submission

Patient-level socio-demographic (SDS) variables were not captured as part of the testing.

2015 Submission

We encourage the results of this measure to be stratified by race, ethnicity, gender, and primary language, and have included these variables as recommended data elements to be collected.

The PCPI advocates that performance measure data should, where possible, be stratified by race, ethnicity, and primary language to assess disparities and initiate subsequent quality improvement activities addressing identified disparities, consistent with recent national efforts to standardize the collection of race and ethnicity data. A 2008 NQF report endorsed 45 practices including stratification by the aforementioned variables.(1) A 2009 IOM report "recommends collection of the existing Office of Management and Budget (OMB) race and Hispanic ethnicity categories as well as more fine-grained categories of ethnicity(referred to as granular ethnicity and based on one's ancestry) and language need (a rating of spoken English language proficiency of less than very well and one's preferred language for health-related encounters)."(2)

References:

(1) National Quality Forum Issue Brief (No.10). Closing the Disparities Gap in Healthcare Quality with Performance Measurement and Public Reporting. Washington, DC: NQF, August 2008.
(2) Race, Ethnicity, and Language Data: Standardization for Health Care Quality Improvement. March 2010. AHRQ Publication No. 10-0058-EF. Agency for Healthcare Research and Quality, Rockville, MD. Available at:
http://www.ahrq.gov/research/iomracereport. Accessed May 25, 2010.

2a2. RELIABILITY TESTING

***Note**: If accuracy/correctness (validity) of data elements was empirically tested, separate reliability testing of data elements is not required – in 2a2.1 check critical data elements; in 2a2.2 enter "see section 2b2 for validity testing of data elements"; and skip 2a2.3 and 2a2.4.*

2a2.1. What level of reliability testing was conducted? (*may be one or both levels*)
☒ **Critical data elements used in the measure** (*e.g., inter-abstractor reliability; data element reliability must address ALL critical data elements*)
☒ **Performance measure score** (e.g., *signal-to-noise analysis*)

2a2.2. For each level 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*)

2019 Submission:

Reliability of the computed measure score was measured as the ratio of signal to noise. The signal in this case is the proportion of the variability in measured performance that can be explained by real differences in physician performance and the noise is the total variability in measured performance. Reliability at the level of the specific physician is given by:

Reliability = Variance (facility-to-facility) / [Variance (facility-to-facility) + Variance (facility-specific-error]

Reliability is the ratio of the facility-to-facility variance divided by the sum of the facility-to-facility variance plus the error variance specific to a facility. A reliability of zero implies that all the variability in

a measure is attributable to measurement error. A reliability of one implies that all the variability is attributable to real differences in practice performance.

Reliability testing was performed by using a beta-binomial model. The beta-binomial model assumes the practice performance score is a binomial random variable conditional on the practice's true value that comes from the beta distribution. The beta distribution is usually defined by two parameters, alpha and beta. Alpha and beta can be thought of as intermediate calculations to get to the needed variance estimates.

To assess signal-to-noise, we employed the beta-binomial model as described by JL Adams (1). Each facility provided numerators and denominators in accordance with the measure specification. Through the estimation of the beta-binomial parameters (often referred to as alpha and beta) as described by Adams (1), we estimated the facility-to-facility variance and the within-facility variance (simply the binomial variance for each facility).

A reliability equal to zero implies that all the variability in a measure is attributable to measurement error. A reliability equal to one implies that all the variability is attributable to real differences in practice performance. A reliability of 0.70 – 0.80 is generally considered the acceptable threshold for reliability, 0.80 – 0.90 is considered high reliability, and 0.90 – 1.0 is considered very high. [1]

1. Adams JL, Mehrotra A, McGlynn EA, Estimating Reliability and Misclassification in Physician Profiling, Santa Monica, CA: RAND Corporation, 2010. www.rand.org/pubs/technical_reports/TR863. (Accessed on February 24, 2012.)

2015 Submission

- PCPI Testing Project
- Data abstracted from patient records were used to calculate inter-rater reliability for the measure.
- 360 patient visits were reviewed.
- Data analysis included:
  - Percent agreement; and
  - Kappa statistic to adjust for chance agreement

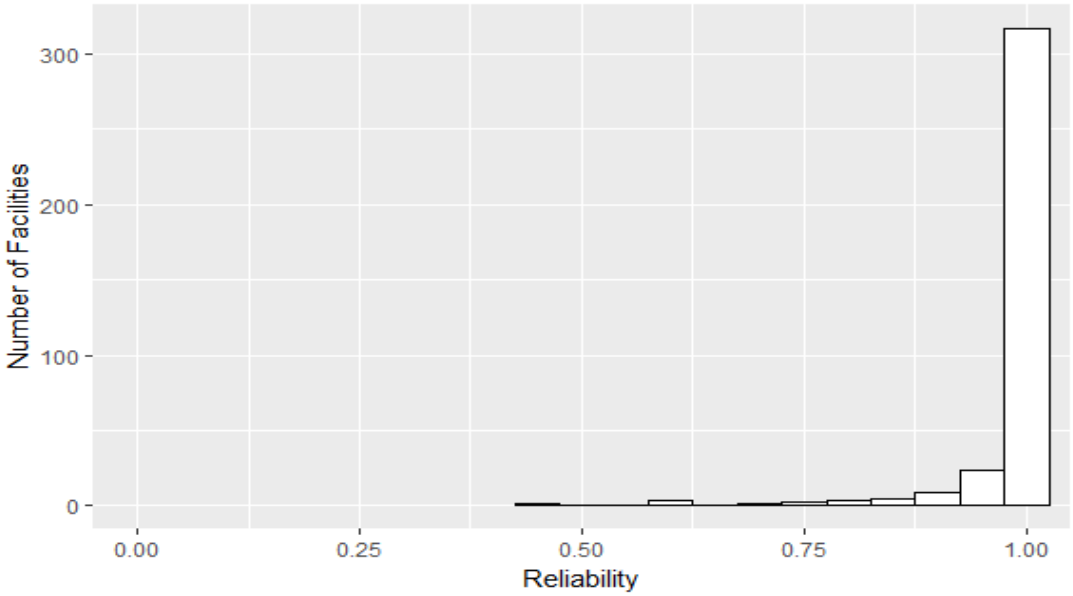
2a2.3. For each level of testing checked above, what were the statistical results from reliability testing? (e.*g., percent agreement and kappa for the critical data elements; distribution of reliability statistics from a signal-to-noise analysis*)

2019 Submission

The reliability above the minimum level of quality reporting events was above 0.95.

- Reliability for 0383 (144):
  - Methods: Signal-to-noise analysis using the Beta-Binomial
  - Results:
  - *Facility-level Reliability*

N	Alpha	Beta	Min	10th Pctl	Median	90th Pctl	Max	Mean
366	0.5552	0.1899	0.4654	0.9633	0.9998	1	1	0.9824



2015 Submission

N, % Agreement, Kappa (95% Confidence Interval)

Overall Reliability: 360, 100.0%, Kappa is noncalculable*

Denominator Reliability: 360, 100.0%, Kappa is noncalculable*

Numerator Reliability: 360, 100.0%, Kappa is noncalculable*

*Kappa Statistics cannot be calculated because of complete agreement. Confidence intervals cannot be calculated because to do so would involve dividing by zero which cannot be done.

2a2.4 What is your interpretation of the results in terms of demonstrating reliability? (i.*e., what do the results mean and what are the norms for the test conducted?*)

2019 Submission

The reliability is excellent. This measure has very high reliability: Mean reliability is 98%; the 10th percentile is 96%.

2015 Submission

This measure demonstrates almost perfect reliability, as shown in results from the above analysis

---

2b1. VALIDITY TESTING

2b1.1. What level of validity testing was conducted? (*may be one or both levels*)

☒ **Critical data elements** (*data element validity must address ALL critical data elements*)

☒ **Performance measure score**

☒ **Empirical validity testing**

☒ **Systematic assessment of face validity of performance measure score as an indicator** of quality or resource use (*i.e., is an accurate reflection of performance on quality or resource use and can distinguish good from poor performance*) **NOTE**: Empirical validity testing is expected at time of maintenance review; if not possible, justification is required.

2b1.2. 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)*

2019 Submission

Oncology: Medical and Radiation – Pain Intensity Quantified (PQRS #143/NQF #0384) was chosen as a suitable candidate for correlation analysis due to the similarities in patient population and domain. We hypothesize that there exists a positive association between patients with a diagnosis of cancer receiving chemotherapy or radiation therapy in which pain intensity is quantified (NQF # 0384) and those with a diagnosis of cancer receiving chemotherapy or radiation therapy who report having pain with a documented plan of care to address pain (PQRS #144/NQF #0383). Providers included in the analysis met the minimum number of quality reporting events (10) and were cleaned in the same process as the PQRS dataset.

Datasets were reviewed to identify shared providers based on NPI and TIN identifiers. Correlation analysis was then performed to evaluate the association between performance scores of these shared providers.

We use the following guidance to describe correlation[1]:

Correlation	Interpretation
> 0.40	Strong
0.20 - 0.40	Moderate
< 0.20	Weak

1. Shortell T. An Introduction to Data Analysis & Presentation. Sociology 712. http://www.shortell.org/book/chap18.html. Accessed July 13, 2018.

2015 Submission

All PCPI performance measures are assessed for content validity by a panel of expert work group members during the development process. Additional input on the content validity of draft measures is obtained through a 30-day public comment period and by also soliciting comments from a panel of consumer, purchaser, and patient representatives convened by the PCPI specifically for this purpose. All comments received are reviewed by the expert work group and the measures adjusted as needed. Other external review groups (eg, focus groups) may be convened if there are any remaining concerns related to the content validity of the measures.

The expert panel was used to assess face validity of the measure. This panel consisted of 31 members, with representation from the following specialties: oncology, radiation oncology, surgical oncology, urology, gastroenterology, hematology, pathology, colon and rectal surgery, otolaryngology, and pain medicine.

The aforementioned panel was asked to rate their agreement with the following statement:

> The scores obtained from the measure as specified will accurately differentiate quality across providers.

The scale 1-5, where 1=Strongly Disagree; 3=Neither Disagree nor Agree; 5=Strongly Agree

This panel consisted of the following 31 members, with representation from a number of specialties including oncology, radiation oncology, surgical oncology, urology, gastroenterology, hematology, pathology, colon and rectal surgery, otolaryngology, and pain medicine.

- Patricia Ganz, MD (Co-Chair) (Clinical Oncology) Los Angeles, CA

- James Hayman, MD (Co-Chair) (Radiation Oncology) Ann Arbor MI
- Joseph Bailes, MD (Clinical Oncology) The Woodlands, TX
- Nancy Baxter, MD, PhD (Colorectal Surgery) Toronto, Ontario Canada
- Joel V. Brill, MD (Gastroenterology) Phoenix, AZ
- Steven B. Clauser, PhD (Outcomes Research) Bethesda, MD
- Charles Cleeland, PhD (Oncology) Houston, TX
- J. Thomas Cross, Jr. MD, MPH (Oncology) Colorado Springs, CO
- Chaitanya R. Divgi, MD (Nuclear Medicine) Philadelphia, PA
- Stephen B. Edge, MD (Surgical Oncology) Buffalo, NY
- Patrick L. Fitzgibbons, MD (Oncology) Fullerton, CA
- Myron Goldsmith, MD (Oncology) Huntington Beach, CA
- Joel W. Goldwein, MD (Oncology) Merion Station, PA
- Alecia Hathaway, MD, MPH (Oncology) Fort Worth, TX
- Kevin P. Hubbard, DO (Oncology) Kansas City, MO
- Nora Janjan, MD, MPSA (Radiation Oncology) Houston, TX
- Maria Kelly, MB, BCh (Radiation Oncology) Earlysville, VA
- Wayne Koch, MD (Head and Neck surgery) Columbia, MD
- Andre Konski, MD (Radiation Oncology) Philadelphia, PA
- Len Lichtenfeld, MD (Oncology) Atlanta, GA
- Norman J. Marcus, MD (Anesthesiology and Psychiatry) New York, NY
- Catherine Miyamoto, RN, BSN (Oncology) Grand Forks, ND
- Michael Neuss, MD (Oncology, Hematology) Cincinnati, OH
- David F. Penson, MD, MPH (Urology) Nashville, TN
- Louis Potters, MD (Radiation Oncology) New Hyde Park, NY
- John M. Rainey, MD (Medical Oncology) Lafayette, LA
- Christopher M. Rose, MD (Radiation Therapy) Beverly Hills, El Segundo, CA
- Lee Smith, MD (Oncology) Washington, DC
- Lawrence A. Solberg, MD, PhD (Oncology) Jacksonville, FL
- Paul E. Wallner, MD (Radiation Oncology) Willingboro, NJ
- J. Frank Wilson, MD (Radiation Oncology) Milwaukee, WI

2b1.3. What were the statistical results from validity testing? (*e.g., correlation; t-test*)

2019 Submission

Oncology: Medical and Radiation – Plan of Care for Pain was positively correlated with Oncology: Pain Intensity Quantified – Medical Oncology and Radiation Oncology

- Coefficient of correlation = 0.69
- P-value = > 0.001

2015 Submission

The results of the expert panel rating of the validity statement were as follows: N = 19; Mean rating = 4.32.

Percentage in the top two categories (4 and 5): 89.47%

Frequency Distribution of Ratings

1- 0

2- 1

3- 1

4- 8

5- 9

2b1.4. What is 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?*)

2019 Submission

Oncology: Care Plan for Pain – Medical Oncology and Radiation Oncology has a strong positive correlation with another evidence-based process of care. The correlation demonstrates the criterion validity of the measure

2015 Submission

The measure is valid, as specified.

ASCO provided measure level concordance based on requirements for previous submission in 2015, which did not require individual data element validity. Given the high rate of agreement at the measure level, no individual data element would have had a low agreement rate.

Face validity testing demonstrated a vast majority of respondents (95%) strongly agree or agree that the measure provided an accurate reflection of quality and can be used to distinguish good and poor quality

---

2b2. EXCLUSIONS ANALYSIS

NA ☒ no exclusions — *skip to section 2b4*

2b2.1. Describe the method of testing exclusions and what it tests (*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*)

2b2.2. What were 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*)

2b2.3. What is your interpretation of the results in terms of demonstrating that exclusions are needed to prevent unfair distortion of performance results? (*i.e., 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*)

2b3. RISK ADJUSTMENT/STRATIFICATION FOR OUTCOME OR RESOURCE USE MEASURES

If not an intermediate or health outcome, or PRO-PM, or resource use measure, skip to section 2b5.

2b3.1. What method of controlling for differences in case mix is used?

☒ **No risk adjustment or stratification**

☐ **Statistical risk model with** Click here to enter number of factors **risk factors**

☐ **Stratification by** Click here to enter number of categories **risk categories**

☐ **Other,** Click here to enter description

2b3.1.1 If using a statistical risk model, provide detailed risk model specifications, including the risk model method, risk factors, coefficients, equations, codes with descriptors, and definitions.

2b3.2. If an outcome or resource use component measure is not risk adjusted or stratified, provide rationale and analyses to demonstrate that controlling for differences in patient characteristics (case mix) is not needed to achieve fair comparisons across measured entities.

2b3.3a. Describe the conceptual/clinical and statistical methods and criteria used to select patient factors (clinical factors or social risk factors) used in the statistical risk model or for stratification by risk (*e.g., potential factors identified in the literature and/or expert panel; regression analysis; statistical significance of p<0.10; correlation of x or higher; patient factors should be present at the start of care*) **Also discuss any "ordering" of risk factor inclusion**; for example, are social risk factors added after all clinical factors?

2b3.3b. How was the conceptual model of how social risk impacts this outcome developed? Please check all that apply:

☐ **Published literature**

☐ **Internal data analysis**

☐ **Other (please describe)**

2b3.4a. What were the statistical results of the analyses used to select risk factors?

2b3.4b. Describe the analyses and interpretation resulting in the decision to select social risk factors *(e.g. prevalence of the factor across measured entities, empirical association with the outcome, contribution of unique variation in the outcome, assessment of between-unit effects and within-unit effects.)* **Also describe the impact of adjusting for social risk (or not) on providers at high or low extremes of risk.**

2b3.5. 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 controlling for differences in patient characteristics (case mix) below.

If stratified, skip to 2b3.9

2b3.6. Statistical Risk Model Discrimination Statistics (*e.g., c-statistic, R-squared*)**:**

2b3.7. Statistical Risk Model Calibration Statistics (*e.g., Hosmer-Lemeshow statistic*):

2b3.8. Statistical Risk Model Calibration – Risk decile plots or calibration curves:

2b3.9. Results of Risk Stratification Analysis:

2b3.10. What is your interpretation of the results in terms of demonstrating adequacy of controlling for differences in patient characteristics (case mix)? (i.*e., what do the results mean and what are the norms for*

the test conducted)

2b3.11. Optional Additional Testing for Risk Adjustment (*not required, but would provide additional support of adequacy of risk model, e.g., testing of risk model in another data set; sensitivity analysis for missing data; other methods that were assessed*)

2b4. IDENTIFICATION OF STATISTICALLY SIGNIFICANT & MEANINGFUL DIFFERENCES IN PERFORMANCE

2b4.1. 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 related to performance gap in 1b)*

2019 Submission

The analysis of meaningful differences in performance was analyzed using calculations of several descriptive statistics, including the minimum, maximum, 25th and 75th percentile, median, IQR, and range. Additionally, we calculated the standard deviation, standard error of the mean performance, and confidence interval for the mean performance. Finally, we calculated the percent of facilities whose performance was statistically significantly different from the overall performance mean.

2015 Submission

PCPI Testing Project

Data analysis performed on the measure included:

Average measure performance rate overall and by site, performance rate range by site and overall standard deviation for the measure.

2b4.2. What were the statistical results from testing the ability to identify statistically significant and/or clinically/practically meaningful differences in performance measure scores across measured entities? (e.g., *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*)

2019 Submission

Although a lot perform at 100%, more than half do not, and performance spans the entire distance from 0% to 100%. Mean performance is 78.6% (95% confidence interval of 75% to 82%): that leaves plenty of room for improvement, and most fall outside of the 95% confidence interval, suggesting statistically significant differences between facilities.

Frequency (Unique Number of TINs): 366

Distribution of Measure Denominators and Measure Performance:

Denominators:

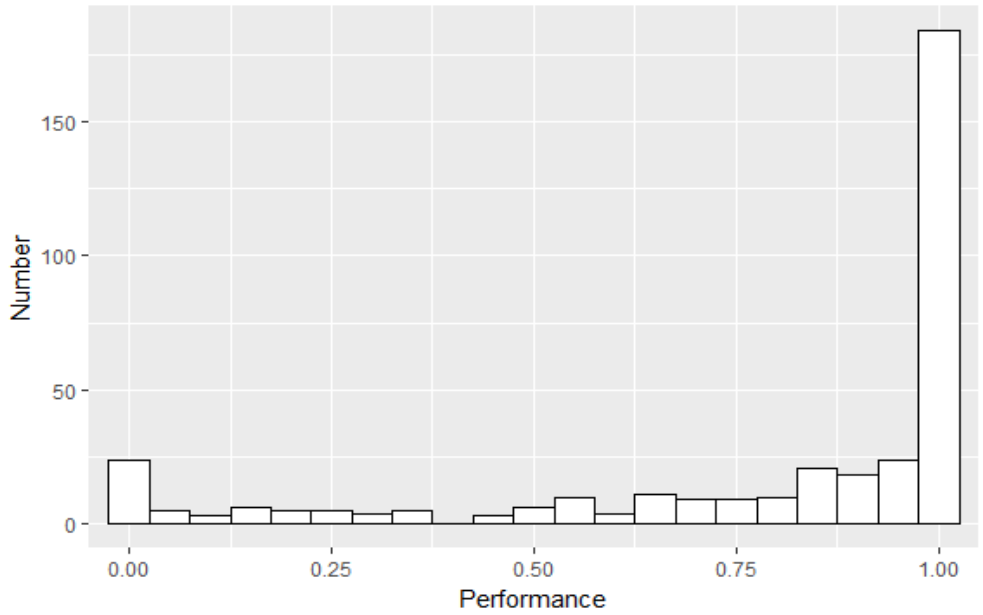
Min	Q1	Median	Mean	Q3	Max	Total

1	10	50.5	297.2	242	8731	108,765

Measure Distribution:

Min	Q1	Median	Mean	Q3	Max	CI.for.mean	Percent.outside.CI
0	0.6771	0.977	0.7864	1	1	(0.75, 0.82)	96.72

Measure Distribution:



2015 Submission

- PCPI Testing Project
- Average Measure performance rate without exceptions: N= 360 Mean = 88.3% Standard Deviation= 0.3215
- The performance rate by site is as follows, where n is the number of performance events by site:

A	**0.8940**	**n=94**
B	0.8870	n=62
C	0.8450	n=71
D	0.8540	n=89
E	0.9770	n=44

2b4.3. What is 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? (i.*e., what do the results mean in terms of statistical and meaningful differences?*)

2019 Submission

Interpretation: Although a lot perform at 100%, more than half do not, and performance spans the entire distance from 0% to 100%. Mean performance is 78.6% (95% confidence interval of 75% to 82%): that leaves plenty of room for improvement, and most fall outside of the 95% confidence interval, suggesting statistically significant differences between facilities.

2015 Submission

The performance rate range is .1320. Although this study captured performance on 360 events, the data were not captured at the physician level, restricting reporting of variation in performance to the organization level only. Additionally, we are unable to present a meaningful calculation of variation in performance across organizations due to the small sample size of sites (n=5) in this study.

2b5. COMPARABILITY OF PERFORMANCE SCORES WHEN MORE THAN ONE SET OF SPECIFICATIONS

If only one set of specifications, this section can be skipped.

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 eMeasures). 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.***

2b5.1. 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; what statistical analysis was used*)

2b5.2. What were the statistical results from testing comparability of performance scores for the same entities when using different data sources/specifications? (*e.g., correlation, rank order*)

2b5.3. What is your interpretation of the results in terms of the differences in performance measure scores for the same entities across the different data sources/specifications? (i.*e., what do the results mean and what are the norms for the test conducted*)

2b6. MISSING DATA ANALYSIS AND MINIMIZING BIAS

2b6.1. Describe the method of testing conducted to 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 nonresponders) and how the specified handling of missing data minimizes bias (*describe the steps—do not just name a method; what statistical analysis was used*)

2019 Submission

The PQRS/MIPS dataset provided to us by CMS did not contain missing data so this test was not performed. Due to data completeness requirements, we suspect that missing data would have been rejected when submitted to CMS, in which case those values would not be counted towards measure performance. While data that may have been missing prior to submission to CMS is unknown and therefore precluded any analysis, there is no indication that this missing data was systematic, thus their omission would lead to unbiased performance results.

2015 Submission

There are no missing data to report on this measure

2b6.2. What is the overall frequency of missing data, the distribution of missing data across providers, and the results from testing related to missing data? (*e.g., results of sensitivity analysis of the effect of various rules for missing data/nonresponse; if no empirical sensitivity analysis, identify the approaches for handling missing data that were considered and pros and cons of each*)

2019 Submission

This test was not performed for this measure as there was no missing data.

2015 Submission

Not Applicable

2b6.3. What is 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 nonresponders) and how the specified handling of missing data minimizes bias**?** (i.*e., 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, provide rationale for the selected approach for missing data*)

2019 Submission

This test was not performed for this measure as there was no missing data.

3. Feasibility

3a. Byproduct of Care Processes

For clinical measures, the required data elements are routinely generated and used during care delivery (e.g., blood pressure, lab test, diagnosis, medication order).

3a.1. Data Elements Generated as Byproduct of Care Processes.

Generated or collected by and used by healthcare personnel during the provision of care (e.g., blood pressure, lab value, diagnosis, depression score), Abstracted from a record by someone other than person obtaining original information (e.g., chart abstraction for quality measure or registry)

If other:

3b. Electronic Sources

The required data elements are available in electronic health records or other electronic sources. If the required data are not in electronic health records or existing electronic sources, a credible, near-term path to electronic collection is specified.

3b.1. To what extent are the specified data elements available electronically in defined fields (*i.e., data elements that are needed to compute the performance measure score are in defined, computer-readable fields*) Update this field for **maintenance of endorsement**.

Some data elements are in defined fields in electronic sources

3b.2. 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 other than electronic sources. For **maintenance of endorsement**, if this measure is not an eMeasure (eCQM), please describe any efforts to develop an eMeasure (eCQM).

All the data elements needed for this measure are collected through electronic data or through the use of keyword searches. ASCO is in the process of assessing the feasibility of developing an electronic clinical quality measure.

3b.3. If this is an eMeasure, provide a summary of the feasibility assessment in an attached file or make available at a measure-specific URL. Please also complete and attach the NQF Feasibility Score Card.

Attachment:

3c. Data Collection Strategy

Demonstration that the data collection strategy (e.g., source, timing, frequency, sampling, patient confidentiality, costs associated with fees/licensing of proprietary measures) can be implemented (e.g., already in operational use, or testing demonstrates that it is ready to put into operational use). For eMeasures, a feasibility assessment addresses the data elements and measure logic and demonstrates the eMeasure can be implemented or feasibility concerns can be adequately addressed.

3c.1. Required for maintenance of endorsement. 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.

IF instrument-based, consider implications for both individuals providing data (patients, service recipients, respondents) and those whose performance is being measured.

Apart from the lack of availability of disparities data for analyses, we have not identified any areas of concern or made any modifications as a result of testing and operational use of the measures in relation to data collection, availability of data, missing data, timing and frequency of data collection, sampling, patient confidentiality, time and cost of data collection, and other feasibility issues unless otherwise noted.

3c.2. Describe 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)*.

ASCO requests interested parties seek a licensing agreement prior to commercial use of this measure.

4. Usability and Use

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

4a. Accountability and Transparency

Performance results are used in at least one accountability application within three years after initial endorsement and are publicly reported within six years after initial endorsement (or the data on

performance results are available). If not in use at the time of initial endorsement, then a credible plan for implementation within the specified timeframes is provided.

4.1. Current and Planned Use

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

Specific Plan for Use	Current Use (for current use provide URL)
Public Reporting	Payment Program Merit-based Incentive Payment System (MIPS) https://qpp.cms.gov/mips/quality-measures ASCO Qualified Clinical Data Registry https://practice.asco.org/sites/default/files/drupalfiles/QCDR-2019-Measure-Summary.pdf

4a1.1 For each CURRENT use, checked above (update for maintenance of endorsement), provide:

- Name of program and sponsor
- Purpose
- Geographic area and number and percentage of accountable entities and patients included
- Level of measurement and setting

Merit-based Incentive Payment System (MIPS)-Sponsored by the Centers for Medicare and Medicaid Services (CMS) ?Prior to 2016, this measure was used for Eligible Providers (EPs) in the Physician Quality Reporting System (PQRS). As of 2017, MIPS replaced the PQRS program. MIPS is a national performance-based payment program that uses performance scores across several categories to determine payment rates for EPs. MIPS takes a comprehensive approach to payment by basing consideration of quality on a set of evidence-based measures that were primarily developed by clinicians, thus encouraging improvement in clinical practice and supporting advances in technology that allow for easy exchange of information. Data on geographic area and number and percentage of accountable entities and patients, including level of measurement and setting, are unavailable for analysis.

QOPI® Qualified Clinical Data Registry

This measure has been reported to CMS by the registry as a Qualified Clinical Data Registry. The Quality Oncology Practice Initiative (QOPI®) was deemed as a registry for oncology measures group reporting and as a QCDR to report to PQRS in 2015 and 2016 and to report to MIPS in 2017, 2018 and 2019. Eligible professionals will be considered to have satisfactorily participated in MIPS if they submit quality measures data or results to CMS via a qualified clinical data registry. In 2017 and 2018, a total of 19 practices representing approximately 50,000 patient charts submitted to MIPS through QOPI. CMS has implemented a phased approach to public reporting performance information on the Physician Compare website.

4a1.2. If not currently publicly reported OR used in at least one other accountability application (e.g., payment program, certification, licensing) what are the reasons? (*e.g., Do policies or actions of the developer/steward or accountable entities restrict access to performance results or impede implementation?*)

According to the CY 2019 Quality Payment Program final rule, Physician Compare has continued to pursue a phased approach to public reporting under MACRA. CMS intends to make all measures under MIPS quality performance category available for public reporting on Physician Compare. These measures include those reported via all available submission methods for MIPS-eligible clinicians and groups. Because this measure has been in use for at least one year and meets the minimum sample size requirement for reliability, this measure meets criteria for public reporting but has not yet been included in Physician Compare.

4a1.3. If not currently publicly reported OR used in at least one other accountability application, provide a credible plan for implementation within the expected timeframes -- any accountability application within 3 years and publicly reported within 6 years of initial endorsement. (*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.)

Despite not yet being included in Physician Compare, this measure meets criteria for public reporting because it has been in use for at least one year and meets the minimum sample size requirement for reliability, this measure meets criteria for public reporting

4a2.1.1. Describe how performance results, data, and assistance with interpretation have been provided to those being measured or other users during development or implementation.

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.

ASCO's measure development process is rigorous, evidence-based, and utilizes the clinical expertise of multiple standing multi-disciplinary Technical Expert Panels (TEPs) dedicated to development and maintenance of measures across the cancer continuum. During measure maintenance, TEP members are provided with full measure specifications, applicable evidence, historical measure performance data, and any external feedback or requests for clarification or updates that have been received for the measure.

Staff on ASCO's measure development team are available to receive comments and questions from measure implementers and clinicians reporting the measures. As comments and questions are received, they are shared with appropriate staff for follow up. If comments or questions require expert input, these are shared with ASCO's TEPs to determine if measure modifications may be warranted. Additionally, for ASCO measures included in federal reporting programs, there is a system that has been established to elicit timely feedback and responses from ASCO staff in consultation with TEP members, as appropriate.

4a2.1.2. Describe the process(es) involved, including when/how often results were provided, what data were provided, what educational/explanatory efforts were made, etc.

See description in 4a2.1.1 above.

4a2.2.1. Summarize the feedback on measure performance and implementation from the measured entities and others described in 4d.1.

Describe how feedback was obtained.

See description in 4a2.1.1 above.

4a2.2.2. Summarize the feedback obtained from those being measured.

In addition to the feedback obtained from a multi-disciplinary technical expert panel during the measure development and maintenance process, ASCO obtains feedback and receives measure inquiries from implementers and reporters via email. No specific feedback has been received by ASCO on this measure other than recommendations to clarify the denominator population.

4a2.2.3. Summarize the feedback obtained from other users

No additional feedback has been received by ASCO on this measure. However, we will continue to solicit feedback as we perform maintenance on this measure.

4a2.3. Describe how the feedback described in 4a2.2.1 has been considered when developing or revising the measure specifications or implementation, including whether the measure was modified and why or why not.

Feedback received on how users have implemented the measure indicated there was an inconsistent approach to applying the measure criteria. To address this feedback, this measure was modified in 2019 to have two populations based on the type of treatment the patient is receiving. This change was made to more clearly delineate the denominator requirements to promote accurate implementation.

Improvement

Progress toward achieving the goal of high-quality, efficient healthcare for individuals or populations is demonstrated. If not in use for performance improvement at the time of initial endorsement, then a credible

rationale describes how the performance results could be used to further the goal of high-quality, efficient healthcare for individuals or populations.

4b1. Refer to data provided in 1b 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, what are the reasons? 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.

Analysis of the MIPS data from 2015 to 2017 demonstrates that practices continue to improve in the frequency with which patients who report pain have a plan of care in place. For example, the highest average (mean) performance rate was in 2016 with just under 90% performance and the third quartile (75th) was consistently reported at 100% across the three years. Interestingly, the average score was 75% in 2017 with double the number of practices and more than 70,000 additional patients, indicating that improvement in this process is possible but still needed.

4b2. Unintended Consequences

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).

4b2.1. Please explain any unexpected findings (positive or negative) during implementation of this measure including unintended impacts on patients.

At this time, we are not aware of any unintended consequences related to this measure. We take unintended consequences very seriously and therefore continuously monitor to identify actions that can be taken to mitigate them.

4b2.2. Please explain any unexpected benefits from implementation of this measure.

We have not observed any unexpected benefits associated with implementation of this measure.

5. Comparison to Related or Competing Measures

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

5. Relation to Other NQF-endorsed Measures

Are there related measures (conceptually, either same measure focus or target population) or competing measures (conceptually both the same measure focus and same target population)? If yes, list the NQF # and title of all related and/or competing measures.

Yes

5.1a. List of related or competing measures (selected from NQF-endorsed measures)

0420 : Pain Assessment and Follow-Up

1628 : Patients with Advanced Cancer Screened for Pain at Outpatient Visits

5.1b. If related or competing measures are not NQF endorsed please indicate measure title and steward.

5a. Harmonization of Related Measures

The measure specifications are harmonized with related measures;

OR

The differences in specifications are justified

5a.1. If this measure conceptually addresses EITHER the same measure focus OR the same target population as NQF-endorsed measure(s):
Are the measure specifications harmonized to the extent possible?
Yes
5a.2. If the measure specifications are not completely harmonized, identify the differences, rationale, and impact on interpretability and data collection burden.
Measure #420 is broadly applicable to any patients 18 years of age and older using claims. Measure #383 is examines whether a plan of care is present and maintained for a population who frequently experience pain – a population in which adequate pain management is crucial. In addition, it uses registry data in addition to paper medical records. Measure #1628 targets only patients with Stage IV cancer. Our measure looks at any stage of cancer for purposes of managing pain for which chemotherapy or radiation may be appropriate.
5b. Competing Measures

The measure is superior to competing measures (e.g., is a more valid or efficient way to measure);
OR
Multiple measures are justified.

5b.1. If this measure conceptually addresses both the same measure focus and the same target population as NQF-endorsed measure(s):
Describe why this measure is superior to competing measures (e.g., a more valid or efficient way to measure quality); OR provide a rationale for the additive value of endorsing an additional measure. (Provide analyses when possible.)
An environmental scan did not identify competing measures.

Appendix

A.1 Supplemental materials may be provided in an appendix. All supplemental materials (such as data collection instrument or methodology reports) should be organized in one file with a table of contents or bookmarks. If material pertains to a specific submission form number, that should be indicated. Requested information should be provided in the submission form and required attachments. There is no guarantee that supplemental materials will be reviewed.

No appendix **Attachment:**

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): American Society of Clinical Oncology

Co.2 Point of Contact: Angela, Kennedy, angela.kennedy@asco.org, 571-483-1656-

Co.3 Measure Developer if different from Measure Steward: American Society of Clinical Oncology

Co.4 Point of Contact: Angela, Kennedy, angela.kennedy@asco.org, 571-483-1656-

Additional Information

Ad.1 Workgroup/Expert Panel involved in measure development

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

Michael Hassett, MD, MPH; Measure Panel Chair;

Dana Farber Cancer Institute Boston, MA

Laura Chow, MD; Panel Member;

University of Texas; Austin, TX

Kristen Fessele, PhD, RN, ANP-BC, AOCN;

Panel Member; Memorial Sloan Kettering Cancer Center NYC, NY

Jennifer J. Griggs, MD, MPH, FASCO;

Panel Member; University of Michigan Ann Arbor, MI

Bonnie Labdi, PharmD, BCOP, RPh;

Panel Member; Memorial Hermann Houston, TX

Merry Jennifer Markham, MD, FACP;

Panel Member; University of Florida Gainesville, FL

Michael Soble, MD; Panel Member;

North Shore Oncology Chicago, IL

Jessica Zerillo, MD, MPH; Panel Member;

Beth Israel Deaconess Medical Center Boston, MA

Caitlin Drumheller; Measure Development Specialist;

American Society of Clinical Oncology Alexandria, VA

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2007

Ad.3 Month and Year of most recent revision: 04, 2019

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

Ad.5 When is the next scheduled review/update for this measure? 04, 2020

Ad.6 Copyright statement: The Measure is not clinical guidelines, does not establish a standard of medical care, and has not been tested for all potential applications.

The Measure, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measures for commercial gain, or incorporation of the Measure into a product or service that is sold, licensed or distributed for commercial gain.

Commercial uses of the Measure require a license agreement between the user and the American Society of Clinical Oncology (ASCO) and American Medical Association (AMA), [on behalf of the Physician Consortium for Performance Improvement® (PCPI®)] and prior written approval of ASCO, AMA, or PCPI. Neither ASCO, AMA, or PCPI, nor its members shall be responsible for any use of the Measures.

The AMA's and PCPI's significant past efforts and contributions to the development and updating of the Measure is acknowledged. ASCO is solely responsible for the review and enhancement ("Maintenance") of the Measures as of January 2015.

ASCO encourages use of the Measures by other health care professionals, where appropriate.

THE MEASURES AND SPECIFICATIONS ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND.

Limited proprietary coding is contained in the Measure specification for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. ASCO, AMA, , PCPI and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT®) or other coding contained in the specification.

Ad.7 Disclaimers: See copyright statement in Ad.6 above.

Ad.8 Additional Information/Comments: