

MEASURE WORKSHEET

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

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

Brief Measure Information

NQF #: 3490

Corresponding Measures:

De.2. Measure Title: Admission and Emergency Department (ED) Visits for Patients Receiving Outpatient Chemotherapy

Co.1.1. Measure Steward: Centers for Medicare and Medicaid Services (CMS)

De.3. Brief Description of Measure: The Admission and Emergency Department (ED) Visits for Patients Receiving Outpatient Chemotherapy Measure, hereafter referred to as the chemotherapy measure, estimates hospital-level, risk-adjusted rates of inpatient admissions or ED visits for cancer patients ≥ 18 years of age for at least one of the following diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of hospital-based outpatient chemotherapy treatment. Rates of admission and ED visits are calculated and reported separately.

1b.1. Developer Rationale: The primary purpose of this measure is to assess the extent to which cancer patients receiving outpatient chemotherapy treatment experience complications resulting in a hospital visit (either an inpatient admission or ED visit). By identifying these events, the measure seeks to encourage quality improvement across facilities to reduce the number of potentially avoidable inpatient admissions and ED visits and increase transparency in the quality of care patients receive. The measure is envisioned to promote effective communication and coordination of care, which is both a Meaningful Measures quality category and a National Quality Strategy priority. It also meets an additional National Quality Strategy priority of promoting the most effective prevention and treatment practices for the leading causes of mortality.

Chemotherapy treatment can have severe, predictable side effects, which, if inappropriately managed, can reduce patients' quality of life and increase healthcare utilization and costs. On average, cancer patients receiving chemotherapy have one hospital admission and two ED visits per year; approximately 40 percent of these admissions, and 50 percent of these ED visits stem from complications of chemotherapy, respectively [1]. The literature suggests that ten symptoms in particular—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—are primary reasons for hospital visits among cancer patients receiving chemotherapy, and are potentially avoidable with proper outpatient management [3 - 5]. Improved management of these symptoms, through improved adherence to clinical treatment guidelines and enhanced care coordination, has been shown to reduce admissions and ED visits and increase patients' quality of care and quality of life [2] [3] [4].

Admissions and ED visits are costly to payers, with one study estimating that, on average, those experiencing chemotherapy-related adverse events incurred \$12,907 in additional hospitalization expenditures per person per year [6]. In addition to increased cost to payers, unplanned admissions and ED visits related to chemotherapy treatment reduce cancer patients' quality of life. Measuring potentially avoidable admissions

and ED visits for cancer patients receiving outpatient chemotherapy will provide hospitals with an incentive to improve the quality of care for these patients, by taking steps to prevent and better manage side effects and complications from treatment. Hospitals that provide outpatient chemotherapy should implement appropriate care to minimize the incidence of these adverse events and the subsequent need for acute hospital care.

Evidence suggests that coordination of care and better management of these symptoms in the outpatient setting can decrease hospital visits among patients receiving chemotherapy. Studies have indicated that in outpatient settings, where established guidelines are not properly followed and structured protocols are not put into place, there is a higher likelihood for adverse events [7] [8] [9]. This measure will encourage hospitals to use guidelines from the American Society of Clinical Oncology, National Comprehensive Cancer Network, Oncology Nursing Society, Infectious Diseases Society of America, and other professional societies with evidence-based interventions to prevent and treat common side effects and complications of chemotherapy [10]. This risk-standardized measure seeks to increase transparency in the quality of care patients receive, and to provide information to help physicians and hospitals mitigate patients' need for acute care, which can be a burden on patients, and increase patients' quality of life [11 – 12].

Citations

1. Vandervelde A, Miller H, Younts J. Impact on Medicare payments of shift in site of care for chemotherapy administration. Washington, DC: Berkeley Research Group; June 2014.
http://www.communityoncology.org/UserFiles/BRG_340B_SiteofCare_ReportF_6-9-14.pdf. Accessed September 16, 2015
2. Burke, T.A., Wisniewski, T., and Ernst, F.R. "Resource utilization and costs associated with chemotherapy-induced nausea and vomiting (CINV) following highly or moderately emetogenic chemotherapy administered in the US outpatient hospital setting." *Supportive Care in Cancer*, vol. 19, 2011, pp. 131-141.
3. McKenzie, H., L. Hayes, K. White, K. Cox, J. Fethney, M. Boughton, and J. Dunn. "Chemotherapy Outpatients' Unplanned Presentations to Hospital: A Retrospective Study." *Support Care Cancer*, vol. 19, 2011, pp. 963–969
4. Foltran, L., Aprile, G., Pisa, F.E. et al. "Risk of unplanned visits for colorectal cancer outpatients receiving chemotherapy: a case-crossover study." *Supportive Care in Cancer*, September 2014, Volume 22, Issue 9, PP 2527-2533.
5. Daly, B. and Abougergi, M.S. "National trends in admissions for potentially preventable conditions among patients with metastatic solid tumors, 2004-2014." *Journal of Clinical Oncology* 36, no. 30_suppl (October 20 2018) 1-1.
6. Hassett, M.J., J. O'Malley, J.R. Pakes, J.P. Newhouse, and C.C. Earle. "Frequency and Cost of Chemotherapy-Related Serious Adverse Effects in a Population Sample of Women with Breast Cancer." *Journal of the National Cancer Institute*, vol. 98, no. 16, 2006, pp. 1108–1117.
7. Abernethy, A.P., Barbour, S.Y., Uronis, H. et al. "Quality management of potential chemotherapy-induced neutropenic complications: evaluation of practice in an academic medical center." *Support Care Cancer* (2009) 17: 735
8. Molassiotis, A., Brearley, S.G. & Stamatakis, Z. "Use of antiemetics in the management of chemotherapy-related nausea and vomiting in current UK practice." *Support Care Cancer* (2011) 19: 949.
9. Patil, Vijay M., Vanita Noronha, Amit Joshi, Anant Ramaswamy, Sudeep Gupta, Arvind Sahu, Vipul Doshi, Tarachand Gupta, Sushmita Rath, Shripad Banavali, and Kumar Prabhash. "Adherence to and Implementation of ASCO Antiemetic Guidelines in Routine Practice in a Tertiary Cancer Center in India." *Journal of Oncology Practice*, vol. 13, no. 6, 2017, pp. e574-e581
10. National Comprehensive Cancer Network. "Prevention and Treatment of Cancer-Related Infections." NCCN Clinical Practice Guidelines in Oncology Version 1.2013. 2013. Available at http://www.nccn.org/professionals/physician_gls/f_guidelines.asp. Accessed Sept. 26, 2014.

11. Osoba, D., B. Zee, D. Warr, J. Latreille, L. Kaizer, and J. Pater. "Effect of Postchemotherapy Nausea and Vomiting on Health-Related Quality of Life." *Support Care Cancer*, vol. 5, 1997, pp. 307–313.

12. Wu, H.S., T. Natavio, J.E. Davis, and H.N. Yarandi. "Pain in Outpatients Treated for Breast Cancer: Prevalence, Pharmacological Treatment, and Impact on Quality of Life." *Cancer Nursing*, vol. 36, no. 3, 2013, pp. 229–235. Available at <http://www.cancernursingonline.com/>. Accessed September 24, 2012.

S.4. Numerator Statement: This measure involves calculating two mutually exclusive outcomes among cancer patients receiving chemotherapy treatment in a hospital outpatient setting: (1) one or more inpatient admissions for any of the following 10 diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of chemotherapy treatment or (2) one or more ED visits for any of the following 10 diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of chemotherapy treatment. These 10 conditions are potentially preventable through appropriately managed outpatient care. To be counted as an outcome, the qualifying diagnosis on the admission or ED visit claim must be (1) the principal diagnosis or (2) a secondary diagnosis accompanied by a principal diagnosis of cancer.

S.6. Denominator Statement: The measure cohort includes Medicare Fee-for-Service (FFS) patients, aged 18 years and older at the start of the performance period, with a diagnosis of any cancer (except leukemia), who received at least one outpatient chemotherapy treatment at the reporting hospital during the performance period.

S.8. Denominator Exclusions: The measure excludes the following patients from the cohort:

- 1) Patients with a diagnosis of leukemia at any time during the performance period.
- 2) Patients who were not enrolled in Medicare FFS Parts A and B in the year prior to the outpatient chemotherapy treatment during the performance period.
- 3) Patients who were not enrolled in Medicare FFS Parts A and B for the 30 days following any chemotherapy treatment.
- 4) Cases in which patients receive chemotherapy to treat conditions other than cancer. Note that this is a case-level exclusion; as long as the patient has additional cases that meet inclusion criteria, they will remain in the cohort.

De.1. Measure Type: Outcome

S.17. Data Source: Claims, Enrollment Data

S.20. Level of Analysis: Facility

IF Endorsement Maintenance – Original Endorsement Date: Most Recent Endorsement Date:

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

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

De.4. IF PAIRED/GROUPED, what is the reason this measure must be reported with other measures to appropriately interpret results? Not applicable.

Preliminary Analysis: New Measure

Criteria 1: Importance to Measure and Report

1a. [Evidence](#)

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

Evidence Summary

- Timely access to chemotherapy side effect management leads to decreased likelihood of preventable admissions and ED visits for patients receiving outpatient chemotherapy. Data show that improved symptom management and coordination of care reduces hospital visits.
- Admissions and ED visits for the ten diagnoses captured in the measure—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—are among the most common reasons that cancer patients receiving chemotherapy visit the hospital. Treatment plans and guidelines exist to support the management of these conditions.

Question for the Committee:

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

Guidance from the Evidence Algorithm

Health outcome measure→The relationship between the outcome and at least one healthcare action is demonstrated by empirical data→Pass

Preliminary rating for evidence: ☒ Pass ☐ No Pass

1b. [Gap in Care/Opportunity for Improvement](#) and 1b. [Disparities](#)

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

- The developer provided performance data from national Medicare Fee-for-Service (FFS) claims and enrollment data for short-term acute hospitals using a period of performance of October 1, 2015 to September 30, 2016.
 - The risk-standardized inpatient admission rate (RSAR) for non-cancer hospitals ranged from 8.9% to 18.5% (median 12.5%, 25th and 75th percentiles are 12.2% and 13.0%, respectively) while the risk-standardized inpatient admission rate for PCHs ranged from 12.3% to 15.2% (median 13.7%, 25th and 75th percentiles are 13.4% and 14.8%, respectively).
 - The risk-standardized ED visit rate (RSEDR) for non-cancer hospitals ranged from 2.9% to 15.2% (median 5.6%, 25th and 75th percentiles are 5.6% and 6.2%, respectively) while the risk-standardized ED visit rate for PCHs ranged from 3.6% to 9.1% (median 6.7%, 25th and 75th percentiles are 4.4% and 8.9%, respectively).
 - The developer also provided distributions of facility scores (RSARs for non-cancer and cancer hospitals, RSEDRs for cancer and non-cancer hospitals).

Disparities

- The developer did not provide disparities data from the measure as specified but did examine associations between outcomes and social risk factors. The developer evaluated two indicators of social risk for impact on the measure score: race, specifically African American or not; and the Agency for Healthcare Research and Quality (AHRQ) Socio-Economic Status (SES) Composite index.
- At the patient level, the developer found that black patients are more likely to have an inpatient admission or ED visit than non-black patients and low AHRQ SES Composite Index patients are more likely to have an inpatient admission or ED visit than higher SES Composite Index patients.

- At the hospital level there was no significant impact of disparities on hospital-level measure scores.

Questions for the Committee:

- Is there a gap in care that warrants a national performance measure?
- Are you aware of evidence that disparities exist in this area of healthcare?

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

Committee Pre-evaluation Comments:

Criteria 1: Importance to Measure and Report (including 1a, 1b, 1c)

Evidence (*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.”*)

- Because a variety of adverse events are considered together and the measure uses risk-adjustment, while individual adverse events may have some evidence relating them to potentially effective interventions, it seems like a stretch to presume that lumping them together can produce a metric that identifies potentially useful interventions.
- This is a CMS, claims-based outcome measure. Forty-one references were provided to address one or more of the 10 ED visit or inpatient admission diagnoses experienced by patients receiving chemotherapy for cancer (within 30 days of a treatment). This is intended to assess the extent to which cancer patients receiving outpatient chemotherapy treatment experience complications resulting in a hospital visit (inpatient admission or ED visit). The ultimate goal is to decrease ED visits and inpatient admissions that could be prevented through prevention and management of side effects (e.g., preventing nausea, vomiting, dehydration). Treatment plans and guidelines currently exist to address the 10 complications.
- Well
- Evidence relates to the outcomes measure and directly applies; if the 10 AE's are managed by clinicians appropriately and proactively according to NCCN, ASCO, ASH guidelines, this should in turn risk mitigate hospital admissions and ER visits due to the 10 AE's for patients receiving hospital outpatient IV chemo. Suggest that refs also include the most recent NCCN supportive care guidelines for all 10 AE's. Suggest to include new 2018 NCCN immune related toxicities as ref to support as well.
- There is adequate literature support provided by the developers to suggest that the complications being addressed by this measure are relevant to the medical evidence. This measure addresses a significant clinical problem and is an outcome measure. Identification and management of these symptoms would result in improved cancer care.
- Evidence provided directly relates to the outcome being measured (admission and ED visits).
- Indirectly
- Direct evidence available
- I believe the data directly relates to the outcome to be measured.

Performance Gap (*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?*)

- When the measure was applied, resulting statistics indicated different levels of performance. However, since the true quality is unknown (i.e. was the lowest scoring hospital truly bad?) it isn't clear whether the supposed gap indicates true differences in performance or is simply an inevitable result of applying the statistical algorithm.
- Performance data from Medicare FFS claims and enrollment data were used for a period of October 1, 2015 to September 30, 2016. Data from non-cancer and cancer hospitals were reported (inpatient admissions were

reported separately from ED visits). The percentages indicate a less than optimal performance in terms of ED visits and inpatient admissions for diagnoses that, for the most part, could be addressed via outpatient interventions (e.g., navigator/symptom manager, improved patient education about prevention complications or managing side effects following chemotherapy treatment). Disparities data were not reported, but two indicators of social risk and the impact on the measure score were evaluated (i.e., race - AA and AHRQ SES Composite Index). They noted that African American patients were more likely to have ED visits or inpatient admissions and those with lower SES scores were more likely to have inpatient admissions or ED visits compared to those with higher SES or non-AA patients.

- Unclear if gap is due to provider decision making or increased resources available to oncologists practicing in a large outpatient center than can manage many of these diagnoses in the outpatient setting. solo practitioners will have to rely more heavily on the ED and hospital to provide the same services.
- Yes current performance data provided; demonstrates performance gap based on 1 year medicare data and opportunity for improvement in both cancer and non cancer hospitals based on inclusion and exclusion criteria of outcomes measure. The developer did not provide disparities data from the measure as specified but did examine associations between outcomes and social risk factors and reported that black patients are more likely to have an inpatient admission or ED visit than non-black patients and low AHRQ SES Composite Index patients are more likely to have an inpatient admission or ED visit than higher SES Composite Index patients.
- The developers demonstrate a gap in performance with a wide range of inpatient admissions and ED visits both in non-cancer hospitals and cancer hospitals. They also demonstrated a performance gap for blacks vs. other and low SES vs. other. Data presented warrants the use of a national performance measure.
- Performance data provided on the measure demonstrating a gap in care. Disparity data directly related to the measure was not provided but the relationship to social factors was examined related to likelihood of admission.
- Provided; modest disparities; provided across some subgroups, but minimized.
- There is a performance gap
- Performance data was provided and it does demonstrate a gap in care (patients taking out patient chemotherapy having to be admitted to ED). I guess there could be a question on disparities in health care and how that affects the data. Data on disparities was provided but I am not clear on how it might impact the measures effectiveness.

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.

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

Evaluators: Methods Panel Subgroup #3

[Subgroup #3 Combined Preliminary Analysis](#)

Evaluation of Reliability and Validity:

This measure was reviewed the Methods Panel. The measure passed on reliability and validity. A summary of the measure and Standing Committee Action Items are provided below:

Reliability

- Reliability was conducted at the measure score-level.
 - Score-level reliability was demonstrated in two ways: Signal-to-noise (SNR) ratio using Adams method and via a split-sample ICC (2,1)
 - NOTE: For both signal-to-noise and split-sample, testing was limited to hospitals with at least 25 and 50 patients, respectively. As such, testing was not consistent with the measure's specifications.
 - **Signal-to-noise results:**
 - Cancer hospitals (n=11): Admissions measure median reliability=0.7848; ED measure median reliability=0.9808
 - Non-cancer hospitals (n=1,524): Admissions measure median reliability=0. 6027; ED measure median reliability=0. 7326
 - **Split-sample results:**
 - Cancer hospitals (n=11): Admissions measure ICC=0.6704; ED measure ICC=0.8904
 - Non-cancer hospitals (n=1,099): Admissions measure ICC=0. 4314; ED measure ICC=0. 3585

Validity

- Validity was conducted via face validity.
- The developer conducted various assessments to demonstrate face validity. Only the assessment by the 2018 Expert Workgroup (EWG) meets NQF's requirements for face validity.
- The Methods Panel had questions about the risk-adjustment approach, which NQF staff clarified based on the testing results (e.g. the definition of concurrent radiology). Staff highlighted the developer's extensive analysis and discussion about consideration of social risk factors and reiterated that the inclusion, or lack of, specific risk-factors should not be a reason to reject a measure.
- **Meaningful differences results:**
 - Cancer hospitals (n=11), admission measure: 1 identified as performing significantly better than the national rate
 - Cancer hospitals (n=11), ED measure: 3 identified as performing significantly better than the national rate; 3 identified as performing significantly worse than the national rate
 - Non-cancer hospitals (n=3,562), admission measure: 13 identified as performing significantly better than the national rate; 65 identified as performing significantly worse than the national rate

Non-cancer hospitals (n=3,562), ED measure: 26 identified as performing significantly better than the national rate; 33 identified as performing significantly worse than the national rate

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 Scientific Methods Panel is 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 Scientific Methods Panel is 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

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

Reliability- Specifications (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?)

- The fact that different factors play into the risk adjustment depending upon whether you're looking at ED visit or inpatient admission, gives one the impression that the risk adjustment was done on a purely statistical basis without any indication that the factor used for risk adjustment predictably modify rates of the adverse effects addressed here.
- The Scientific Methods Panel reviewed this measure and had some diversity in responses. This is a complex measure to evaluate. The developers tested the reliability of the facility by calculating the ICC of the measure score using a split-sample (test-retest) method. Signal to noise results were also calculated with median reliability in PCH as 0.78 and ED = 0.98.
- high reliability
- The specifications are measuring 2 separate outcomes, admissions and ER visits 30 days post IV HOPD chemo yet they are combined in the numerator of 1 measure. Concerns about consistently implementing how the data will be captured as an admission or ER visit across hospital systems. While it appears simple in that the developer explained that 'Outcomes are counted separately for the inpatient admission and ED visit categories; a patient can only qualify for an outcome in either category, but not both. Patients who experience both an inpatient admission and an ED visit during the performance period are counted towards the inpatient admission outcome. Among those with no qualifying inpatient admissions, qualifying ED visits are counted. As a result, the rates can be viewed as additive to provide a comprehensive performance estimate of quality of care following hospital-based outpatient chemotherapy treatment.'. Regarding the denominator exclusion criteria, Patients with a diagnosis of leukemia at any time during the performance period- can the developer explain the clinical rationale more clearly with citations? If this is based on not being able to decipher if the chemo caused the AE or the underlying disease state, then additional hematologic malignancies may also be considered for exclusion criteria such as multiple myeloma and other lymphomas.
- The numerator and denominator statements are clear and well-defined. The exclusions are well-described and appropriate for the measure and for the clinical scenarios. One limitation that was described by the developers was the exclusion of oral chemotherapy. While difficult to obtain this data due to the nature of the measure and population, the use of oral chemotherapy in the management of cancer is increasingly rapidly and can account for significant toxicity as well. Also, I was unable to determine if the definition of chemotherapy included all antineoplastic therapy such as immunotherapy. Immunotherapy is a different ICD 10 code and may not be included in this measure but does account for significant toxicity. Please clarify.
- No significant concerns

- Have to consider whether to risk adjust further based on the actual diagnosis (e.g. chemotherapy for metastatic disease v. neoadjuvant therapy) and which line of treatment and expected toxicity based on drugs given
- OK
- Data elements seem to be defined and I do not have any concerns that the measure can be consistently implemented.

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

- Isn't this two different measures? (1) ED (2) Inpatient
- One question I have focuses on the separation of ED visits and inpatient admissions. I'm in agreement with one of the members of the Scientific Methods Panel who questioned whether the same patient may be counted who has an ED visit and then is directly admitted to the hospital. The direct admission from the ED would not be uncommon.
- No
- see 6.2a1
- The developers provided reliability testing. They also described the risk stratification methods. I have no concerns with the reliability of the measure.
- No significant concerns
- No
- No
- No concerns

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

- I think validity should be discussed and put up for a vote. This is a very complex measure, perhaps so complex as to produce metrics that are uninterpretable.
- Primarily demonstrated via face validity. The developers involved Technical Expert Panel, Expert Work Groups, as well as Public Comment feedback.
- No
- Moderate validity based on EWG feedback; extensive public commenting; agree with one EWG member's comment regarding validity of directly comparing non cancer vs. cancer hospitals performance with this measure (separate risk models).
- The statistical models used suggest that the measure is valid. Face validity was used. I have no concerns with the validity of the measure.
- No significant concerns
- No
- No
- No concerns

Validity- Threats to Validity (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?*)

- Because of the "at least one" adverse event and "one or more" visits/admissions counting strategy, a patient with one adverse effect and one ED visit is counted similarly to a patient with 10 adverse effects and 10 ED visits. It's unclear how the scores permit a meaningful understanding of differences between facilities when such situations are weighed equally.
- No concerns
- Concerns about comparability between practice settings.
- 2b4. based on construct, it appears that this measure provides face validity data to assess "The risk-standardized admissions rates and risk-standardized emergency department rates obtained from the chemotherapy measure as specified can be used to distinguish between better and worse quality facilities' ; no empiric validity exists. 2b5. -2b6- face validity only.
- The differences among the centers are likely to be meaningful. I do not have any concerns about threats to validity.
- No concerns
- Not as currently constructed

- No
- I do not have any concerns related to threats to validity.

Validity- Other Threats to Validity (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?)

- See earlier remark re: risk adjustment.
- Used statistical risk model with 21 risk factors for the inpatient admission outcome model and 16 risk factors for the ED visit outcome model.
- Acceptable
- 2b2. consider expanding exclusion criteria to other patient populations as stated above (MM and other lymphomas); also, immunotx diagnoses were excluded; consider coding for CART in exclusion criteria 2b3. social risk factor variables- tangential to conceptual measure design. Risk adjustment while extremely complex b/c involves 2 outcomes in 1 measure, appears appropriate. SDS was tested at length and it appears despite effect in cancer hospitals, no effect overall when assessed.
- The developers present extensive data regarding risk adjustment. This is appropriate given the differences in the populations, the hospitals being measured, and ED vs. inpatient components of the measure. I think that the exclusions are appropriate. One limitation of the measure is that it is based on hospital-based care rather than community-based care. There may be differences in this population when compared to patients who are managed in non-hospital settings. The measure is not intended to address this difference and therefore it is a valid measure for patients treated in the hospital setting.
- N/A
- Biggest threat is in case mix as it pertains to specialty hospitals and low volume hospitals. Likewise, may unfairly penalize hospitals that are hospitals of last resort; some concern about undercounting of ER visits based on hospital structure as some PPS-exempt hospitals do not have EDs per se but have 24 hr clinics.
- 2b3 - Some integrated oncology systems may have their own embedded urgent care/ER systems. That would not make them comparable to other systems which rely on other ERs.
- There are several patient groups excluded (leukemia, BMT, etc) but it was explained sufficiently. The only other concern (as stated by panel a panel member) is the 8 people who reviewed the measure were involved in creating the measure. That being said, I am not all that concerned.

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.

- All data elements are in defined fields in electronic claims and generated or collected by and used by healthcare personnel during the provision of care. The data are coded by someone other than person obtaining original information.
- There are no fees, licensing, or other requirements to use any aspect of the measure as specified.
- Measure development and testing showed that the measure cohort can be defined and outcomes reported using routinely collected Medicare claims data.

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?

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

Committee Pre-evaluation Comments:

Criteria 3: Feasibility

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?)

- Yes
- No concerns - all data elements are in defined fields. Outcomes can be reported using routinely collected Medicare claims data.
- All fields. No concerns
- All data elements are in defined fields in electronic claims ; based on dry run, The measure is primarily based on key fields in the claims data that are used for payment and, therefore, have a high level of completeness across Medicare claims and are considered reliable.
- Already implemented an appears feasible since the required elements are available in electronic form.
- Collected via routinely collected data.
- None
- I think this is OK
- I would rate feasibility as "High"; the data elements appear to be routinely generated and available in electronic format.

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

Accountability program details:

- The measure has been adopted for use in two CMS programs, the Hospital Outpatient Quality Reporting (OQR) Program and PPS-Exempt Cancer Hospital Quality Reporting (PCHQR) Program. OQR public reporting to start January 2020; PCHQR confidential reporting to start January 2019.

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

Feedback on the measure by those being measured or others

- The developer stated that the majority of feedback received from those being measured involved three topics:
 - Patients were included in the measure cohort who were receiving chemotherapy treatment for an autoimmune disease and not cancer;
 - Concern over patients being included in the outcome who were admitted for planned procedures (e.g., for stem cell transplantation); and,
 - Concern over patients being included in the cohort who had Leukemia in remission.
- To address the feedback received the developers took the following action:
 - Implemented a new case-level exclusion in which patients receiving chemotherapy to treat a qualifying autoimmune condition rather than cancer are excluded from the measure.
 - Implemented new logic into the measure that identifies and excludes outcomes identified as “always planned.”
 - Reviewed and revised the code set for exclusion of patients with leukemia to also exclude patients with leukemia in remission
 - Added a new risk-adjustment variable to the risk models for both outcomes that assesses whether a patient is receiving concurrent radiotherapy and chemotherapy.

Questions for the Committee:

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

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

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

4b. Usability 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.

Improvement results

- The measure has been adopted for public reporting in the Hospital OQR and PCHQR program beginning in CY 2020 and FY 2019, respectively. The developer expects there to be improvement in measure scores over time since publicly reported measure scores can reduce adverse patient outcomes associated with poorly managed outpatient care by capturing and making more visible to providers and patients all potentially preventable hospital visits following chemotherapy treatment in the hospital outpatient setting.

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

Potential harms

- The developer did not identify any unintended consequences during measure development or testing. However, during the NQF Measure Applications Partnership (MAP) review of this measure in December 2015, the MAP expressed concerns about a possible unintended consequence related to treatment decisions and underuse of appropriate care. The MAP’s concern was that the measure might indirectly discourage more aggressive treatment plans that would have had clinical benefits.

However, the purpose of the measure is to open lines of communication between the patient and provider on risks and preventative actions that can be taken for each type of treatment, and set the expectations for the patient so they can make more informed decisions on healthcare utilization as well. Furthermore, the measure is risk adjusted to help account for the variation in patient mix and aggressiveness of treatment. Lastly, the measure rate is not intended to be zero and CMS recognizes that not all admissions and ED visits are avoidable. Improving patient/provider communication and appropriately adjusting the model mitigates the risk of the unintended consequences. MAP advised that the measure undergo review and endorsement by NQF, with a special consideration from the Standing Committee of the exclusions and risk-adjustment methods. The measure has since been added to the two CMS programs noted above.

Additional Feedback:

- During the measure's first NQF endorsement review in 2016, members of the NQF Cancer Committee expressed concern over inclusion of patients in the measure receiving concurrent chemotherapy and radiotherapy, noting that these patients are at higher risk for an outcome due to increased exposure to toxins. In response to this feedback, the 2018 EWG recommended revising the risk-adjustment model to ensure that facilities treating a higher proportion of patients receiving concurrent chemotherapy and radiotherapy were not penalized for providing treatment to higher risk patients.

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

Committee Pre-evaluation Comments:

Criteria 4: Usability and Use

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?)

- Yes
- This measure is used in 2 CMS programs (OQR and PCHQR). Developers have addressed feedback from those being measured.
- IDK
- The measure is currently used in two CMS programs, the Hospital Outpatient Quality Reporting (OQR) Program and PPS-Exempt Cancer Hospital Quality Reporting (PCHQR) Program. Measure developers incorporated feedback from both public commenting and 2017 dry run to refine measure appropriately, incorporating additional exclusion criteria (such as patients on chemo for immune dx and additional risk stratification for patients receiving concurrent radiation).
- Already in use for public reporting. Useful for quality improvement and for developing information about the range and frequency of symptoms. Will improve quality and efficiency if actionable information is identified through these measures.
- Publicly reported via outpatient quality reporting program and the PPS exempt Cancer Hospital Quality Reporting Program
- Currently publicly reported; feedback seems appropriate.
- OK
- No concerns, there appears to be a credible plan for implementing.

Usability (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.)

- Perhaps, but there remains the concern that some of the significantly worse or significantly better hospitals may achieve that designation by virtue of referral/geographic constraints rather than initiatives/lapses related to care coordination
- No unintended consequences noted by developers. NQF Measure Application Partnership voiced concern about possible treatment decisions and underuse of appropriate care. Would this measure discourage more aggressive treatment plan that would have had clinical benefit? The measure is risk-adjusted to account for variation in patient mix and the ultimate score is not to be zero.
- Agree with concern regarding aversion to ED referral when appropriate
- 2015 MAP's concern was that the measure might indirectly discourage more aggressive treatment plans that would have had clinical benefits. However, the purpose of the measure is to open lines of communication between the patient and provider on risks and preventative actions that can be taken for each type of treatment, and set the expectations for the patient so they can make more informed decisions on healthcare utilization as well; need to monitor to ensure that patients have access to therapy and that this measure ensures AE management to prevent potentially avoidable complications for ER and admission and NOT prevent access to IV chemo for patients in need.
- I don't see any unintended harms from implementing this measure.
- Can be used to improve unnecessary admissions/ED admissions.
- May drive some to try to perform more care in the non-ED/non inpatient setting which may be inappropriate sometimes. Likewise may drive some hospitals to de-escalate chemotherapy sooner which may be less efficacious; this is balanced against better care coordination. Proper risk adjustment would mitigate harms.
- It is possible that there could be delays in care if a facility preferred the patient to wait to be seen at the oncology office rather than the ER.
- I think it was a good idea to use a risk-adjusted model to account for patients with a higher risk due to receiving both chemotherapy and radiation treatment.

Criterion 5: [Related and Competing Measures](#)

Related measures

- 0383 : Oncology: Plan of Care for Pain – Medical Oncology and Radiation Oncology (paired with 0384)
- 0384 : Oncology: Medical and Radiation - Pain Intensity Quantified
- 1628 : Patients with Advanced Cancer Screened for Pain at Outpatient Visits

Harmonization

- All four related measures (NQF 0383a, NQF 0384a, NQF 1628, and Cancer – fatigue/anemia) focus on cancer patients receiving outpatient chemotherapy; however, there are some key differences in measure scope and measure type.
 - Measure scope: Each of the four related measures (NQF 0383a, NQF 0384a, NQF 1628, and Cancer – fatigue/anemia) narrowly focuses on pain management and/or fatigue/anemia. Measure 3490 does not target a specific symptom, but rather assesses the overall management of 10 important symptoms and complications that was identified as being more frequently cited in literature as reasons for ED visits and inpatient admissions following outpatient chemotherapy.
 - Measure type: The four related measures (NQF 0383a, NQF 0384a, NQF 1628, and Cancer – fatigue/anemia) are all process measures encouraging the use of screening and care plans to

improve care. Measure 3490 is an outcome measure not encouraging or measuring specific processes to detect and treat these conditions, but rather assessing the outcomes of the care being provided. The four process measures, which are not risk-adjusted, support the intent of the measure by reinforcing that those providing outpatient care should screen for and manage symptoms such as pain.

Committee Pre-evaluation Comments: Criterion 5:

Related and Competing Measures

Relate 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?*)

- No
- No related or competing measures.
- IDK
- There are 3 related process measures for cancer pain. However, these measures do not compete as they are process measures, not outcomes measures.
- No competing measures appear to have been identified.
- Yes, but there is minimal overlap.
- Not that I know of.
- There are related measures (as noted) but I believe this measure takes different approach by assessing the outcome of care and not a narrow focus on specific symptom.

Public and Member Comments

NQF received two public comments from two organizations as of January 30, 2019

1. The Federation of American Hospitals (FAH) appreciates the opportunity to comment on this measure prior to the Standing Committee's evaluation. While FAH agrees with the potential for this measure to support quality improvement efforts, we have several concerns regarding the measure and its intended use for accountability purposes.

The FAH questions why an assessment of similarity (some kind of analysis of variance or inter-class reliability) between the two groups (PPS-exempt cancer hospitals and non-cancer hospitals) was not made. For example, the risk-standardized admissions rate was 0.3116 lower for non-cancer hospitals and the risk-standardized emergency department visit rate was 0.3932 less than the PPS-exempt cancer hospitals at the 25th percentile. It is not clear whether these differences indicate whether there are group level effects that impact the measure. FAH understands it is important to account for the effects of clusters and whether there are differences in the repeatability of the measure. A difference which may suggest whether additional review is needed to determine if further refinements should be made to the measure to enable similar findings across the two distinct groups.

In addition, the FAH was disappointed to see that the risk adjustment model continues to include the identification and testing of social risk factors as supplementary. Given that this is a new measure, it provided an opportunity for the measure developer to include these factors within the testing of the model rather than the previous approach of "adding on"; factors after the model is developed. This type of approach would assist hospitals and others in understanding how their inclusion could impact the model and provide additional information for groups examining this issue such as the NQF and Office of the Assistant Secretary for Planning and Evaluation.

As a result, the FAH does not believe that this measure is not appropriate for use for accountability purposes and lacks sufficient information on the social risk factors in the risk adjustment approach.

FAH does not support endorsement of this measure at this time. The FAH thanks you for the opportunity to comment.

2. The Alliance of Dedicated Cancer Centers (ADCC) represents the premier cancer centers in the nation, all of whom participate in the PPS-Exempt Cancer Hospital Quality Reporting (PCHQR) Program. Unlike other hospitals that care for patients suffering from any condition, the Dedicated Cancer Centers treat cancer patients exclusively. Much of the progress in understanding cancer's biology and successful treatment methods is directly attributable to the work of ADCC members. Our institutions are at the forefront of innovative treatment options in precision medicine, immunotherapies, and other state of the art diagnostic and patient care technologies. The Dedicated Cancer Centers are committed to delivering the highest standard of cancer care and share the Centers for Medicare & Medicaid Services' (CMS) focus on cancer care delivery that is safe, effective, high-quality, and patient-centered. We are committed to achieving the best outcomes for our patients through novel therapies and excellent care delivery. Our members serve as regional, national, and international resources in developing the most effective and efficient ways to treat cancer patients.

We had provided extensive comment to the measure developer after the official dry run of this measure was conducted in Fall 2017. Furthermore, several members of the Expert Work Group (EWG) formed to respond to the findings of this dry run were from ADCC member institutions. This resulted in the updated measure specifications that are currently being reviewed in this review cycle by the NQF. Unfortunately, as the measure steward has disclosed in the materials submitted, these updated measure specifications were not used to produce the most recent measure results that CMS shared with the PCHQR program participants. Thus, we are limited in our ability to comment on the proposed updated specifications. According to the measure developer, it is anticipated that the ADCC members will receive data using the revised specifications in the Summer of 2019.

With that caveat in mind, we offer the following input pertaining to this measure based upon a review of the new technical specifications, and a review of the data received in the Fall of 2019 from CMS.

Denominator Validation of Patients Who Received Chemotherapy:

Past analysis revealed that for 5-7% of patients included in the denominator, the date of administration of the outpatient chemotherapy could not be confirmed. This issue was identified in the latest round of Facility-Specific Reports (FSRs) as well. If the updated specifications include removal of the ICD-10 code Z51.11, that may reduce the number of cases in which patients who did not actually receive chemotherapy are included in the denominator, as will the exclusion of cases with AHQR CCS codes for bone marrow transplant and chemotherapy.

We continue to identify cases in which patients received biologic response modifiers and hormonal or supportive care agents but no chemotherapy, and strongly recommend chemotherapy specifications be further updated to exclude medications such as BCG, degarelix, goserelin, mesna, leuprolide and histrelin.

Numerator Validation

Patients Whose Admission Was Planned, Not the Result of an Adverse Event Associated with Chemotherapy: We identified a number of cases in which patients who had planned admissions (such as surgery after neo-adjuvant therapy, stem cell transplantation, and CAR-T cell therapy) were included in the numerator. We anticipate the inclusion of the AHRQ CCS exclusion codes will reduce the number of cases in which planned admissions are included in the numerator. In addition, or alternatively, we also recommend the specifications take into account whether an admission is coded as “elective” (classification “3” on UB-04) to further reduce the number of planned admissions included erroneously in the numerator.

Patients Admitted for Reasons Other Than Adverse Events from Chemotherapy: Several of the Dedicated Cancer Centers conducted chart and/or clinical reviews to ascertain whether potentially preventable diagnoses were reasonably attributable to the chemotherapy. These Centers identified multiple instances in which the potentially preventable diagnosis, particularly pain, was attributed to factors other than chemotherapy, such as disease progression including pericardial effusion, bladder rupture, or cord compression. One solution to better capture symptoms related to prior chemotherapy administration is to incorporate Present on Admission (POA) codes. We proposed this update in our comments to the proposed Final Rule and we understand the measure developer was receptive to this recommendation. Nevertheless, we were disappointed to learn that these updated specifications do not consider whether the qualifying symptoms are present on admission.

As noted in the CMS Measure Dry Run Facility Specific Report (FSR) User Guide for Admissions and Emergency Department (ED) Visits for Patients Receiving Outpatient Chemotherapy document released in August 2017, the goal of the measures is to stimulate efforts “to improve the quality of care delivered to patients undergoing chemotherapy in the hospital outpatient department (HOPD).” Therefore, attribution of these adverse events to the chemotherapy is a critical component of the measure. This point is reinforced under question #22 in the CMS Measure Dry Run Frequently Asked Questions (FAQs) for Admissions and Emergency Department (ED) Visits for Patients Receiving Outpatient Chemotherapy document released in August 2017, which states “...relating the time frame to a specific chemotherapy administration supports the idea that the admission stems from the management of side effects of treatment and ongoing care, rather than progression of the disease or other unrelated events.” Furthermore, question #33 notes that “facilities that provide outpatient chemotherapy should proactively implement appropriate care to minimize the need for acute hospital care for these adverse events. Guidelines from the American Society of Clinical Oncology, National Comprehensive Cancer Network...and other professional societies recommend evidence-based interventions to prevent and treat common side effects and complications of chemotherapy.” In light of the intent of these measures, we strongly recommend that the measures be further refined to capture accurately side effects and complications of the outpatient chemotherapy accurately, and distinguish from the effects of the cancer in general.

Another potential to consider to reduce the excessive noise of adverse events that are not under the control of the clinical team caring for the patient receiving outpatient chemotherapy would be to exclude those patients with metastatic disease. Patients with metastatic disease have more advanced and/or aggressive disease and may in fact present with adverse events not associated with the outpatient chemotherapy.

Existing code sets are limited in their ability to identify adverse events resulting from chemotherapy. Although the add-on adverse effect chemotherapy code T451X5A can be used to indicate instances

in which a condition is chemotherapy-related, it is not possible to identify which specific condition occurred as a result of the chemotherapy when multiple conditions are listed. Given the current lack of specificity of codes available, the ADCC submitted a suggestion for a more durable and sustainable solution in their July 2017 letter to the ICD-10 Coordination and Maintenance Committee, recommending that additional ICD-10-CM diagnosis codes be created that specifically identify diagnoses related to chemotherapy to improve CMS quality measure reporting.

Limitations to Risk Adjustment:

Although numerous factors are taken into consideration in the risk adjustment model for this measure, basic cancer-specific factors, such as the cytotoxicity of chemotherapy regimens and disease stage of patients, are not included. The absence of these factors in the risk adjustment is concerning within the context of the PCHQR Program; this absence poses a particularly serious limitation in the risk adjustment method used in the Hospital Outpatient Quality Reporting Program, given the variation in patient populations across the hospitals included in that program. In the absence of adequate risk adjustment, public reporting of these measures in both Programs could lead to inaccurate benchmark comparisons. An example of this are myeloma and lymphoma patients, who oftentimes have aggressive conditioning regimens.

We thank you for this opportunity to comment. While we are supportive of the intent of this measure to reduce potentially preventable harm and associated costs, at this time we cannot support endorsement. We strongly recommend further testing of the improved measure specifications in the cancer-hospital setting. We also strongly encourage adding the requirement that the adverse events be POA to be included in the numerator and the exclusions for planned readmissions be expanded to include all elective admissions. We also ask that consideration be given to the exclusion of patients with metastatic disease. As currently tested, there are too many “false positives” attributed to the numerator, making productive performance improvement efforts difficult. We are happy to continue to offer supportive guidance to CMS and their partners in the further refinement of this measure.

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

[Chemotherapy_Evidence_11.07.18.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): N/A

Measure Title: [Admissions and Emergency Department \(ED\) Visits for Patients Receiving Outpatient Chemotherapy](#)

IF the measure is a component in a composite performance measure, provide the title of the Composite Measure here: N/A

Date of Submission: [11/15/2018](#)

Instructions

- Complete 1a.1 and 1a.2 for all measures. If instrument-based measure, complete 1a.3.
- Complete **EITHER 1a.2, 1a.3 or 1a.4** as applicable for the type of measure and evidence.
- For composite performance measures:
 - A separate evidence form is required for each component measure unless several components were studied together.
 - If a component measure is submitted as an individual performance measure, attach the evidence form to the individual measure submission.
- All information needed to demonstrate meeting the evidence subcriterion (1a) must be in this form. An appendix of *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
- If you are unable to check a box, please highlight or shade the box for your response.
- Contact NQF staff regarding questions. Check for resources at [Submitting Standards webpage](#).

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

1a. Evidence to Support the Measure Focus

The measure focus is evidence-based, demonstrated as follows:

- **Outcome:** [3](#) Empirical data demonstrate a relationship between the outcome and at least one healthcare structure, process, intervention, or service. If not available, wide variation in performance can be used as

evidence, assuming the data are from a robust number of providers and results are not subject to systematic bias.

- **Intermediate clinical outcome:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence ⁴ that the measured intermediate clinical outcome leads to a desired health outcome.
- **Process:** ⁵ a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence ⁴ that the measured process leads to a desired health outcome.
- **Structure:** a systematic assessment and grading of the quantity, quality, and consistency of the body of evidence ⁴ that the measured structure leads to a desired health outcome.
- **Efficiency:** ⁶ evidence not required for the resource use component.
- For measures derived from patient reports, evidence should demonstrate that the target population values the measured outcome, process, or structure and finds it meaningful.
- **Process measures incorporating Appropriate Use Criteria:** See NQF's guidance for evidence for measures, in general; guidance for measures specifically based on clinical practice guidelines apply as well.

Notes

3. Generally, rare event outcomes do not provide adequate information for improvement or discrimination; however, serious reportable events that are compared to zero are appropriate outcomes for public reporting and quality improvement.

4. The preferred systems for grading the evidence are the Grading of Recommendations, Assessment, Development and Evaluation ([GRADE](#)) guidelines and/or modified GRADE.

5. Clinical care processes typically include multiple steps: assess → identify problem/potential problem → choose/plan intervention (with patient input) → provide intervention → evaluate impact on health status. If the measure focus is one step in such a multistep process, the step with the strongest evidence for the link to the desired outcome should be selected as the focus of measurement. Note: A measure focused only on collecting PROM data is not a PRO-PM.

6. Measures of efficiency combine the concepts of resource use and quality (see NQF's [Measurement Framework: Evaluating Efficiency Across Episodes of Care](#); [AQA Principles of Efficiency Measures](#)).

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

Outcome

☒ Outcome: [One or more inpatient admissions or emergency department \(ED\) visits for one of the following diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of chemotherapy treatment among cancer patients receiving treatment in a hospital outpatient setting.](#)

☐ 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: [Click here to name what is being measured](#)

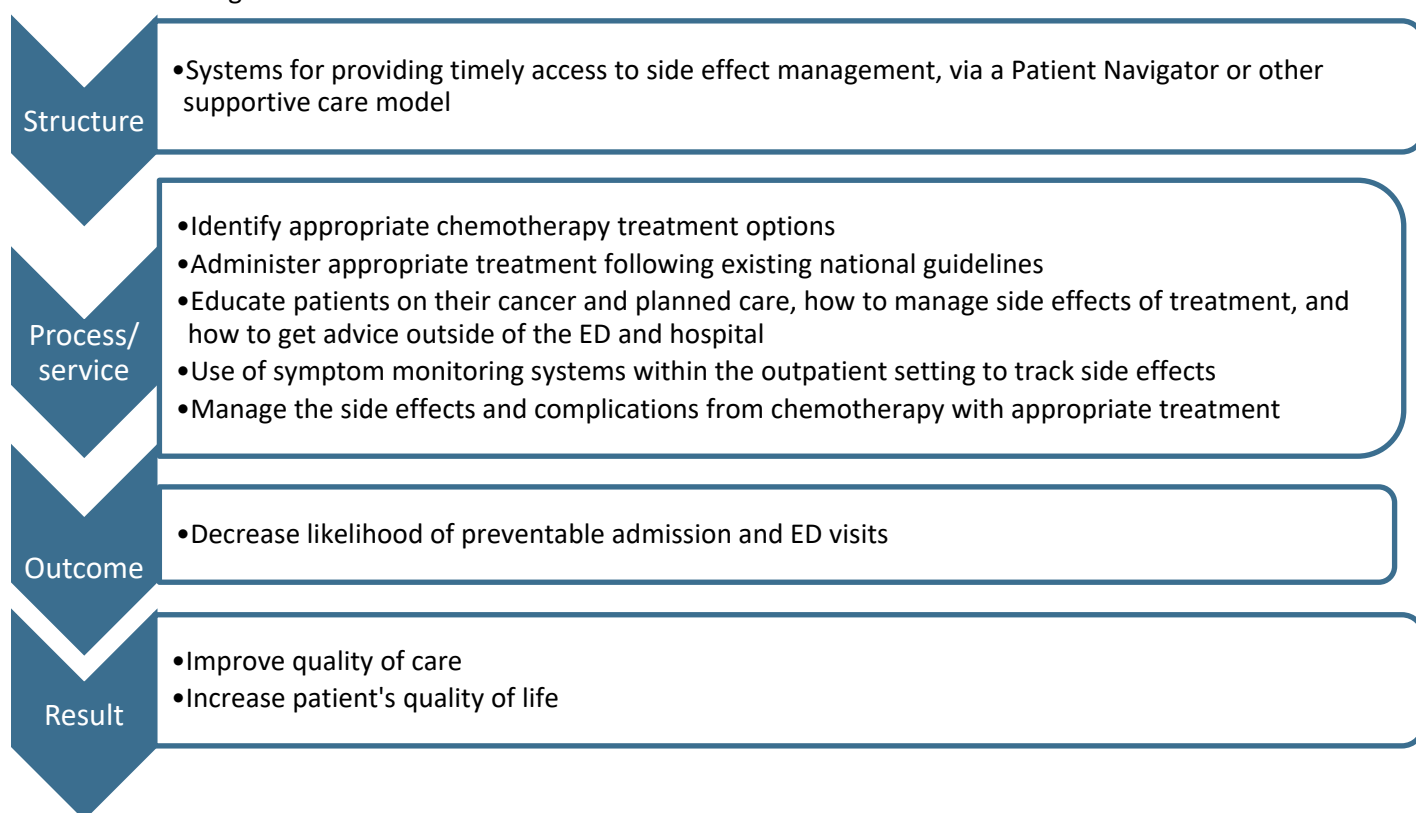
☐ Appropriate use measure: [Click here to name what is being measured](#)

☐ 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.



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.)

Not applicable. This is a claims-based measure.

****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.

To demonstrate the relationship between the measure outcome and current healthcare processes/services as they relate to chemotherapy care, we provide evidence that: (1) better management of these symptoms and enhanced coordination of care reduces outcome rates among patients receiving chemotherapy; (2) chemotherapy patients frequently seek emergency department (ED) care or experience inpatient hospital admissions due to the ten diagnoses/symptoms comprising the measure outcome; and (3) all ten symptoms can be managed in the outpatient setting using national clinical guidelines and established best care practices.

Improved Symptom Management and Coordination of Care Reduces Hospital Visits

Chemotherapy treatment can have severe, predictable side effects, and hospital admissions and ED visits among patients receiving treatment in a hospital outpatient setting are often caused by manageable side effects and complications. Improved management of these symptoms and coordination of care in the outpatient setting can decrease hospital visits among patients receiving chemotherapy.

Divergence from established guidelines for use of antiemetic medications to manage chemotherapy-related nausea can result in adverse outcomes. A 2011 study identified great variability in the use of antiemetic medications to manage chemotherapy-related nausea. Most medications prescribed in this study did not follow American Society of Clinical Oncology Guidelines, and researchers suggested that the low level of agreement between actual clinical practice and evidence-based consensus guidelines may be contributing

significantly to the incidence of chemotherapy-related nausea and vomiting [1]. In another study, nonadherence to established, evidence-based guidelines for antiemetic medications was associated with increased occurrence of chemotherapy-induced nausea, and patients who received proper medications were much less likely to experience vomiting (6.6% v 21.9%; $P < .001$), emergency department visits (2.6% v 5.8%; $P = .006$), and hospitalization for emesis (0.9% v 4.9%; $P < .001$) [2].

Enhanced care coordination can also decrease hospital visits and ED visits among cancer patients receiving chemotherapy. According to a 2017 study, implementation of a hospital-based, dedicated supportive care service to monitor and assist outpatient chemotherapy patients with treatment-related symptoms showed decreases of 18.5% in unplanned hospital admissions (from 17.3% to 14.1%) and 7.6% in ED visits (from 66.0% to 61.0%), relative to the pre-implementation period [3]. The authors note that these decreases occurred even though outpatient chemotherapy volume increased by approximately 6.5% (from 1,275 to 1,358) during the study period. In a second study, routine symptom screening of breast cancer patients undergoing adjuvant outpatient chemotherapy was associated with a 43% decrease in ED visits relative to those who were not screened. For each additional prior symptom screening assessment, there was a further 17% decrease in the rate of ED visits [4]. Chemotherapy patients with access to enhanced electronic care monitoring systems also experience fewer hospital visits relative to those without access. According to a 2016 random control trial, patients who were able to report symptoms using tablet computers, which triggered an email alert to the clinical nurse and were summarized for review during clinic visits with the treating oncologist, had 17% fewer ED visits (from 41% to 34%; $P = .02$) and 8% fewer hospitalizations (from 49% to 45%; $P = .08$) relative to patients receiving usual symptom monitoring [5].

The introduction of additional hospital-based coordination and monitoring systems have also been associated with declines in adverse events among chemotherapy patients. At the University of Alabama at Birmingham Health System Cancer Community Network, patient navigators were assigned to high-risk cancer patients to improve their access to care, enhance coordination, and overcome barriers to obtaining timely, high-quality care [6]. The study authors found that relative to matched, non-navigated patients, those with a navigator had fewer emergency department visits (6.0% decrease), hospitalizations (7.9% decrease), and intensive care unit admissions (10.6% decrease). According to another study, utilization of an oncology management program that prioritizes survival, minimizing toxicity, and avoiding unnecessary healthcare, along with a telephonic nursing intervention – wherein oncology-certified nurses contacted, assessed, and educated patients in between treatments – resulted in decreases of 28.6% in ED visits (from 14% to 10%) and 25.0% in inpatient admissions (from 24% to 18%), relative to the control group [7].

Facility-wide, alternative delivery models focused on coordinating care can also improve the management of chemotherapy related adverse events. According to a 2013 study examining breast cancer patients, patients who were treated at a facility using a patient-centered medical home delivery model were significantly less likely to experience an inpatient admission with chemotherapy-related adverse events compared to patients who were provided with usual care [8]. According to a second study, patients in an oncology medical home demonstration project had 68% fewer ED visits (0.07 relative to 0.22) and 47% fewer inpatient admissions (0.18 relative to 0.34) per patient, relative to historical control data. The study's authors concluded that in addition to reducing hospital visits and reducing costs, the model encouraged adherence to national guidelines, advance care planning, and standardized symptom management [9].

Reasons for Admissions and ED Visits among Cancer and Chemotherapy Patients

Admissions and ED visits for the ten diagnoses captured in the measure—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—are among the most common reasons that cancer patients receiving chemotherapy visit the hospital [10] [11-19].

The frequency of and reasons for hospital admissions and ED visits among cancer patients overall, and among specific subpopulations of cancer patients receiving chemotherapy, are well documented in the literature. An analysis by Rivera et al of Nationwide Emergency Department Sample (NEDS) data from 2006 – 2012 determined that 4.2% of all ED visits ($n = 29.5$ million) were made by patients with cancer, and the most

common primary reasons for these visits were pneumonia (4.5%), nonspecific chest pain (3.7%), and urinary tract infection (3.2%) [16]. Among visits where maintenance chemotherapy or radiotherapy was reported, the primary reasons for the visit were deficiency and other anemia (5.7%), fluid and electrolyte disorders (4.7%), nausea and vomiting (4.3%), diseases of white blood cells (3.3%), and fever of unknown origin (3.1%). A smaller assessment of ED visits among cancer patients living in North Carolina similarly found that the top 3 most frequent complaints were: (1) pain (chest pain, abdominal pain, back pain, extremity pain, other), (2) respiratory (respiratory distress/shortness of breath, cough, hemoptysis, fever/pneumonia, chronic obstructive pulmonary disease, or other), and (3) gastro-intestinal (nausea/vomiting, diarrhea, constipation, bowel obstruction, other) [13].

Additional studies focusing on specific populations of cancer patients receiving chemotherapy show similar results. Among breast cancer patients receiving chemotherapy, one study reported the most common reasons for hospital visits were fever or infection (8.4%), neutropenia or thrombocytopenia (5.5%), dehydration or electrolyte disorders (2.5%), nausea, emesis, or diarrhea (2.4%), and anemia (2.2%) [12], while a second study found that fever (23.3%), pain (12.8%), and febrile neutropenia (9%) were the most frequent reasons for hospital visits [15]. For colorectal cancer patients receiving chemotherapy, the majority of unplanned visits occurred within 30 days of treatment and the most frequent complaints were pain, fatigue, and anorexia [10]. In another study of 233 cancer patients receiving chemotherapy who visited the hospital, the authors reported the most frequent symptoms were: nausea and/or vomiting (45.2%), pain (27%), fever and/or febrile neutropenia (23.4%), shortness of breath (19.3%), dehydration (12.1%), anemia (8.8%), fatigue (8.8%), diarrhea (8.8%), and anxiety and/or depression (5.5%) [14]. Furthermore, 70% of all hospital visits occurred within 4 weeks of receiving chemotherapy, and the majority (87.6%) resulted in hospital admission.

Guidelines to Support Outpatient Management of Measure Outcome Conditions/Symptoms

Treatment plans and guidelines exist to support the outpatient management of the ten conditions captured in the outcome. Guidelines from the American Society of Clinical Oncology, National Comprehensive Cancer Network, Oncology Nursing Society, Infectious Diseases Society of America, and other professional societies recommend evidence-based interventions to improve the quality of disease and symptom management [20] [21]. Proper management of symptoms associated with outpatient chemotherapy reduces the risk of admissions and ED visits for side effects and complications such as nausea and vomiting, anemia, and neutropenic fever [22] [23] [24]. Below we provide more detail on clinically proven treatment plans used to prevent and manage the side effects and symptoms of cancer and outpatient chemotherapy treatment that decrease the risk of admissions and ED visits.

Anemia: There are many therapeutic agents (e.g., epotein beta) available to treat anemia as well as clinical guidelines on how to prevent and manage anemia in patients receiving chemotherapy treatment [25] [26] [27].

Dehydration: Dehydration can be prevented by educating patients on the importance of fluid intake and monitoring patients that have reduced oral intake or appetite loss. Healthcare professionals should also closely monitor patients at risk for chemotherapy-induced diarrhea and vomiting for signs of dehydration [28].

Diarrhea: Providers can often treat chemotherapy-induced diarrhea on an outpatient basis, and effective treatment of diarrhea can prevent dehydration [28]. Existing evidence enables management of diarrhea, and evidence about prevention continues to evolve as research focuses on identifying predictive factors of chemotherapy-induced diarrhea [29].

Nausea/emesis: Chemotherapy-induced nausea and emesis can be prevented and effectively managed in the outpatient setting [30]. Studies and reviews have shown the effectiveness of specific drugs (e.g. serotonin receptor antagonists, dexamethasone, and aprepitant) for prevention and management of nausea and emesis resulting from particular chemotherapy regimens and their effects on quality of life [31] [30] [32] [33] [34] [35] [36].

Neutropenic fever: A systematic review and meta-analysis of randomized controlled trials concluded that prophylactic granulocyte colony-stimulating factors significantly reduce neutropenic fever [37]. Additionally, a

2017 update to the standard treatment guidelines published by the American Society of Clinical Oncology and Infectious Diseases Society of America Clinical Practice recommends the use of validated tools such as the Multinational Association of Support Care in Cancer Risk Index when determining candidacy for outpatient management of neutropenic fever [20].

Pain: A number of pharmacological treatments for pain exist, including opioids. However, many patients receive inadequate analgesia [38] [39]. Optimal pain control can be achieved through combining pharmacological and non-pharmacological approaches, in addition to assessing and reassessing patients' pain [40].

Pneumonia/Sepsis: The relationship between neutrophil count and the risk of infection is well established and studies have shown that risk factors can be identified and appropriate prophylactic measures, such as use of colony-stimulating factor, implemented to prevent neutropenia and associated complications [41]. Because of this relationship and the need for lab results to confirm neutropenia, neutropenia is often captured on the claim as the related infection, such as pneumonia and sepsis. The measure includes pneumonia and sepsis as outcomes to capture this population [37] [41].

Conclusion

We have shown that specific healthcare structures, processes, and services have a demonstrated relationship with the measure outcome. There is clear evidence that better management of the ten diagnoses/symptoms captured by this measure and enhanced coordination of care reduces the rate of inpatient admissions and ED visits among patients receiving chemotherapy. In addition, there is strong evidence that these ten symptoms are primary factors in chemotherapy patients seeking emergency department care or experiencing inpatient hospital admissions, indicating that the measure focus is appropriate and important for cancer patients receiving outpatient chemotherapy. Finally, established national clinical guidelines and best practices on appropriate care underlying effective symptom management in the outpatient setting suggests that there are specific evidence-based interventions that will reduce hospital visits. This evidence supports the relevance and need for this measure.

Citations

1. Molassiotis, A., Brearley, S.G. & Stamatakis, Z. "Use of antiemetics in the management of chemotherapy-related nausea and vomiting in current UK practice." *Support Care Cancer* (2011) 19: 949.
2. Patil, Vijay M., Vanita Noronha, Amit Joshi, Anant Ramaswamy, Sudeep Gupta, Arvind Sahu, Vipul Doshi, Tarachand Gupta, Sushmita Rath, Shripad Banavali, and Kumar Prabhash. "Adherence to and Implementation of ASCO Antiemetic Guidelines in Routine Practice in a Tertiary Cancer Center in India." *Journal of Oncology Practice*, vol. 13, no. 6, 2017, pp. e574-e581.
3. Antonuzzo, A., Vasile, E., Sbrana, A. et al. "Impact of a supportive care service for cancer outpatients: management and reduction of hospitalizations. Preliminary results of an integrated model of care." *Support Care Cancer* (2017) 25: 209.
4. Barbera, L., Sutradhar, R., Howell, D. et al. "Does routine symptom screening with Edmonton Symptom Assessment Scale (ESAS) decrease ED visits in breast cancer patients undergoing adjuvant chemotherapy?" *Support Care Cancer* (2015) 23: 3025
5. Basch, E., Deal, A., Kris, M.G., et al. "Symptom Monitoring With Patient-Reported Outcomes During Routine Cancer Treatment: A Randomized Controlled Trial." *Journal of Clinical Oncology* 34, no. 6 (February 2016) 557-565.
6. Rocque, G.B., Pisu, M., Jackson, B.E., et al. "Resource Use and Medicare Costs During Lay Navigation for Geriatric Patients With Cancer." *JAMA Oncol.* 2017; 3(6):817-825.
7. Hoverman, J.R., Klein, I., Harrison, D.W., et al. "Opening the Black Box: The Impact of an Oncology Management Program Consisting of Level I Pathways and an Outbound Nurse Call System." *Journal of Oncology Practice* 10, no. 1 (January 2014) 63-67.

8. Goyal, Ravi K., Stephanie B. Wheeler, Racquel E. Kohler, Kristen H. Lich, Ching-Ching Lin, Katherine Reeder-Hayes, Anne-Marie Meyer, and Deborah K. Mayer. "Health Care Utilization from Chemotherapy-Related Adverse Events among Low-Income Breast Cancer Patients: Effect of Enrollment in a Medical Home Program." *North Carolina Medical Journal*, vol. 75, no. 4, 2014, pp. 231-8.
9. Kuntz, G., Tozer, J.M., Snegosky, J., Fox, J., and Neumann, K. "Michigan Oncology Medical Home Demonstration Project: First-Year Results." *Journal of Oncology Practice* 10, no. 5 (September 1 2014) 294-297.
10. Foltran, L., Aprile, G., Pisa, F.E. et al. "Risk of unplanned visits for colorectal cancer outpatients receiving chemotherapy: a case-crossover study." *Supportive Care in Cancer*, September 2014, Volume 22, Issue 9, PP 2527-2533.
11. Burke, T.A., Wisniewski, T., and Ernst, F.R. "Resource utilization and costs associated with chemotherapy-induced nausea and vomiting (CINV) following highly or moderately emetogenic chemotherapy administered in the US outpatient hospital setting." *Supportive Care in Cancer*, vol. 19, 2011, pp. 131-141.
12. Hassett, M.J., J. O'Malley, J.R. Pakes, J.P. Newhouse, and C.C. Earle. "Frequency and Cost of Chemotherapy-Related Serious Adverse Effects in a Population Sample of Women with Breast Cancer." *Journal of the National Cancer Institute*, vol. 98, no. 16, 2006, pp. 1108–1117.
13. Mayer, D.K., D. Travers, A. Wyss, A. Leak, A. Waller. "Why Do Patients with Cancer Visit Emergency Departments? Results of a 2008 Population Study in North Carolina." *Journal of Clinical Oncology*, vol. 26, no. 19, 2011, pp. 2683–2688.
14. McKenzie, H., L. Hayes, K. White, K. Cox, J. Fethney, M. Boughton, and J. Dunn. "Chemotherapy Outpatients' Unplanned Presentations to Hospital: A Retrospective Study." *Support Care Cancer*, vol. 19, 2011, pp. 963–969.
15. Pittman, N.M., Hopman, W.M., and Mates, M. "Emergency Room Visits and Hospital Admission Rates After Curative Chemotherapy for Breast Cancer." *Journal of Oncology Practice* 11, no. 2 (March 1 2015) 120-125.
16. Rivera, D.R., Gallicchio, L., and Brown, J. "Trends in Adult Cancer-Related Emergency Department Utilization: an Analysis of Data from the Nationwide Emergency Department Sample." *JAMA Oncol.* 2017; 3(10):e172450.
17. Solares, S.S., Vidal, M., Ferrer, A., Tobena, M., Millastre, E., Ruiz Echarri, M., Pajares, I. Lopez, D., Tres, A., and Lambea, J. "Adverse Events due to Chemotherapy in Emergency Room." *Journal of Clinical Oncology* 27, no. 15S – published online before print.
18. Daly, B., Nicholas, K., Gorenshteyn, D., et al. "Misery Loves Company: Presenting Symptom Clusters to Urgent Care by Patients Receiving Antineoplastic Therapy." *Journal of Oncology Practice* 14, no. 8 (August 1 2018) e484-e495.
19. Daly, B. and Abougergi, M.S. "National trends in admissions for potentially preventable conditions among patients with metastatic solid tumors, 2004-2014." *Journal of Clinical Oncology* 36, no. 30_suppl (October 20 2018) 1-1.
20. Taplitz, R.A., Kennedy, E.B., Bow, E.J., et al. "Outpatient Management of Fever and Neutropenia in Adults Treated for Malignancy: American Society of Clinical Oncology and Infectious Diseases Society of America Clinical Practice Guideline Update." *Journal of Clinical Oncology* 36, no. 14 (May 2018) 1443-1453.
21. Brahmer, J.R., Lacchetti, C., Schneider, B.J., et al. "Management of Immune-Related Adverse Events in Patients Treated With Immune Checkpoint Inhibitor Therapy: American Society of Clinical Oncology Clinical Practice Guideline." *Journal of Clinical Oncology* 36, no. 17 (June 2018) 1714-1768.

22. Abernethy, A.P., Barbour, S.Y., Uronis, H. et al. "Quality management of potential chemotherapy-induced neutropenic complications: evaluation of practice in an academic medical center." *Support Care Cancer* (2009) 17: 735.
23. Goodman, L.M., Estfan, B., Montero, A., et al. "Improving the Management of Patients With Low-Risk Neutropenic Fever at the Cleveland Clinic Taussig Cancer Institute." *Journal of Oncology Practice* 13, no. 3 (March 1 2017) e259-e265.
24. Ying, F.L.M., Ping, M.C.Y., Tong, M. et al. "A cohort study on protocol-based nurse-led out-patient management of post-chemotherapy low-risk febrile neutropenia." *Support Care Cancer* (2018) 26: 3039
25. Crowley, K., and K. Augustin. "Chemotherapy-Induced Anemia." *US Pharmacist*, vol. 28, no. 04, 2003.
26. Groopman, J.E., and L.M. Itri. "Chemotherapy-Induced Anemia in Adults: Incidence and Treatment." *Journal of the National Cancer Institute*, vol. 91, 2000, pp. 1616–1634.
27. Wang, J., Wu, Y., Yu, D., et al. "Multicenter open label study of epoetin beta prevention and treatment for anaemia in non-small cell lung cancer patients treated with chemotherapy prone to induce anaemia." *Journal of Clinical Oncology* 26, no. 15_suppl (May 20 2008) 20743-20743.
28. Richardson, G., and R. Dobish. "Chemotherapy Induced Diarrhea." *Journal of Oncology Pharmacy Practice*, vol. 13, no.4, 2007, pp. 181–98.
29. Stein, A., W. Voigt, and K. Jordan. "Chemotherapy-Induced Diarrhea: Pathophysiology, Frequency, and Guideline-Based Management." *Therapeutic Advances Medical Oncology*, vol. 2, 2010, pp. 51–63.
30. Billio, A., E. Morello, and M.J. Clarke. "Serotonin Receptor Antagonists for Highly Emetogenic Chemotherapy in Adults." *Cochrane Database of Systematic Reviews*. Available at <http://onlinelibrary.wiley.com/doi/10.1002/14651858.CD006272.pub2/abstract>. Accessed on September 24, 2012.
31. Craver, Chris, Julie Gayle, Sanjeev Balu, and Deborah Buchner. "Palonosetron Versus Other 5-HT₃ Receptor Antagonists for Prevention of Chemotherapy-Induced Nausea and Vomiting in Patients with Hematologic Malignancies Treated with Emetogenic Chemotherapy in a Hospital Outpatient Setting in the United States." *Journal of Medical Economics*, vol. 14, no. 3, 2011, pp. 341-9.
32. Lohr, L. "Chemotherapy-Induced Nausea and Vomiting." *Cancer Journal*, vol. 14, no. 2, 2008, pp. 85–93.
33. Navari, R.M. "Prevention of Emesis from Multiple-Day and High-Dose Chemotherapy Regimens." *Journal of the National Comprehensive Cancer Network*, vol. 5, no. 1, January 2007, pp. 51–59.
34. Osoba, D., B. Zee, D. Warr, J. Latreille, L. Kaizer, and J. Pater. "Effect of Postchemotherapy Nausea and Vomiting on Health-Related Quality of Life." *Support Care Cancer*, vol. 5, 1997, pp. 307–313.
35. Rapoport, B.L., Jordan, K., Boice, J.A. et al. "Aprepitant for the prevention of chemotherapy-induced nausea and vomiting associated with a broad range of moderately emetogenic chemotherapies and tumor types: a randomized, double-blind study." *Support Care Cancer* (2010) 18: 423.
36. Hesketh, P.J., Kris, M.G., Basch, E., et al. "Antiemetics: American Society of Clinical Oncology Clinical Practice Guideline Update." *Journal of Clinical Oncology* 35, no. 28 (October 2017) 3240-3261.
37. Kuderer, N.M., D.C. Dale, J. Crawford, and G.H. Lyman. "Impact of Primary Prophylaxis with Granulocyte Colony-Stimulating Factor on Febrile Neutropenia and Mortality in Adult Cancer Patients Receiving Chemotherapy: A Systematic Review." *Journal of Clinical Oncology*, vol. 25, 2007, pp. 3158–3167.
38. Fisch, M.J., J.W. Lee, M. Weiss, L.I. Wagner, V.T. Chang, D. Cella, J.B. Manola, L.M. Minasian, W. McCaskill-Stevens, T.R. Mendoza, and C.S. Cleeland. "Prospective, Observational Study of Pain and Analgesic Prescribing in Medical Oncology Outpatients with Breast, Colorectal, Lung, or Prostate Cancer." *Journal of Clinical Oncology*, vol. 30, no. 16, 2012, pp. 1980–1988.

39. Wu, H.S., T. Natavio, J.E. Davis, and H.N. Yarandi. "Pain in Outpatients Treated for Breast Cancer: Prevalence, Pharmacological Treatment, and Impact on Quality of Life." *Cancer Nursing*, vol. 36, no. 3, 2013, pp. 229–235. Available at <http://www.cancernursingonline.com/>. Accessed September 24, 2012.
40. Chapman, S. "Assessment and Management of Patients with Cancer Pain." *Cancer Nursing Practice*, vol. 10, no. 10, 2011, pp. 28–36.
41. Crawford, J.C., D.C. Dale, and G.H. Lyman. "Chemotherapy-Induced Neutropenia." *Cancer*, vol. 15, 2004, pp. 228–237.

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

Not applicable. This is an outcome measure.

Source of Systematic Review: • Title • Author • Date • Citation, including page number • URL	
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.	
Grade assigned to the evidence associated with the recommendation with the definition of the grade	
Provide all other grades and definitions from the evidence grading system	
Grade assigned to the recommendation with definition of the grade	
Provide all other grades and definitions from the recommendation grading system	

Body of evidence: - Quantity – how many studies? - Quality – what type of studies?	
Estimates of benefit and consistency across studies	
What harms were identified?	
Identify any new studies conducted since the SR. Do the new studies change the conclusions from the SR?	

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.

Not applicable. This is an outcome measure.

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

Not applicable. This is an outcome measure.

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

Not applicable. This is an outcome measure.

1b. Performance Gap

Demonstration of quality problems and opportunity for improvement, i.e., data demonstrating:

- considerable variation, or overall less-than-optimal performance, in the quality of care across providers; and/or
- Disparities in care across population groups.

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.

The primary purpose of this measure is to assess the extent to which cancer patients receiving outpatient chemotherapy treatment experience complications resulting in a hospital visit (either an inpatient admission or ED visit). By identifying these events, the measure seeks to encourage quality improvement across facilities to reduce the number of potentially avoidable inpatient admissions and ED visits and increase transparency in the quality of care patients receive. The measure is envisioned to promote effective communication and coordination of care, which is both a Meaningful Measures quality category and a National Quality Strategy priority. It also meets an additional National Quality Strategy priority of promoting the most effective prevention and treatment practices for the leading causes of mortality.

Chemotherapy treatment can have severe, predictable side effects, which, if inappropriately managed, can reduce patients' quality of life and increase healthcare utilization and costs. On average, cancer patients

receiving chemotherapy have one hospital admission and two ED visits per year; approximately 40 percent of these admissions, and 50 percent of these ED visits stem from complications of chemotherapy, respectively [1]. The literature suggests that ten symptoms in particular –anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis –are primary reasons for hospital visits among cancer patients receiving chemotherapy, and are potentially avoidable with proper outpatient management [3 - 5]. Improved management of these symptoms, through improved adherence to clinical treatment guidelines and enhanced care coordination, has been shown to reduce admissions and ED visits and increase patients’ quality of care and quality of life [2] [3] [4].

Admissions and ED visits are costly to payers, with one study estimating that, on average, those experiencing chemotherapy-related adverse events incurred \$12,907 in additional hospitalization expenditures per person per year [6]. In addition to increased cost to payers, unplanned admissions and ED visits related to chemotherapy treatment reduce cancer patients’ quality of life. Measuring potentially avoidable admissions and ED visits for cancer patients receiving outpatient chemotherapy will provide hospitals with an incentive to improve the quality of care for these patients, by taking steps to prevent and better manage side effects and complications from treatment. Hospitals that provide outpatient chemotherapy should implement appropriate care to minimize the incidence of these adverse events and the subsequent need for acute hospital care.

Evidence suggests that coordination of care and better management of these symptoms in the outpatient setting can decrease hospital visits among patients receiving chemotherapy. Studies have indicated that in outpatient settings, where established guidelines are not properly followed and structured protocols are not put into place, there is a higher likelihood for adverse events [7] [8] [9]. This measure will encourage hospitals to use guidelines from the American Society of Clinical Oncology, National Comprehensive Cancer Network, Oncology Nursing Society, Infectious Diseases Society of America, and other professional societies with evidence-based interventions to prevent and treat common side effects and complications of chemotherapy [10]. This risk-standardized measure seeks to increase transparency in the quality of care patients receive, and to provide information to help physicians and hospitals mitigate patients’ need for acute care, which can be a burden on patients, and increase patients’ quality of life [11 – 12].

Citations

1. Vandervelde A, Miller H, Younts J. Impact on Medicare payments of shift in site of care for chemotherapy administration. Washington, DC: Berkeley Research Group; June 2014.
http://www.communityoncology.org/UserFiles/BRG_340B_SiteofCare_ReportF_6-9-14.pdf. Accessed September 16, 2015
2. Burke, T.A., Wisniewski, T., and Ernst, F.R. “Resource utilization and costs associated with chemotherapy-induced nausea and vomiting (CINV) following highly or moderately emetogenic chemotherapy administered in the US outpatient hospital setting.” *Supportive Care in Cancer*, vol. 19, 2011, pp. 131-141.
3. McKenzie, H., L. Hayes, K. White, K. Cox, J. Fethney, M. Boughton, and J. Dunn. “Chemotherapy Outpatients’ Unplanned Presentations to Hospital: A Retrospective Study.” *Support Care Cancer*, vol. 19, 2011, pp. 963–969
4. Foltran, L., Aprile, G., Pisa, F.E. et al. “Risk of unplanned visits for colorectal cancer outpatients receiving chemotherapy: a case-crossover study.” *Supportive Care in Cancer*, September 2014, Volume 22, Issue 9, PP 2527-2533.
5. Daly, B. and Abougergi, M.S. “National trends in admissions for potentially preventable conditions among patients with metastatic solid tumors, 2004-2014.” *Journal of Clinical Oncology* 36, no. 30_suppl (October 20 2018) 1-1.
6. Hassett, M.J., J. O’Malley, J.R. Pakes, J.P. Newhouse, and C.C. Earle. “Frequency and Cost of Chemotherapy-Related Serious Adverse Effects in a Population Sample of Women with Breast Cancer.” *Journal of the National Cancer Institute*, vol. 98, no. 16, 2006, pp. 1108–1117.

7. Abernethy, A.P., Barbour, S.Y., Uronis, H. et al. "Quality management of potential chemotherapy-induced neutropenic complications: evaluation of practice in an academic medical center." *Support Care Cancer* (2009) 17: 735
8. Molassiotis, A., Brearley, S.G. & Stamataki, Z. "Use of antiemetics in the management of chemotherapy-related nausea and vomiting in current UK practice." *Support Care Cancer* (2011) 19: 949.
9. Patil, Vijay M., Vanita Noronha, Amit Joshi, Anant Ramaswamy, Sudeep Gupta, Arvind Sahu, Vipul Doshi, Tarachand Gupta, Sushmita Rath, Shripad Banavali, and Kumar Prabhash. "Adherence to and Implementation of ASCO Antiemetic Guidelines in Routine Practice in a Tertiary Cancer Center in India." *Journal of Oncology Practice*, vol. 13, no. 6, 2017, pp. e574-e581
10. National Comprehensive Cancer Network. "Prevention and Treatment of Cancer-Related Infections." NCCN Clinical Practice Guidelines in Oncology Version 1.2013. 2013. Available at http://www.nccn.org/professionals/physician_gls/f_guidelines.asp. Accessed Sept. 26, 2014.
11. Osoba, D., B. Zee, D. Warr, J. Latreille, L. Kaizer, and J. Pater. "Effect of Postchemotherapy Nausea and Vomiting on Health-Related Quality of Life." *Support Care Cancer*, vol. 5, 1997, pp. 307–313.
12. Wu, H.S., T. Natavio, J.E. Davis, and H.N. Yarandi. "Pain in Outpatients Treated for Breast Cancer: Prevalence, Pharmacological Treatment, and Impact on Quality of Life." *Cancer Nursing*, vol. 36, no. 3, 2013, pp. 229–235. Available at <http://www.cancernursingonline.com/>. Accessed September 24, 2012.

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.*

We assessed hospital-level performance scores using 100% national Medicare Fee-for-Service (FFS) claims and enrollment data for short-term acute hospitals (please see Measure Testing Attachment Section 1.2 and 1.7 for full description of the datasets used).

We estimated the measure score for hospitals using Medicare FFS claims with a performance period of October 1, 2015 to September 30, 2016. We estimated separate scores for qualifying patients receiving outpatient chemotherapy treatment at two facility types: (1) non-cancer hospitals included in calculations for the Outpatient Quality Reporting (OQR) program and (2) Prospective Payment System-Exempt Cancer hospitals (PCHs) participating in the Prospective Payment System-Exempt Cancer Hospital Quality Reporting (PCHQR) program. The total number of hospitals with at least one attributed patient was 3,562 in non-cancer hospitals and 11 in PCHs. The total number of patients meeting inclusion and exclusion criteria across these hospitals was 266,066 non-cancer hospital patients and 23,477 PCH patients.

The risk-standardized inpatient admission rate (RSAR) for non-cancer hospitals ranged from 8.9% to 18.5 % (median 12.5%, 25th and 75th percentiles are 12.2% and 13.0%, respectively) while the risk-standardized inpatient admission rate for PCHs ranged from 12.3% to 15.2% (median 13.7%, 25th and 75th percentiles are 13.4% and 14.8%, respectively).

The risk-standardized ED visit rate (RSEDR) for non-cancer hospitals ranged from 2.9% to 15.2% (median 5.6%, 25th and 75th percentiles are 5.6% and 6.2%, respectively) while the risk-standardized ED visit rate for PCHs ranged from 3.6% to 9.1% (median 6.7%, 25th and 75th percentiles are 4.4% and 8.9%, respectively).

The distributions of facility scores (RSARs for non-cancer and cancer hospitals, RSEDRs for cancer and non-cancer hospitals) are provided below.

Distribution of RSARs and RSEDRs for Non-Cancer and Cancer Hospitals

Non-Cancer RSAR (%)

Minimum: 8.9

1st: 10.2

5th: 11.1
10th: 11.6
25th: 12.2
50th (Median): 12.5
75th: 13.0
90th: 13.9
95th: 14.8
99th: 16.4

Maximum: 18.5

PCHs: RSAR (%)

Minimum: 12.3

1st: 12.3

5th: 12.3

10th: 13.4

25th: 13.4

50th (Median): 13.7

75th: 14.8

90th: 14.8

95th: 15.2

99th: 15.2

Maximum: 15.2

Non-Cancer RSEDR (%)

Minimum: 2.9

1st: 4.2

5th: 4.8

10th: 5.2

25th: 5.6

50th (Median): 5.6

75th: 6.2

90th: 6.8

95th: 7.4

99th: 8.6

Maximum: 15.2

PCHs: RSEDR (%)

Minimum: 3.6

1st: 3.6

5th: 3.6

10th: 4.1

25th: 4.4

50th (Median): 6.7

75th: 8.9

90th: 9.1

95th: 9.1

99th: 9.1

Maximum: 9.1

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.

Not applicable, performance data provided above demonstrating gap.

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.*

Please note that the following describes disparities analyses performed in 2018 using updated measure specifications. For a description of methods and results of the original disparities analyses performed in 2016 please see the testing attachment.

Our analysis of disparities examined the impact of social risk factors on the measure score. We evaluated two indicators of social risk: 1) race, specifically African-American or not and 2) the Agency for Healthcare Research and Quality (AHRQ) Socio-Economic Status (SES) Composite index, which was derived from January 2009 – December 2013 American Community Survey (ACS) data. The AHRQ SES Composite index score is calculated using 7 different variables which generally represent the socio-economic well-being of populations within each zip code in the ACS data. These variables are: (1) median household income, (2) percentage of persons living below the federal poverty level, (3) percentage of persons who are aged >16 years and in the labor force but not employed, (4) median value of owner-occupied homes, (5) percentage of persons aged >25 years who completed at least a 12th grade education, (6) percentage of persons aged >25 years who completed at least four years of college, and (7) percentage of households that average one or more persons per room. SES composite scores range from 0 to 100 with higher scores indicating higher socio-economic well-being and lower scores indicating lower socio-economic well-being. An SES score of below 42.7 is considered “low” socio-economic well-being for the purpose of this analysis. Dual status was evaluated at the time of initial measure development as described in Section 2b3.4b of the Testing Attachment, but was not re-examined for the current measure specification (2018 reevaluation).

These data included 3,562 OPD facilities, 11 PCH facilities, and 289,543 unique patients. Our goal for these analyses were twofold: 1) to examine whether these factors were associated with increased risk in inpatient admissions and ED visits after adjusting for other risk factors and 2) to evaluate the impact of social risk factors on facility-level measure scores. Key findings are detailed below. We examined associations between outcomes and sociodemographic status (SDS) factors using both bivariate and multivariate analyses. At the patient-level, our analysis shows that “low SDS” patients (as characterized by two individual indicators: race as black and low AHRQ SES Composite Index) receiving hospital-based outpatient chemotherapy are more likely to have an inpatient admission and emergency department (ED) visit within 30 days than “non-low SDS” patients.

- Black patients are more likely to have an inpatient admission or ED visit than non-black patients (14.2 percent of black patients versus 12.6 percent of non-black for inpatient admission, and 7.6 percent of black patients versus 5.8 percent of non-black for ED visits)

- Low AHRQ SES Composite Index patients are more likely to have an inpatient admission or ED visit than higher SES Composite Index patients (14.4 percent of patients with low AHRQ SES Composite Index compared to 12.4 percent of patients with higher AHRQ SES Composite Index for inpatient admission, and 7.1 percent of patients with low AHRQ SES Composite Index versus 5.7 percent of patients with high AHRQ SES Composite Index for ED visits).

When evaluating the hospital-level, there was no significant impact of disparities on hospital-level measure scores. No clear relationship between the median risk-standardized rates and hospitals' case mix by these two SDS factors was observed. Additionally, the distributions of risk-standardized rates overlapped significantly across hospitals grouping by these two SDS factors, suggesting that hospitals caring for a greater percentage of low SDS patients have similar rates of inpatient admission and ED visits within 30 days of hospital-based outpatient chemotherapy. See Section 2b4.4b of the Testing Attachment, Section 2b4.4b and in the separate appendix titled "ChemoMeasure_NQF Appendix_SDS" for more information on the analysis and results.

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

Not applicable.

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):

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

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

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.)

<https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1228776192121> - See 2018 Chemotherapy Measure Updates and Specifications Report and accompanying Data Dictionary.

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:

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: [2018_Chemotherapy_Measure_Data_Dictionary_082218-636771841901551813.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:**

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.

No

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.

Not applicable.

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).

This measure involves calculating two mutually exclusive outcomes among cancer patients receiving chemotherapy treatment in a hospital outpatient setting: (1) one or more inpatient admissions for any of the following 10 diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of chemotherapy treatment or (2) one or more ED visits for any of the following 10 diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of chemotherapy treatment. These 10 conditions are potentially preventable through appropriately managed outpatient care. To be counted as an outcome, the qualifying diagnosis on the admission or ED visit claim must be (1) the principal diagnosis or (2) a secondary diagnosis accompanied by a principal diagnosis of cancer.

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).

The chemotherapy measure is a risk-adjusted outcome measure and does not have a traditional numerator like a process measure; thus we use this field to define the measured outcomes of interest as this measure separately reports hospital rates of two outcomes: (1) inpatient admission and (2) ED visits.

Outcome Definition

The chemotherapy measure has two reported outcomes. The outcomes for this measure are:

(1) one or more inpatient admissions for any of the following 10 diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of chemotherapy treatment, and (2) one or more ED visits without an admission, for one of the 10 following diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of receiving hospital-based outpatient chemotherapy treatment for cancer. These 10 conditions are potentially preventable through appropriately managed outpatient care.

Outcome Identification and Counting

Outcomes are identified using Medicare Part A Inpatient and Part B Outpatient hospital claims. The qualifying diagnosis on the admission or ED visit claim must be (1) the principal diagnosis or (2) a secondary diagnosis

accompanied by a principal diagnosis of cancer. The ICD-9 and ICD-10-CM codes that identify these diagnoses are in the 2018 Chemotherapy Measure_Data Dictionary on sheets “S.6 Numerator-Anemia,” “S.6 Numerator-Dehydration,” “S.6 Numerator-Diarrhea,” “S.6 Numerator-Emesis,” “S.6 Numerator-Fever,” “S.6 Numerator-Nausea,” “S.6 Numerator-Neutropenia,” “S.6 Numerator-Pain,” “S.6 Numerator-Pneumonia,” and “S.6 Numerator-Sepsis.” The ICD-9 codes were used during development and testing of the measure; the Data Dictionary also includes the mapping from these ICD-9 codes to ICD-10 codes.

Inpatient admissions that are considered always planned do not qualify for the measure. Planned admissions are defined as those planned by providers for anticipated medical treatment or procedures that must be provided in the inpatient setting. The measure counts only unplanned admissions in the measure outcome because variation in planned admissions does not reflect quality differences. For the chemotherapy measure, inpatient hospital admissions with the following AHRQ CCS procedures or diagnoses are considered always planned and do not qualify for the measure:

Procedures

- 64 – Bone marrow transplant
- 105 – Kidney transplant
- 176 – Other organ transplantation (other than bone marrow corneal or kidney)

Diagnoses

- 45 – Maintenance chemotherapy; radiotherapy
- 254 – Rehabilitation care; fitting of prostheses; and adjustment of devices

Outcomes are counted separately for the inpatient admission and ED visit categories; a patient can only qualify for an outcome in either category, but not both. Patients who experience both an inpatient admission and an ED visit during the performance period are counted towards the inpatient admission outcome. Among those with no qualifying inpatient admissions, qualifying ED visits are counted. As a result, the rates can be viewed as additive to provide a comprehensive performance estimate of quality of care following hospital-based outpatient chemotherapy treatment. The rates are calculated separately because the severity and cost of an inpatient admission is different from that of an ED visit, but both adverse events are important signals of quality and represent important outcomes of care.

Outcome Time Frame

The measure limits the outcome time frame to the 30 days following the date of each chemotherapy treatment (including the day of treatment) in an outpatient setting for four reasons. First, existing literature suggests the vast majority of adverse events occur within 30 days after treatment [1, 2, 3, 4], indicating that a 30-day period is a reasonable timeframe to observe the side effects of treatment. Second, we observed in our own data that the highest rates of hospital visits occur within 30 days after chemotherapy treatment. Third, restricting the time period ensures that patients’ experiences are attributed to the hospitals that provided their recent treatment while accounting for variations in duration between outpatient treatments. Fourth, relating the time frame to a specific chemotherapy administration supports the idea that the admission stems from the management of side effects of treatment and ongoing care, rather than progression of the disease or other unrelated events.

Citations

1. Aprile, G., F.E. Pisa, A. Follador, L. Foltran, F. De Pauli, M. Mazzer, S. Lutrino, C.S. Sacco, M. Mansutti, and G. Fasola. “Unplanned Presentations of Cancer Outpatients: A Retrospective Cohort Study.” *Supportive Care in Cancer*, vol. 21, no. 2, 2013, pp. 397–404.
2. Foltran, L., G. Aprile, F.E. Pisa, P. Ermacora, N. Pella, E. Iaiza, E. Poletto, S.E. Lutrino, M. Mazzer, M. Giovannoni, G.G. Cardellino, F. Puglisi, and G. Fasola. “Risk of Unplanned Visits for Colorectal Cancer Outpatients Receiving Chemotherapy: A Case-Crossover Study.” *Supportive Care in Cancer*, vol. 22, no. 9, 2014, pp. 2527–2533.

3. McKenzie, H., L. Hayes, K. White, K. Cox, J. Fethney, M. Boughton, and J. Dunn. "Chemotherapy Outpatients' Unplanned Presentations to Hospital: A Retrospective Study." *Supportive Care in Cancer*, vol. 19, no. 7, 2011, pp. 963–969.
4. Oatley, M., M. Fry, and L. Mullen. "A Cross-Sectional Study of the Clinical Characteristics of Cancer Patients Presenting to One Tertiary Referral Emergency Department." *International Emergency Nursing*, vol. 24, 2016, pp. 35 – 38.

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

The measure cohort includes Medicare Fee-for-Service (FFS) patients, aged 18 years and older at the start of the performance period, with a diagnosis of any cancer (except leukemia), who received at least one outpatient chemotherapy treatment at the reporting hospital during the performance period.

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).

The target population is Medicare Fee-for-Service (FFS) patients aged 18 and older with a diagnosis of cancer receiving chemotherapy treatment in a hospital outpatient setting at any point during the measurement year.

The measure uses the International Classification of Diseases, Ninth Revision, Clinical Modification (ICD-9) and ICD-10 codes that identify cancer diagnoses. The measure identifies chemotherapy treatment using ICD-9 and ICD-10 procedure and encounter codes; and Current Procedural Terminology (CPT®)/Healthcare Common Procedure Coding System (HCPCS) procedure and medication procedure codes.

Code sets used for cohort identification are attached in the 2018 Chemotherapy Measure_Data Dictionary on sheets "S.9 Denominator-Cancer," "S.9 Denominator-Chemo Procedure," "S.9 Denominator – Chemo Encounter," and "S.9 Denominator – Chemo Medicine".

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

The measure excludes the following patients from the cohort:

- 1) Patients with a diagnosis of leukemia at any time during the performance period.
- 2) Patients who were not enrolled in Medicare FFS Parts A and B in the year prior to the any outpatient chemotherapy treatment during the performance period.
- 3) Patients who were not enrolled in Medicare FFS Parts A and B for the 30 days following any chemotherapy treatment.
- 4) Cases in which patients receive chemotherapy to treat conditions other than cancer. Note that this is a case-level exclusion; as long as the patient has additional cases that meet inclusion criteria, they will remain in the cohort.

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.)*

- 1) Patients with a diagnosis of leukemia at any time during the performance period – exclusions are identified using the codes listed in the 2018 Chemotherapy Measure Data Dictionary on sheet "S.11 Denominator Exclusion – Leukemia." If a patient has a claim with any of the diagnosis codes within the code set, at any point during the performance period, they are excluded from the cohort.

Rationale: Patients with leukemia are excluded due to the high toxicity of treatment and recurrence of disease so that admissions do not reflect poorly managed outpatient care for this population. Patients with leukemia

have an expected admission rate due to relapse, so including leukemia patients in the cohort could be conceptualized as a planned admission, which does not align with the intent of the measure.

2) Patients who were not enrolled in Medicare FFS Parts A and B in the year prior to any outpatient chemotherapy treatment during the performance period. The Medicare Enrollment database is used to determine if a patient was enrolled in Medicare FFS Parts A and B in the year prior to the first outpatient chemotherapy treatment during the performance period.

Rationale: We exclude these patients to ensure complete patient diagnosis data for the risk-adjustment models, which use the year prior to the chemotherapy treatment during the period to identify comorbidities.

3) Patients who do not have at least one outpatient chemotherapy treatment followed by continuous enrollment in Medicare FFS Parts A and B in the 30 days after the procedure. The Medicare Enrollment database is used to determine if a patient was enrolled in Medicare FFS Parts A and B in the 30-days after a qualifying outpatient chemotherapy treatment during the performance period.

Rationale: We exclude these patients to ensure full data availability for outcome assessment.

4) Cases in which patients receive chemotherapy to treat conditions other than cancer. If a case includes a chemotherapy procedure code from the “S.11 Denominator Exclusion – ChemoNonCancer” code set, a diagnosis code from the “S.11 Denominator Exclusion - AutoImmuneDiags” code set, and no cancer diagnosis from the “S.9 Denominator-Cancer” code set in any position on the claim, the case is excluded from the cohort. Note that this is a case-level exclusion; as long as the patient has additional cases that meet inclusion criteria, they will remain in the cohort.

Rationale: We exclude these patients because cases where chemotherapy is administered for non-cancer conditions, such as treatment of auto-immune diseases, is not aligned with the measure’s intent. The measure is intended to assess the quality of care provided to cancer patients receiving outpatient chemotherapy.

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.)*

Not applicable. This measure is not stratified.

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

Statistical risk model

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 = Lower 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.)*

Calculation of the Observed Rate

Denominator

Steps to Identify Cohort

Step 1: Identify all Medicare Fee-for-Service (FFS) patients age 18 and older with a diagnosis of cancer receiving chemotherapy treatment in a hospital outpatient setting during the performance period.

Step 2: Remove all patients with a diagnosis of leukemia at any time during the performance period.

Step 3: Remove all chemotherapy cases that are not preceded by 12 months of Medicare FFS Parts A and B.

Step 4: Remove all chemotherapy cases that are not followed by continuous enrollment in Medicare FFS Parts A and B in the 30 days after the treatment.

Step 5: Remove all cases in which patients receive chemotherapy to treat a qualifying autoimmune condition, rather than to treat cancer. Note that this is a case-level exclusion; as long as the patient has additional cases that meet inclusion criteria, they will remain in the cohort.

Step 6: Identify the unique number of patient-level provider ID/Facility ID combinations for the remaining cases.

Step 7: The remaining unique patients the measure denominator (cohort) at each facility.

Numerator

Steps to Identify Qualifying Inpatient Hospital Admissions and ED Visits

Step 1: Identify the first qualifying outpatient chemotherapy administration for each patient in each facility. [Note: a patient may be included at multiple facilities.]

Step 2: Determine whether that outpatient chemotherapy treatment was followed by either an inpatient hospital admission or ED visit within 30 days with either:

- A primary diagnosis of anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis, or
- A primary diagnosis of cancer and a secondary diagnosis of anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis

Step 3: Remove any qualifying inpatient admissions with an “always planned” diagnosis or procedure.

Step 4: If a patient had both a qualifying inpatient admission and an ED visit within 30 days, select the inpatient admission.

Step 5: If a patient multiple qualifying inpatient admissions, select the first one.

Step 6: Sum the number of patients in the cohort with an inpatient admission. This is the numerator for the inpatient admissions outcome.

Step 7: Sum the number of patients in the cohort who had an ED visit, but no inpatient admission. This is the numerator for the ED visit outcome.

Calculation of the Observed Performance Rate

Calculate the inpatient admissions observed rate by dividing the number of patients with an inpatient hospital admission by the total number of patients in the cohort for a given facility.

Calculate the ED visits observed rate by dividing the number of patients with an ED visit by the total number of patients in the cohort for a given facility.

Calculation of the Predicted and Expected Rates

The measure’s two-level hierarchical logistic regression model accounts for the clustering of patients within hospitals and variation in sample size. The measure calculates the hospital-specific risk-adjusted rate as the ratio of a hospital’s “predicted” number of outcomes to “expected” number of outcomes multiplied by the national observed outcome rate.

- **Predicted Rate:** The measure estimates the predicted number of outcomes for each hospital using the same patient mix, but an estimated hospital-specific intercept. It calculates the predicted number of outcomes for each hospital by summing the predicted probabilities for all patients in the hospital. The measure

calculates the predicted probability for each patient through the hierarchical model, which applies the estimated regression coefficients to the observed patient characteristics and adds the hospital-specific intercept.

- **Expected Rate:** This rate estimates the expected number of outcomes for each hospital using the hospital's patient mix and the average hospital-specific intercept (that is, the average intercept among all hospitals in the sample). Operationally, the measure obtains the expected number of outcomes for each hospital by summing the expected probabilities of outcomes for all patients treated at the hospital. It calculates the expected probability of outcomes for each patient via the hierarchical model, which applies the estimated regression coefficients to the observed patient characteristics and adds the average of the hospital-specific intercept.

If a hospital's ratio of predicted to expected outcomes is less than 1, it indicates that the hospital is performing better than expected given its case mix. If a hospital's ratio of predicted to expected outcomes is greater than 1, it indicates that the hospital is performing worse than expected given its case mix. The risk factors included in the Inpatient Admission and ED Visit models are listed below.

Inpatient Admission Model Variables

The patient-level risk-adjustment variables are:

1. Age (continuous)
2. Sex (male)
3. Number of Outpatient Chemotherapy Treatments
4. Receipt of Concurrent Radiotherapy
5. Respiratory Disorder
6. Renal Disease
7. Diabetes
8. Other Injuries
9. Metabolic Disorder
10. Gastrointestinal Disorder
11. Psychiatric Disorder
12. Neurological Conditions
13. Cardiovascular Disease
14. Breast Cancer
15. Digestive Cancer
16. Respiratory Cancer
17. Lymphoma
18. Prostate Cancer
19. Secondary Cancer of Lymph Nodes
20. Secondary Cancer of Solid Tumors
21. Other Cancer

ED Visits Model Variables

The patient-level risk-adjustment variables are:

1. Age (years above 18, continuous)

2. Sex (male)
3. Number of Outpatient Chemotherapy Treatments
4. Receipt of Concurrent Radiotherapy
5. Respiratory Disorder
6. Other Injuries
7. Gastrointestinal Disorder
8. Psychiatric Disorder
9. Neurological Conditions
10. Cardiovascular Disease
11. Breast Cancer
12. Digestive Cancer
13. Respiratory Cancer
14. Secondary Cancer of Lymph Nodes
15. Secondary Cancer of Solid Tumors
16. Other Cancer

Calculation of the Risk-Adjusted Rates

The risk-standardized admissions rate (RSAR) is calculated as the ratio of the number of “predicted” qualifying inpatient admissions to the number of “expected” qualifying inpatient admissions multiplied by the national observed qualifying inpatient admission rate. Similarly, the risk-standardized ED visits rate (RSEDR) is calculated as the ratio of the number of “predicted” qualifying ED visits to the number of “expected” qualifying ED visits multiplied by the national observed qualifying ED visit rate.

For each rate, this approach is analogous to a ratio of “observed” to “expected” outcomes used in other types of statistical analyses. It conceptually allows for a comparison of a particular facility’s performance given its case mix to an average facility’s performance with the same case mix. Thus, a predicted/expected ratio of less than one indicates a lower-than-expected visit rate (or better quality), and a ratio of greater than one indicates a higher-than-expected visit rate (or worse quality).

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.

This measure is not based on a sample or survey.

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.

This measure is not based on a sample or survey.

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

If other, please describe in S.18.

Claims, Enrollment 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.

The numerator (outcome), denominator (cohort), and risk factors for this measure are based on Medicare administrative claims and enrollment data.

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)

Facility

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.)

Not applicable.

2. Validity – See attached Measure Testing Submission Form

[Chemotherapy_TestingAttachment_081718-636771081788287184.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.

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.

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.

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

NATIONAL QUALITY FORUM—Measure Testing (subcriteria 2a2, 2b1-2b6)

Measure Number (if previously endorsed): N/A

Measure Title: Admissions and Emergency Department Visits for Patients Receiving Outpatient Chemotherapy

Date of Submission: [8/1/2018](#)

Type of Measure:

☒ Outcome (including PRO-PM)	☐ Composite – STOP – use composite testing form
--	--

☐ Intermediate Clinical Outcome	☐ Cost/resource
☐ Process (including Appropriate Use)	☐ Efficiency
☐ Structure	

Instructions

1. Measures must be tested for all the data sources and levels of analyses that are specified. ***If there is more than one set of data specifications or more than one level of analysis, contact NQF staff*** about how to present all the testing information in one form.
2. For all measures, sections 1, 2a2, 2b1, 2b2, and 2b4 must be completed.
3. For outcome and resource use measures, section 2b3 also must be completed.
4. If specified for multiple data sources/sets of specifications (e.g., claims and EHRs), section 2b5 also must be completed.
5. Respond to all questions as instructed with answers immediately following the question. All information on testing to demonstrate meeting the subcriteria for reliability (2a2) and validity (2b1-2b6) must be in this form. An appendix for *supplemental* materials may be submitted, but there is no guarantee it will be reviewed.
6. If you are unable to check a box, please highlight or shade the box for your response.
7. Maximum of 25 pages (including questions/instructions; minimum font size 11 pt; do not change margins). ***Contact NQF staff if more pages are needed.***
8. Contact NQF staff regarding questions. Check for resources at [Submitting Standards webpage](#).
9. For information on the most updated guidance on how to address social risk factors variables and testing in this form refer to the release notes for version 7.1 of the Measure Testing Attachment.

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

2a2. Reliability testing ¹⁰ demonstrates the measure data elements are repeatable, producing the same results a high proportion of the time when assessed in the same population in the same time period and/or that the measure score is precise. For **instrument-based measures** (including PRO-PMs) and **composite performance measures**, reliability should be demonstrated for the computed performance score.

2b1. Validity testing ¹¹ demonstrates that the measure data elements are correct and/or the measure score correctly reflects the quality of care provided, adequately identifying differences in quality. For **instrument-based measures (including PRO-PMs) and composite performance measures**, validity should be demonstrated for the computed performance score.

2b2. Exclusions are supported by the clinical evidence and are of sufficient frequency to warrant inclusion in the specifications of the measure; ¹²

AND

If patient preference (e.g., informed decisionmaking) is a basis for exclusion, there must be evidence that the exclusion impacts performance on the measure; in such cases, the measure must be specified so that the information about patient preference and the effect on the measure is transparent (e.g., numerator category computed separately, denominator exclusion category computed separately). ¹³

2b3. For outcome measures and other measures when indicated (e.g., resource use):

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

OR

2. rationale/data support no risk adjustment/ stratification.

2b4. Data analysis of computed measure scores demonstrates that methods for scoring and analysis of the specified measure allow for **identification of statistically significant and practically/clinically meaningful differences in performance**;

OR

there is evidence of overall less-than-optimal performance.

2b5. If multiple data sources/methods are specified, there is demonstration they produce comparable results.

2b6. Analyses identify the extent and distribution of **missing data** (or nonresponse) and demonstrate that performance results are not biased due to systematic missing data (or differences between responders and nonresponders) and how the specified handling of missing data minimizes bias.

Notes

10. Reliability testing applies to both the data elements and computed measure score. Examples of reliability testing for data elements include, but are not limited to: inter-rater/abstractor or intra-rater/abstractor studies; internal consistency for multi-item scales; test-retest for survey items. Reliability testing of the measure score addresses precision of measurement (e.g., signal-to-noise).

11. Validity testing applies to both the data elements and computed measure score. Validity testing of data elements typically analyzes agreement with another authoritative source of the same information. Examples of validity testing of the measure score include, but are not limited to: testing hypotheses that the measures scores indicate quality of care, e.g., measure scores are different for groups known to have differences in quality assessed by another valid quality measure or method; correlation of measure scores with another valid indicator of quality for the specific topic; or relationship to conceptually related measures (e.g., scores on process measures to scores on outcome measures). Face validity of the measure score as a quality indicator may be adequate if accomplished through a systematic and transparent process, by identified experts, and explicitly addresses whether performance scores resulting from the measure as specified can be used to distinguish good from poor quality. The degree of consensus and any areas of disagreement must be provided/discussed.

12. Examples of evidence that an exclusion distorts measure results include, but are not limited to: frequency of occurrence, variability of exclusions across providers, and sensitivity analyses with and without the exclusion.

13. Patient preference is not a clinical exception to eligibility and can be influenced by provider interventions.

14. Risk factors that influence outcomes should not be specified as exclusions.

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

1. DATA/SAMPLE USED FOR ALL TESTING OF THIS MEASURE

Often the same data are used for all aspects of measure testing. In an effort to eliminate duplication, the first five questions apply to all measure testing. If there are differences by aspect of testing, (e.g., reliability vs. validity) be sure to indicate the specific differences in question 1.7.

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: Enrollment database files	☒ other: Enrollment database files

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).

The measure requires a data source that allows us to link patient data across care settings in order to identify qualifying patients receiving chemotherapy in hospital outpatient departments (OPDs) for inclusion, comorbidities for risk adjustment, and the outcomes of inpatient hospital admissions and emergency department (ED) visits. Therefore, we used Medicare Fee-for-Service (FFS) claims and enrollment data, as they support these linkages and were available for the population of interest.

1. The primary dataset used to support measure testing included Fiscal Year 2016 (October 1, 2015 – September 2016) Medicare Outpatient, Inpatient, and Carrier (Part B Physician) claims and enrollment data from the Health Account Joint Information (HAJI) database. This same dataset was utilized to support the measure’s August 2017 Dry Run, when facilities had the opportunity to review measure results prior to future reporting in the PPS-Exempt Cancer Hospital Reporting (PCHQR) and Hospital Outpatient Quality Reporting (OQR) programs. The measure has been adopted for public reporting in the PCHQR and Hospital OQR programs beginning in FY 2019 and CY 2020, respectively (81 FR 57190 and 81 FR 79764).

a) Datasets used to define the cohort:

- Outpatient chemotherapy procedures performed at qualifying PPS-Exempt Cancer Hospitals (PCHs) and non-cancer hospital outpatient departments were identified using Outpatient and Inpatient hospital claims data from FY 2016. Inpatient hospital claims data are used to capture outpatient chemotherapy treatment that may be bundled on an inpatient claim due to the CMS 3-day payment window policy; only chemotherapy procedures occurring within the 3-day window prior to an inpatient admission are included.
- Outpatient hospital, Inpatient hospital, and Carrier (Part B Physician) claims were also used to identify cancer diagnoses during FY 2016.
- Medicare Enrollment Database data was used to determine Medicare Fee-For-Service (FFS) enrollment status, demographic, and death information during FY 2016.

b) Datasets used to capture the outcome (inpatient hospital admissions and ED visits)

- Inpatient and outpatient hospital claims data from FY 2016 were used to identify qualifying hospital admissions and ED visits, respectively.

- Qualifying inpatient hospital admissions and ED visits are those that occur within 30 days of a qualifying chemotherapy procedure with either: (1) a primary discharge diagnosis of anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis, or (2) a primary discharge diagnosis of cancer and a secondary diagnosis of one of those 10 diagnoses on the same claim.

c) Datasets used to identify comorbidities for risk adjustment:

- Inpatient hospital, Outpatient hospital, and Carrier (Part B Physician) claims were used to identify comorbidities during the prior 365 days (Fiscal Year 2015: October 1, 2014 – September 30, 2015) for risk adjustment for these patients.

2. We also utilized datasets from other performance periods to support other aspects of measure development and testing, as follows:

- FY 2016 - 2017 (October 1, 2015 – September 30, 2017) data were used to calculate split-sample reliability statistics, as detailed in Section 2.a.2 Reliability.
- July 2012 – June 2013 data were used to support the development and testing of the initial risk-adjustment models, as described in Section 2.b.3 Risk-Adjustment/Stratification. These data were derived from the Medicare Standard Analytic Files, rather than HAJI, but were otherwise identical in terms of identifying the measure cohort, outcomes, and comorbidities for risk adjustment.

To evaluate the inclusion of social risk factors/socio-demographic status (SDS) in our risk-adjustment algorithms, we used American Community Survey (ACS) data from the United States Census Bureau to derive the Agency for Healthcare Research and Quality (AHRQ) Socio-economic status (SES) index for each patient ZIP code. The following datasets were used:

- January 2008 – December 2012 ACS data were used to calculate the AHRQ SES Index to evaluate the inclusion of social risk factors during the initial development of the risk-adjustment algorithms.
- January 2009 – December 2013 ACS data were used to calculate the AHRQ SES Index to re-evaluate the inclusion of social risk factors with the risk-adjustment algorithms during the measure's 2018 re-evaluation cycle.

1.3. What are the dates of the data used in testing? The datasets vary by testing type, as described above. The data used to test the chemotherapy measure spanned July 1, 2012 – September 30, 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)

The number of measured entities (outpatient departments at PCHs and non-cancer hospitals) varies by testing type; see Section 1.7 for details.

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)*

The number of patients varies by testing type; see Section 1.7 for details

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.

We used Medicare data described in Section 1.2 to develop and test aspects of this measure. We applied the measure specifications to each dataset to develop the measure's cohort (see the Notice of Intent to Submit or Measure Submission Forms, Sections S.4 to S.11, for full inclusion and exclusion criteria) and test the measure. We present measure information testing results for both types of facilities scored on the measure (PCHs, non-cancer hospital OPDs) as well as separately for the PCHs and non-cancer hospital OPDs to reflect how the measure will be publicly reported. The information below represents the final cohort, after measure exclusions were applied.

The datasets, number of measured entities, and demographic profiles for the patients used in each type of testing are as follows:

(1) Medicare FFS FY 2016 dataset

- **Dates:** October 1, 2015 - September 30, 2016
- **Dataset used for:** testing facility-level reliability (Section 2a2), analyses to address potential threats to validity (Section 2b1), face validity review (Section 2b1), testing the exclusion criteria (Section 2b2), re-evaluation of risk-adjustment algorithm (Section 2b3), and demonstrating meaningful differences in performance (see Section 2b4).

Table 1. Number of patients and patient characteristics

	Overall	PCHs	Non-Cancer Hospitals
Hospitals (n)	3,573	11	3,562
Patients (n)	289,543	23,477	266,066
Age (average)	72.1	71.6	72.1
Male (%)	51.5	54.6	51.3
Chemotherapy Exposure: Number of Treatments during Performance Period			
25th Percentile	2	2	1
Median	4	4	3
75th Percentile	7	9	7
Most Frequent Cancer Diagnoses			
Solid Tumor	44.2	58.7	42.9
Other Cancers	28.0	33.9	27.4
Digestive Cancer	20.7	22.3	20.5
Lymph Node	19.9	37.2	18.4

See attached Data Dictionary for Cancer category definitions.

(2) Medicare FFS FY 2016 – 2017

- **Dates:** October 1, 2015 - September 30, 2017
- **Dataset used for:** conducting split-sample reliability (see Section 2a2).

Table 2. Number of patients and patient characteristics

	Sample 1: PCHs	Sample 2: PCHs	Sample 1: Non-Cancer Hospitals	Sample 2: Non-Cancer Hospitals
Hospitals (n)	11	11	1,439	1,439
Patients (n)	19,316	19,323	204,160	204,898
Age (average)	71.4	71.4	72.0	71.9
Male (%)	54.3	53.9	51.0	51.1
Chemotherapy Exposure: Number of Treatments during Performance Period				
25th Percentile	2	2	1	1
Median	4	4	2	2
75th Percentile	9	9	7	7
Most Frequent Cancer Diagnoses				
Solid Tumor	56.5	56.5	42.4	42.3
Other Cancers	34.4	34.8	28.4	28.3
Digestive Cancer	22.9	21.8	21.1	21.0
Lymph Node	37.0	36.8	19.9	19.7

See attached Data Dictionary for Cancer category definitions.

(3) Medicare FFS July 2012 – June 2013 Datasets

- **Dates:** July 1, 2012 - June 30, 2013
- **Dataset used for:** the development and testing of the initial risk-adjustment models (see Section 2b3).

Table 3. Number of patients and patient characteristics for 2012 – 2013 Full Sample

	Overall	PCHs	Non-Cancer Hospitals
Hospitals (n)	3,765	11	3,754
Patients (n)	240,446	18,400	223,719
Age (average)	72.2	71.6	72.2
Male (%)	50.2	54.6	49.8
Chemotherapy Exposure: Number of Treatments during Performance Period			
25th Percentile	2	2	1
Median	3	4	3
75th Percentile	7	8	7
Most Frequent Cancer Diagnoses			
Solid Tumors	42.2	60.4	40.9

Other Cancer	39.8	45.5	39.4
Digestive Cancer	24.2	27.8	23.9
Respiratory Cancer	21.8	21.6	21.8

See attached Data Dictionary for Cancer category definitions.

Table 4. Number of patients and patient characteristics for 2012 – 2013 Development and Validation Split Samples

	Development	Validation
Hospitals (n)	3,483	3,469
Patients (n)	123,149	123,115
Age (average)	72.2	72.1
Male (%)	50.1	50.0
Chemotherapy Exposure: Number of Treatments during Performance Period		
25th Percentile	1	1
Median	3	3
75th Percentile	7	7
Most Frequent Cancer Diagnoses		
Solid Tumors	42.4	42.5
Other Cancer	28.2	28.3
Digestive Cancer	24.3	24.1
Respiratory Cancer	21.8	21.7

See attached Data Dictionary for Cancer category definitions.

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.

We considered social risk factors that align with the NQF guidelines for developers and can be linked to claims data. The NQF-convened Expert Panel that considered risk-adjustment for social risk factors recognized risk adjustment may be constrained by data limitations and data collection burden [1]. The variables that are available within, or that can be linked directly, to Medicare administrative claims data used for this measure include those listed below. In selecting variables for analysis, our intent was to be responsive to the National Quality Forum (NQF) guidelines for measure developers and the findings of recent work funded by the IMPACT Act [2,3]. Our approach was to examine patient-level indicators of both socio-economic status (SES) and race that are reliably available for all Medicare beneficiaries and linkable to claims data and to select those that have established validity.

As detailed below and in section 2b3.4b, we considered one patient-level social risk factor variable (African-American race), Medicare-Medicaid dual-eligibility status, and a composite measure (the AHRQ-validated SES index score). In addition, we examined the facility-level proportions of African-American patients, facility-level proportions of dual-eligible patients, and of low-SES patients based on the AHRQ SES index. These analyses were performed with the Medicare FFS July 2012 – June 2013 and datasets and 2008 – 2012 data from the

Census Bureau's American Community Survey. We also repeated these evaluations using Medicare FY 2016 data and 2009 – 2013 American Community Survey data (described in Section 1.7).

Potential pathways for SES and race variables' effects are described below in Section 2b3.3a.

The SES and race variable that we examined, further described below, are:

- African-American race
- Dual-eligible status. Note, this analysis was only performed with the July 2012 – June 2013 dataset, as dual-eligibility status is not routinely provided in the HAJI data that are now used to support measure calculations for CMS outpatient hospital quality reporting programs.
- AHRQ-validated SES Index score

African-American race (black, other)

Data source: Medicare enrollment database

The particular case of race as a predictor of health outcomes illuminates the complexity of the role social risk factor variables play in assessing hospital performance. The association between patient race and high symptom burden has been observed among cancer patients [4,5]. Those that are undertreated for their symptoms might be more likely to seek care in the ED or be admitted to the hospital compared to those with adequate symptom management. It is possible this association may also confound with other social risk factors such as SES and geographic access to care.

Medicaid dual-eligible status (Medicaid-Medicare dual, Medicare only)

Data source: Medicare enrollment database

The dual-status patient-level variable provides a reliably-obtained indication of patients with low income/assets and high health care spending. Income is an important social risk factor to consider for this measure because, for example, it may reflect access to resources, ability to purchase medications to manage symptoms, adherence to scheduled follow-up appointments for routine check-ins and timely care, and availability of family support that may impact the measure outcome [2]

AHRQ-validated SES Index score: neighborhood SES factors as proxies for patient-level SES [6]

Data source: Enrollment database and Census data (American Community Survey)

The American Community Survey (ACS) provides a number of social risk indicators that are available at the ZIP code level and can be linked directly to Medicare claims at the 9-digit ZIP code level. We used the Agency for Healthcare Research and Quality (AHRQ)-validated composite index of SES which has been used and tested among Medicare beneficiaries [1]. This index is a composite of seven different variables found in the Census data which may capture social risk better than any single variable. The variables are: (1) median household income, (2) percentage of persons living below the federal poverty level, (3) percentage of persons who are aged >16 years and in the labor force but not employed, (4) median value of owner-occupied homes, (5) percentage of persons aged >25 years who completed at least a 12th grade education, (6) percentage of persons aged >25 years who completed at least four years of college, and (7) percentage of households that average one or more persons per room.

This neighborhood-level variable, which we use as a proxy for patient-level social risk factors, are important to consider for this measure because they may reflect patient income (discussed above), patient's health literacy level, which is associated with higher ED use [2, 7], and home environment, where outpatients handle their care out of sight of the hospital.

References:

1. National Quality Forum. "Risk Adjustment for Socioeconomic Status or Other Sociodemographic Factors: Technical Report." August 15, 2014.

http://www.qualityforum.org/Publications/2014/08/Risk_Adjustment_for_Socioeconomic_Status_or_Other_Sociodemographic_Factors.aspx

2. Department of Health and Human Services, Office of the Assistant Secretary of Planning and Evaluation. Report to Congress: Social Risk factors and Performance Under Medicare's Value-based Payment Programs. 2016; <https://aspe.hhs.gov/pdf-report/report-congress-social-risk-factors-and-performance-under-medicare-value-based-purchasing-programs>. Accessed November 10, 2017.
3. National Academies of Sciences, Engineering, and Medicine (NASEM); Accounting for Social Risk Factors in Medicare Payment: Data. Washington DC: National Academies Press; 2016.
4. Fisch, M.J., J.W. Lee, M. Weiss, L.I. Wagner, V.T. Chang, D. Cella, J.B. Manola, L.M. Minasian, W. McCaskill-Stevens, T.R. Mendoza, and C.S. Cleeland. "Prospective, Observational Study of Pain and Analgesic Prescribing in Medical Oncology Outpatients with Breast, Colorectal, Lung, or Prostate Cancer." *Journal of Clinical Oncology*, vol. 30, no. 16, 2012, pp. 1980–1988.
5. Miaskowski, C., B. A. Cooper, M. Melisko, L. M. Chen, J. Mastick, C. West, S. M. Paul, L. B. Dunn, B. L. Schmidt, M. Hammer, F. Cartwright, F. Wright, D. J. Langford, K. Lee, and B. E. Aouizerat. "Disease and Treatment Characteristics do Not Predict Symptom Occurrence Profiles in Oncology Outpatients Receiving Chemotherapy." *Cancer*, vol. 120, no. 15, 2014, pp. 2371-2378.
6. Creation of New Race-Ethnicity Codes and Socioeconomic Status (SES) Indicators for Medicare Beneficiaries: Final Report. August 2012. Agency for Healthcare Research and Quality, Rockville, MD. <http://archive.ahrq.gov/research/findings/final-reports/medicareindicators/index.html>. Accessed August 24, 2015.
7. Griffey, Richard T., Sarah K. Kennedy, Lucy McGownan, Melody Goodman, and Kimberly A. Kaphingst. "Is Low Health Literacy Associated with Increased Emergency Department Utilization and Recidivism?" *Academic Emergency Medicine*, vol. 21, no. 10, 2014, pp. 1109-1115.

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)

Measure Score Reliability

Split-Sample Reliability. We tested the reliability of the facility measure score by calculating the intra-class correlation coefficient (ICC) of the measure score using a split-sample (i.e., test-retest) approach. To calculate the ICC, we used the Medicare FFS FY 2015-2017 Dataset. The two years of data from the Medicare FFS FY 2015-2017 Dataset were then randomly split into two samples (1 year of combined data for each sample). We calculated ICC (2,1) described by Shrout and Fleiss, which evaluates the agreement between the risk-standardized admissions rates (RSARs) and risk-standardized emergency department rates (RSEDR) calculated in the two randomly selected samples assuming that all patients are rated by the same raters who are assumed to be a random subset of all possible raters [1, 2]. The formula for ICC (2,1) described in Shrout & Fleiss (1979) utilizes a two-way ANOVA to calculate the ICC as a measure of absolute agreement. The formula is implemented as follows:

$$ICC(2,1) = \frac{MS_R - MS_E}{MS_R + (k - 1)MS_E + k(MS_C - MS_E)/N}$$

where MS_R = Between Measure Mean Square Error, MS_C
= Between Sample Mean Square Error, MS_E = Average Error Variance, k
= # of samples = 2, and N = # of facilities

The split-sample reliability testing methods were aligned with the specifications used for measure implementation in CMS's PCHQR and OQR programs. All patients meeting the measure inclusion and exclusion criteria were included in the split-sample measure score calculations to ensure that the measure cohort was as comprehensive as possible. However, because CMS has determined that measure scores cannot be reliably determined for facilities with fewer than 25 patients, the ICC analysis was limited to hospitals with at least 25 patients in each of the split samples. This approach is consistent with CMS's current public reporting strategy for the PCHQR and OQR programs that includes smaller hospitals in the measure calculation, but does not publicly release the measure score for hospitals with fewer than 25 patients (i.e., labels them in public reporting as having "too few cases" to support a reliable estimate). We note that the minimum sample size for public reporting is a policy choice that balances competing considerations such as the reliability of the measure score and transparency for consumers.

Facility-Level Reliability. Second, we estimated the facility-level reliability. While split-sample reliability is the most relevant metric from the perspective of overall measure reliability, it is also meaningful to consider the separate notion of "unit" reliability, that is, the reliability with which individual units (here, hospital outpatient departments) are measured. This is because the reliability of any one facility's measure score will vary depending on the number of patients receiving chemotherapy. Facilities with more volume (i.e., with more patients) will tend to have more reliable scores, while facilities with less volume will tend to have less reliable scores. Therefore, we also use the formula presented by Adams and colleagues (2010) to calculate facility-level reliability as an additional, complementary metric [3, 4]. Specifically for each facility we calculate the reliability as:

$$Reliability = \frac{\sigma_{facility-to-facility}^2}{\sigma_{facility-to-facility}^2 + \frac{\sigma_{facility\ error\ variance}^2}{n}}$$

Where facility-to-facility variance is estimated from the model, n is equal to each facilities observed case size, and the facility error variance is estimated using the variance of the logistic distribution. The facility-level reliability testing methods were also aligned with the specifications used in CMS's PCHQR and OQR programs, with the analysis limited to facilities with at least 25 patients.

References

1. Shrout, P.E., and Fleiss, J.L (1979), "Intraclass correlations: uses in assessing rater reliability," *Psychological Bulletin*, 86, 420-428.
2. Hamer, Robert. (1990) "Compute Six Intraclass correlations." <http://support.sas.com/kb/25/031.html>
3. Adams J, Mehrota, A, Thoman J, McGlynn, E. (2010). Physician cost profiling – reliability and risk of misclassification. *NEJM*, 362(11): 1014-1021.
4. Rousson V, Gasser T, Seifert B. Assessing intrarater, interrater and test–retest reliability of continuous measurements. *Statistics in Medicine* 2002; 21:3431-3446.

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)

Measure Score Reliability

Split-Sample Reliability. There were 1,450 hospitals (11 PCHs and 1,439 non-cancer hospitals) with ≥ 25 patients in their cohorts in each half of the two-year, FY 2015 – 2017 sample. This sample was randomly split and the key characteristics were compared to ensure the patients in each sample were similar, as shown in Table 2, in Section 1.7. For the 11 PCHs, the ICC score was 0.6704 for the Risk-Standardized Admissions Rate (RSAR), and 0.8904 for the Risk-Standardized Emergency Department Visit Rate (RSEDR). Among the 1,099

non-cancer hospitals with at least 50 patients in the combined sample, the ICC score was 0.4314 for the RSAR and 0.3585 for the RSED. R.

Facility-Level Reliability. The PCHs had a median reliability of 0.7848 for the RSAR, and 0.9808 for the RSED. R. All 11 PCHs had 25 or more patients in the FFY 2016 (October 1, 2015 – September 30, 2016) dataset. Among the 1,524 non-cancer hospitals with at least 25 patients, the median reliability was 0.6027 for the RSAR, and 0.7326 for the RSED. R. The reliability estimates for the 25th and 75th percentile denominator values (number of patients) are also shown in Table 5 below.

Table 5. Facility-Level Reliability

	RSAR		RSED. R	
ICC Score for Volume Percentile	PCHs (N=11)	Non-Cancer Hospitals (N=1,524)	PCHs (N=11)	Non-Cancer Hospitals (N=1,524)
25th Percentile	0.7485	0.4369	0.9766	0.5834
Median	0.7848	0.6027	0.9808	0.7326
75th Percentile	0.8922	0.7633	0.9914	0.8534

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?)

Measure Score Reliability Results

The ICC score results from the split-sample or test-retest reliability analysis for PCHs were 0.6704 for the RSAR, and 0.8904 for RSED. R, indicating substantial and almost perfect reliability, respectively. Among the 1,439 non-cancer hospitals with at least 25 patients, the ICC score was 0.4314 for the RSAR and 0.3585 for the RSED. R, reflecting moderate and fair reliability, respectively. The ICC [2,1] is a conservative measure of split-sample/test-retest reliability because it assumes that the multiple measurements are drawn from a larger sample of tests, and that the measured providers are drawn from a larger sample of providers. Given the conservative nature of the ICC [2,1] and the complex constructs of risk-adjusted outcome measures, a lower reliability score is expected.

For the facility-level reliability, the median ICC values for PCHs were 0.7848 for the RSAR, and 0.9808 for the RSED. R, showing substantial and almost perfect agreement, respectively. Among the non-cancer hospitals, the median ICC values were 0.6027 for the RSAR and 0.7326 for the RSED. R, indicating substantial agreement.

Our interpretation of these results is based on the standards established by Landis and Koch (1977) [1]:

< 0 – Less than chance agreement;

0 – 0.2 Slight agreement;

0.21 – 0.39 Fair agreement;

0.4 – 0.59 Moderate agreement;

0.6 – 0.79 Substantial agreement;

0.8 – 0.99 Almost Perfect agreement; and

1 Perfect agreement

The split-sample/test-retest reliability scores among PCHs of 0.6704 for the RSAR, and 0.8904 for RSED. R, and 0.4314 for the RSAR and 0.3585 for the RSED. R among non-cancer hospitals, discussed in the previous section, represent the lower bound of estimate of the true chemotherapy measure reliability. Using the approach used by Adams et al (2010) [2], we obtained median reliability scores for PCHs of 0.7848 for the RSAR, and 0.6027 for the RSAR and 0.7326 for the RSED. R among non-cancer hospitals. This pattern was also observed by Yu,

Mehrotra and Adams (2013) [3]. For example, they found mean reliability for a PCP visits utilization measure to be 0.94 using the approach used by Adams and colleagues (2010), although the test-retest reliability score was 0.68.

Taking together the results above indicate that there is sufficient reliability in the measure score.

References

1. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. *Biometrics*. 1977; 33(1): 159-174.
2. Adams J, Mehrotra, A, Thoman J, McGlynn, E. (2010). Physician cost profiling – reliability and risk of misclassification. *NEJM*, 362(11): 1014-1021.
3. Yu H, Mehrotra A, Adams J. (2013). Reliability of utilization measures for primary care physician profiling. *Healthcare*, Jun; 1(1-2):22-9.

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)

We demonstrated measure validity through the application of established measure development guidelines, and through assessment by external groups. Specifically, our Technical Expert Panel (TEP) and PPS-Exempt Cancer Hospital Measure Development Workgroup (Cancer Workgroup) provided input on the measure's initial development, while subsequent Expert Workgroups (EWGs) convened from 2015 – 2018 advised on revisions to the measure specifications during annual measure reevaluation. In addition, we received measure feedback through a formal public comment process in 2015. Also, the 11 PCHs and all hospitals participating in CMS's OQR program had the opportunity to review draft, non-public results, and provide comments on the measure specifications during August/September 2017 during the measure dry run CMS hosted to educate and receive input from facilities on the measure results and data used for measure calculation. Finally, the measure's face validity was systematically assessed by the 2018 EWG members.

Validity Indicated by Established Measure Development Guidelines:

We developed this measure in consultation with national guidelines for publicly reported outcome measures, with input from outside experts and the public. The measure is consistent with the technical approach to outcomes measurement set forth in NQF guidance for outcome measures [1], CMS Measure Management System (MMS) guidance, and guidance articulated in the American Heart Association scientific statement entitled, "Standards for Statistical Models Used for Public Reporting of Health Outcomes" [2].

Validity as Assessed by External Groups:

Throughout the initial measure development and reevaluation processes, we obtained expert and stakeholder input by holding regular discussions with external clinical consultants, consulting our TEP, PPS-Exempt Cancer Hospital Workgroup, and subsequent EWGs convened from 2015 – 2018 (see below and Measure Submission Form, Section Ad.1. for full membership lists). We also held a 46-day public comment period during measure

development in 2013 and a subsequent, 45-day public comment period during the measure's national Dry Run for the PCHQR and OQR programs (August 15 – September 29, 2017). Additional details about these activities are provided below.

Technical Expert Panel:

TEP. In alignment with the CMS Measures Management System (MMS) Blueprint, we convened a TEP to provide input and feedback during measure development. To convene the TEP, we released a public call for nominations and selected individuals to represent a range of perspectives, including clinicians, patients, and individuals with experience in quality improvement, performance measurement, and healthcare disparities. The TEP had 12 members, including physicians, nurses and patient advocates (see below and the Measure Submission Form, Section Ad.1. for full membership list). We held thirteen structured TEP conference calls consisting of presentation of key issues, our proposed approach, and relevant data, followed by open discussion among TEP members. The TEP's role was to provide advice and feedback through all phases of the initial measure development process, including review and comment on evidence provided in an environmental scan, input to and reviews of measure specifications, and review and guidance relating to public comment on and testing of the measure.

Workgroups:

2014 PPS-Exempt Cancer Hospital Workgroup. The Cancer Workgroup consisted of representatives from each of the 11 PPS-exempt cancer hospitals (see below and the Measure Submission Form, Section Ad.1. for full membership list). The purpose of engaging with the workgroup was to understand quality measurement and improvement activities taking place at the PPS-exempt cancer hospitals and to obtain their perspectives on the importance and usefulness of the measure during its initial development.

2015 – 2018 EWGs

Following the initial measure development phase and during measure reevaluation, we convened a series of EWGs to provide input and feedback on potential revisions to the measure specifications during annual reevaluation. While the membership fluctuated over time, these EWGs generally included 2-3 members from the original TEP and Cancer Workgroups, and then additional experts who were added over time to ensure representation from key stakeholders (see below and the Measure Submission Form, Section Ad.1. for full membership lists). We held 2-3 structured EWG conference calls per year, where we presented key reevaluation issues, discussed our proposed approach and relevant data, and then held open discussion among EWG members.

Public Comment Periods:

2013 Public Comment During Initial Measure Development. During development, we solicited public comment on the measure from June 1 through July 19, 2013 using the standard CMS, MMS Blueprint process. The measure specifications were posted for 45 calendar days to allow time for interested stakeholders to review and comment. 13 measure-specific comments were received, including comments from the American Hospital Association and the Alliance of Dedicated Cancer Centers.

Measures' Application Partnership Review & Public Comment. In addition, in December 2015 and January 2016 as part of the NQF Measure Applications Partnership process, the measure underwent a second public comment period. Throughout the MAP process stakeholders submitted a total of 11 unique comments.

CMS Federal Regulation Public Comment Period. Additionally, as part of CMS' Federal rulemaking process, the measure's final rule language for the Inpatient Prospective Payment System (IPPS) and the Outpatient Prospective Payment System (OPPS) was released for review and public comment. Each program's final rule language was made public for 45 calendar days with the IPPS public comment period occurring May 15 – June 30, 2016 and the OPPS public comment period occurring August 1 – September 15, 2016. CMS received 33 measure specific comments were received during the IPPS public comment period and 75 comments were received during the OPPS public comment period. Commenters included the Alliance of Dedicated Cancer Centers and the American Society of Hematology.

2017 Measure Dry Run and Public Comment. Additionally, during the measure’s national Dry Run CMS held a 45-day public comment period from August 15 through September 29, 2017. During this period, facilities participating in the PCHQR and OQR programs had the opportunity to ask questions about the measure specifications, and their non-public, facility-level results for the FY 2016 data period. We received 216 questions during this period, 3 from PCHs and 213 from non-cancer hospitals.

CMS used the feedback from all of these sources – TEP, Workgroups (Cancer Workgroup, EWGs), public comment periods, and measure Dry Run – to refine the measure specifications during the initial development phase and then during reevaluation. They served as a source of ongoing face validity review on key aspects of the measure, including the codes and logic used to define the cohort, outcomes, exclusions, and risk-adjustment model. In addition, CMS conducted a formal, face validity assessment of the measure with the 2018 EWG, the group most recently convened to provide input on and evaluate the measure, as described below.

Face Validity as Determined by the 2018 EWG:

The 2018 EWG includes an interdisciplinary team of clinicians, medical coders, and measurement experts from cancer and non-cancer hospitals. Three of the EWG members have been involved with the measure since the initial development stages, while the remaining five members were added in 2018 to ensure representation from a diverse set of stakeholders with relevant clinical, coding, and quality measurement expertise. The group met twice in May 2018 to review the current measure specifications and provide input on changes under consideration during the 2018 reevaluation cycle.

The measure score as an indicator of quality was systematically assessed for face validity by confidentially soliciting the EWG members’ agreement with the following statement via an online survey: “The risk-standardized admissions rates and risk standardized emergency department rates obtained from the chemotherapy measure as specified can be used to distinguish between better and worse quality facilities.” The survey offered participants six response options ranging from “strongly disagree” to “strongly agree.” EWG members were asked to complete this survey after reviewing the revised measure specifications and distribution of measure performance among PCHs and non-cancer hospitals in the FFY 2016 dataset, as summarized in Section 2.b.4.

Process Used to Identify International Classification of Diseases, Tenth Revision (ICD-10) Codes

This application includes both the International Classification of Diseases, Tenth Revision (ICD-10) and Ninth Revision (ICD-9) codes needed to identify the measure cohort, exclusions, outcome, and risk factors. A multi-phase process was used to map ICD-9 codes to their ICD-10 equivalents. We:

- Used a conversion tool (CMS’s General Equivalency Mappings [GEMs]) to bidirectionally identify (forward and backward map) ICD-10 equivalents of ICD-9 codes.
- Reviewed the codes to ensure they aligned with the measure’s intent.
- Obtained input from clinical and coding experts on the accuracy and appropriateness of that mapping, and evaluated the impact of ICD-10 conversion on the measure results.
- Posted the mapping in public forum for stakeholders.
- Assessed measure results after coding updates and face validity.

Details of the multi-phase process are described below.

During the first phase, ICD-10 equivalents of all ICD-9 codes were identified by forward and backward mapping using CMS’s GEMs tool during 2015. This mapping was reviewed by project staff with medical coding and clinical expertise, and publically posted on CMS’s Measures Methodology [webpage](#) for review and comment in March 2016 with the 2016 Technical Report, which is included as an appendix to this submission (see “ChemoMeasure_NQF Appendix_SDS). A subsequent forward and backward mapping using CMS’s October 2016 GEMs release was then completed in order to include any new codes introduced for FFY 2017.

In January 2017, as part of the second phase, the complete ICD-9/ICD-10 mapping was reviewed by the members of the 2017 EWG (members listed below). The primary aim of this review was to ensure the mapped ICD-9 and ICD-10 codes aligned with the measure's intent. The code set was refined based on the EWG's feedback, and reviewed by two members of the Yale New Haven Health Services Corporation/Center for Outcomes Research and Evaluation [CORE] Code Set team: Elizabeth Triche, PhD, Division Director and Lead of Code Set Team at CORE; and Danielle Purvis, MPH, Project Coordinator and Coordinator of Code Set Team at CORE. The CORE Code Set Team is responsible for reevaluation of all CORE's outpatient measures, including clinical and technical review of coding updates. This revised code set was used to support measure calculations during CMS's dry run of the measure during September 2017. At this time the ICD-9/ICD-10 mapping was published on the QualityNet website (www.qualitynet.org > Hospitals – Outpatient > Measures > Chemotherapy Measure Dry Run > Measure Methodology or www.qualitynet.org > PPS-Exempt Cancer Hospitals > Measures > Chemotherapy Measure Dry Run > Measure Methodology). Stakeholders were welcome to provide feedback on the mapping and other aspects of the measure specifications during the dry run.

During the third phase, consistency across pre/post conversion time periods was assessed by comparing measure results for the October 2014 – September 2015 (ICD-9) and October 2015 – September 2016 (ICD-10) performance periods, using the ICD-9/ICD-10 mapping from the 2017 measure dry run. The results showed minor increases in the measure denominator and numerator after ICD-10 conversion, but only a small net impact on rates, indicating that conversion did not substantially affect the measure.

Finally, the 2018 EWG also conducted targeted review of the code set to ensure correct ICD-9 to ICD-10 mappings.

- ICD-10 diagnosis and procedure codes used to define the Always Planned Procedures and Diagnoses were identified from the 2015 version of the AHRQ Clinical Classification Software (CCS) categories specified for ICD-10, followed by clinician review. The algorithm also includes some individual ICD-9 codes. To create the crosswalk for the ICD-9-level codes, we used the 2015 ICD-9-CM to ICD-10-CM GEMs tool, made available by CMS, followed by team review.

TEP and EWG Members (also listed in Measure Submission Form, Ad.1) represented a range of perspectives, including physicians and nurses with cancer care and chemotherapy expertise, patient advocates, medical coders, and quality improvement and performance measurement professionals:

2018 EWG members

1. Robert Daly, MD, MBA – Memorial Sloan-Kettering Cancer Center (Staff Physician, Medical Oncology)
2. Stephen Edge, MD* – Roswell Park Memorial Institute (Vice President, Healthcare Outcomes and Policy, Professor of Oncology and Surgical Oncology)
3. Michael Hassett, MD, MPH* – Dana Farber Cancer Center (Attending Physician, Medical Oncology; Assistant Professor, Medicine, Harvard Medical School)
4. Scott Huntington, MD, MPH – Yale New Haven Hospital (Attending Physician, Hematology)
5. Denise Morse, MBA – City of Hope Cancer Treatment and Research Center (Senior Manager, Quality Analytics)
6. Joseph Ross, MD, MHS – Yale University School of Medicine (Associate Professor of General Medicine and of Public Health)
7. Weijing Sun, MD – University of Kansas Cancer Center (Director of Medical Oncology and Associate Director of University of Kansas Cancer Center)
8. Allison Snyderman, Ph.D.* - Memorial Sloan-Kettering Cancer Center (Outcomes Researcher)

*Also served as member of TEP and 2014 PPS-Exempt Cancer Workgroup

2017 EWG members

1. Susan Armstrong—City of Hope Cancer Treatment and Research Center (Senior Manager, Coding and Data Quality)
2. Arnold Chen, MD, MPH – Mathematica Policy Research (Clinician, Senior Researcher)
3. Michael J. Hassett, MD, MPH—Dana-Farber Cancer Institute, Outpatient Clinic and Treatment Center (Staff Physician, Breast Oncology)
4. Joseph Ross, MD, MHS – Yale University School of Medicine (Associate Professor of General Medicine and of Public Health)
5. Tracy Spinks, CPH—MD Anderson Hospital, Cancer Care Delivery (Program Director)
6. Denise Stone, RN, MBA – Mathematica Policy Research (Clinician, Lead Program Analyst)

2016 EWG members

1. Peter B. Bach, MD, MAPP—Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Director)
2. Stephen Edge, MD—Baptist Cancer Center (Cancer Center Director)
3. Michael J. Hassett, MD, MPH—Dana-Farber Cancer Institute, Outpatient Clinic and Treatment Center (Staff Physician, Breast Oncology)
4. Karl Lorenz, MD, MSHS—UCLA, Department of Veterans Affairs (Associate Professor); Rand Health
5. Allison Snyderman, PhD —Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Researcher)
6. Tracy Spinks, CPH—MD Anderson Hospital, Cancer Care Delivery (Program Director)

2015 EWG members

1. Peter B. Bach, MD, MAPP—Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Director)
2. Stephen Edge, MD—Baptist Cancer Center (Cancer Center Director)
3. Michael J. Hassett, MD, MPH—Dana-Farber Cancer Institute, Outpatient Clinic and Treatment Center (Staff Physician, Breast Oncology)
4. Karl Lorenz, MD, MSHS—UCLA, Department of Veterans Affairs (Associate Professor); Rand Health
5. Allison Snyderman, PhD —Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Researcher)
6. Tracy Spinks, CPH—MD Anderson Hospital, Cancer Care Delivery (Program Director)

PPS-Exempt Cancer hospital workgroup members

1. J. Robert Beck, MD—American Oncologic Hospital (Fox Chase) (Senior Vice President and Chief Academic Officer)
2. Joe Jacobson, MD—Dana Farber Cancer Institute (Chief Quality Officer)
3. Barbara Jagels, MHA, RN, OCN—Seattle Cancer Care Alliance (Fred Hutchinson Cancer Research Center) (Director of Nursing and Clinical Excellence)
4. Dana Jenkins—Roswell Park Memorial Institute (Vice President of Organizational Improvement)
5. Tricia Kassab, RN, MS, CPHQ, HACCP—City of Hope National Medical Center (Vice President of Quality and Patient Safety)
6. Jeremy Miransky, PhD—Memorial Hospital for Cancer and Allied Disease (MSKCC) (Quality Analytics Manager)
7. Shyroll Morris, MBA, MPH—University of Miami Hospital and Clinics

8. Thomas Ross, MS—H. Lee Moffitt Cancer and Research Institute Hospital, Inc. (Director of Quality and Safety)
9. Anthony Senagore, MD—University of Southern California Kenneth Norris Jr. Cancer Hospital (Chief of Colorectal Surgery)
10. Ron Walters, MD, MHA, MBA—The University of Texas MD Anderson Cancer Center (Associate Vice President of Medical Operations and Informatics)
11. Saul Weingart, MD, PhD—Dana Farber Cancer Institute (Vice President for Quality Improvement and Patient Safety)

TEP members

1. Peter B. Bach, MD, MAPP—Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Director)
2. Stephen Edge, MD—Baptist Cancer Center (Cancer Center Director)
3. Andrew Glass, MD—Kaiser Permanente Northwest, Center for Health Research (Senior Investigator)
4. Mark Gorman—Independent Consultant (Patient Advocate)
5. Michael J. Hassett, MD, MPH—Dana-Farber Cancer Institute, Outpatient Clinic and Treatment Center (Staff Physician, Breast Oncology)
6. Karl Lorenz, MD, MSHS UCLA, Department of Veterans Affairs (Associate Professor); Rand Health
7. Joan McClure, MS—National Comprehensive Cancer Network, Clinical Information and Publications (Senior Vice President)
8. Bruce Minsky, MD—MD Anderson Hospital, Department of Radiation Oncology (Professor and Director of Clinical Research)
9. Shirley Stagner, MSN, ONP, AOCNP—Lawrence Hospital Center, Cancer Survivorship Program (Nurse Practitioner)
10. Janet H. Van Cleave, PhD, MSN, AOCNP—New York University College of Nursing (Assistant Professor)
11. Sandra L. Wong, MD MS—University of Michigan Health System, Division of Surgical Oncology (Physician)

References

1. National Quality Forum. National voluntary consensus standards for patient outcomes, first report for phases 1 and 2: A consensus report. Available at: http://www.qualityforum.org/projects/Patient_Outcome_Measures_Phases1-2.aspx. Accessed June 7, 2017.
 2. Krumholz HM, Brindis RG, Brush JE, et al. Standards for statistical models used for public reporting of health outcomes: An American Heart Association scientific statement from the Quality of Care and Outcomes Research Interdisciplinary Writing Group: cosponsored by the Council on Epidemiology and Prevention and the Stroke Council endorsed by the American College of Cardiology Foundation. *Circulation*. 2006; 113(3):456-462.
- 2b1.3. What were the statistical results from validity testing? (e.g., correlation; t-test)**

Face Validity as Assessed by the 2018 EWG:

All 8 EWG members completed the face validity survey. All 8 respondents (100%) felt the measure had face validity, indicating that they strongly, moderately, or somewhat agreed with the following statement: “The risk-standardized admissions rates and risk-standardized emergency department rates obtained from the chemotherapy measure as specified can be used to distinguish between better and worse quality facilities.” Specifically, 2 members (25%) strongly agreed, 5 members (62.5%) moderately agreed, and 1 member (12.5%) somewhat agreed.

Two members also submitted comments as part of this process. One member noted that it is important to ensure results for cancer hospitals are not compared directly to those for non-cancer hospitals, since separate risk models are used for these two populations. The other comment signaled support for the measure methodology, noting that the measure methods are comprehensive.

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?*)

There was perfect agreement (100%) among EWG members that the measure has face validity, i.e., that it measures what it is intended to measure – the quality of care provided to cancer patients receiving outpatient chemotherapy treatment. The EWG members were actively involved with the measure’s reevaluation during the spring of 2018, and reviewed both the current measure specifications and distribution of performance prior to assessing face validity. We conclude the measure’s validity to be at least moderate based on their assessment.

2b2. EXCLUSIONS ANALYSIS

NA ☐ **no exclusions** — *skip to section [2b3](#)*

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*)

We evaluated five possible exclusions throughout the initial development and reevaluation of the measure:

- First, we considered the removal of patients with leukemia because of the high toxicity of treatment and expected readmissions due to relapse.
- Second, we considered the removal of patients who do not have a full year of prior enrollment data to ensure complete data for the risk-adjustment model.
- Third, we considered removal of patients who do not have at least one chemotherapy treatment followed by 30 days of enrollment for full data availability to identify outcomes.
- Fourth, we considered the removal of patients younger than 65 years of age because patients aged 18-64 enrolled in Medicare may be systematically different than those patients 65 and older.
- Finally, we considered the removal of patients who were receiving chemotherapy and had a cancer diagnosis during the performance period, but did not have a cancer diagnosis on the index chemotherapy claim and did have a diagnosis for an auto-immune condition. Based on stakeholder feedback, these patients are assumed to be receiving chemotherapy to treat an auto-immune condition rather than to treat cancer.

We reviewed each of these exclusions with our TEP and/or EWGs. TEP and EWG members raised concerns about the exclusion of patients aged 18-64, expressing a desire for a broad cohort and indicating that there was no clinical reason to exclude this group. We therefore explored the appropriateness of including these patients by (1) reviewing patient characteristics separately for these two subsets, (2) reviewing the observed performance rates for the two separate subsets, and (3) fitting the risk-adjustment model separately for these two subsets.

Expert input indicated that the remaining four exclusions were clinically appropriate or required for data completeness (see Notice of Intent to Submit or Measure Submission Form, Section S.9 for more information):

- (1) patients with a diagnosis of leukemia at any time during the performance period,
- (2) patients who were not enrolled in Medicare FFS Parts A and B in the year before their first outpatient chemotherapy treatment during the performance period,

- (3) patients who do not have at least one outpatient chemotherapy treatment followed by continuous enrollment in Medicare FFS Parts A and B in the 30 days after the procedure, and
- (4) cases in which patients receive chemotherapy to treat a qualifying autoimmune condition, rather than to treat cancer. Such cases have a qualifying chemotherapy code, and an autoimmune diagnosis, but no cancer diagnosis. Note that this is a case-level exclusion; as long as the patient has additional cases that meet inclusion criteria, they will remain in the cohort.

We examined overall frequencies and proportions of the total cohort excluded for each exclusion criterion. We then looked at the distribution of the exclusions across hospitals. Lastly, we calculated the observed performance rate with and without accounting for exclusions. The results are presented below.

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)*

We explored the potential bias of including patients aged 18-64 in the measure using the 2012–2013 dataset during the measure’s initial development. Specifically, we compared differences in the 18-64 and 65 and older populations by comparing: (1) patient characteristics, (2) observed inpatient admission and ED visit rates, and (3) risk-adjustment model fit statistics for the two populations. We found that patients aged 18-64 represented 13% of the final measure cohort, and while the younger population has higher observed outcome rates, the risk-adjustment model parameter estimates were similar for both age groups. Based on these findings, as well as the recommendation of our TEP, we determined there was not a strong statistical or clinical reason to exclude the younger patients from the measure cohort; all adult patients 18 years and older remain in the eligible cohort.

Applying our measure inclusion criteria (all adult Medicare FFS patients with a diagnosis of cancer aged 18 years or older at the start of the performance period who received a qualifying chemotherapy procedure) to the Medicare FFS FY 2016 Dataset resulted in an initial cohort of 354,849 unique patients overall, with 30,006 patients from PCHs and 324,843 from non-cancer hospitals. We then applied the remaining four exclusion criteria (see the Intent to Submit Notice and Measure Submission Form, Sections S.9, for exclusion rationale):

- (1) patients with a diagnosis of leukemia at any time during the performance period,
- (2) patients who were not enrolled in Medicare FFS Parts A and B in the year before their first outpatient chemotherapy treatment during the performance period,
- (3) patients who do not have at least one outpatient chemotherapy treatment followed by continuous enrollment in Medicare FFS Parts A and B in the 30 days after the procedure, and
- (4) cases in which patients receive chemotherapy to treat a qualifying autoimmune condition, rather than to treat cancer. Such cases have a qualifying chemotherapy code, and an autoimmune diagnosis, but no cancer diagnosis. Note that this is a case-level exclusion; as long as the patient has additional cases that meet inclusion criteria, they will remain in the cohort.

This resulted in excluding 65,306 (18.4%) of patients eligible for the cohort from the measure in the FY2016 dataset overall, as shown in Table 6. Among the excluded patients, 6,529 were from PCHs and 58,777 were from non-cancer hospitals, representing 21.8% and 18.1% of their respective measure-eligible cohorts. Thus, the final Medicare FFS FY 2016 Dataset included 289,543 unique patients overall, with 23,477 at PCHs and 266,066 at non-cancer hospitals.

Table 6. Count and Percent of Excluded Patients

	Overall		PCHs		Non-Cancer Hospitals	
Exclusion	n	%	n	%	n	%
(1) Leukemia	24,924	7.0	2,510	8.4	22,414	6.9

(2) No Medicare FFS A/B Enrollment 12 Months Prior to First Index Chemo	41,841	11.8	4,298	14.3	37,543	11.6
(3) No Medicare FFS A/B Enrollment 30 Days Following Index Chemo	4,326	1.2	245	0.8	4,081	1.3
(4) Receiving Chemotherapy for Autoimmune Condition	337	0.1	4	0.0	333	0.1
All Exclusions*	65,306	18.4	6,529	21.8	58,777	18.1

*Note: Patients are eligible for more than one exclusion, therefore the count of all exclusions is lower than the sum of the individual exclusions.

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)

After extensive literature review, examination of existing measures, consideration of requirements to have adequate risk adjustment and identification of admissions or ED visits, and discussion with the TEP and EWGs, we determined these four exclusion criteria are necessary for a valid measure. The goal was to be as inclusive as possible while creating a clinically coherent cohort.

Testing of the distribution of exclusion criteria across hospitals suggests modest variation among providers. The uneven distribution of excluded populations and procedures supports our decision that these exclusions are required. Failure to exclude these populations may distort the measure score and unfairly disadvantage certain hospitals. Additional rationales for exclusions are detailed in the Intent to Submit Notice and Measure Submission Form, Section S.9. After exclusions were applied, the measure captured the majority (81.6%) of all qualifying patients. The exclusions are very narrowly targeted and necessary to ensure a clinically coherent measure cohort and a cohort with complete data available for risk adjustment and identification of admissions or ED visit outcomes.

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 2b4.

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

☐ No risk adjustment or stratification

☒ Statistical risk model with [21 risk factors for the inpatient admission outcome model and 16 risk factors for the ED visit outcome](#)

☐ 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.

The measure has two mutually exclusive outcomes: (1) patients in the cohort admitted to any acute-care hospital with one of the following diagnoses—anemia, dehydration, diarrhea, emesis, fever, nausea, neutropenia, pain, pneumonia, or sepsis—within 30 days of an outpatient chemotherapy administration at the reporting hospital, and (2) patients in the cohort that did not have a qualifying inpatient admission, but were seen at any ED with one of the qualifying diagnoses within 30 days of an outpatient chemotherapy administration at the reporting hospital. As a result, we developed two risk-adjustment models, one for each dependent variable—inpatient admissions and ED visits.

The measure uses a two-level hierarchical logistic regression model to estimate facility-level risk-standardized admission rates (RSARs) and risk-standardized emergency department rates (RSEDRs). This approach accounts for the clustering of patients within facilities and variation in sample size across facilities.

The measure has four risk-adjustment models, one for each outcome, reported separately for PCHs and non-cancer hospitals (two outcomes each reported for two facility types). The risk-adjustment model for inpatient admissions has 21 variables (age, sex, chemotherapy exposure, concurrent radiotherapy exposure, 9 comorbidity variables, and 8 cancer diagnosis categories), and the risk-adjustment model for ED visits has 16 variables (age, sex, chemotherapy exposure, concurrent radiotherapy exposure, 6 comorbidity variables, and 6 cancer diagnosis categories). The ED visit model does not include the variables for renal disease, diabetes, metabolic disorder, lymphoma, or prostate cancer that the inpatient admission model includes because these variables were not predictive of risk for the outcome in the ED setting. The same risk factors are included for both PCHs and non-cancer hospital models, but the coefficients vary according to differences in the underlying patient populations for these two facility types.

Risk variable definitions:

Chemotherapy exposure is defined as the number of chemotherapy treatments in the performance period for a patient a given provider. Exposure to concurrent radiotherapy assesses whether the first index outpatient chemotherapy case – which is the case included in the measure denominator – was accompanied by concurrent radiotherapy. Chemotherapy treatments are defined in the attached 2018 Chemotherapy Measure_Data Dictionary on sheets using the “S.9 Denominator-Chemo Procedure,” “S.9 Denominator – Chemo Encounter,” and “S.9 Denominator – Chemo Medication” codes.

Concurrent radiotherapy is defined as having a radiotherapy procedure present on the same claim as the first index chemotherapy case or on a separate claim within 14 days prior to the first index chemotherapy case. Individual (ICD-10) diagnosis codes, procedure codes, HCPCS codes and CPT codes are used to identify chemotherapy and radiotherapy exposure. Concurrent Radiotherapy is defined in the attached 2018 Chemotherapy Measure_Data Dictionary on sheets using the “S.15 Risk Factor-Radiotherapy” codes.

We define the comorbidity variables in the models using the CMS Condition Categories (CCs), which are clinically meaningful groupings of more than 69,000 ICD-10 diagnosis codes. In consultation with our TEP, the CCs selected for inclusion were bundled with other clinically related CCs for empirical assessment of significance within the model. The result was nine bundled CCs—diabetes, metabolic disorders, gastrointestinal (GI) disorders, psychiatric disorders, neurological conditions, cardiovascular disease, respiratory disorders, renal disease, and other injuries.

The cancer types included in the model are defined using ICD-10 diagnosis codes (see attached 2018 Chemotherapy Measure_Data Dictionary on sheets using the “S.15 Risk Factor” codes). Based on input from our TEP, these were aggregated into nine clinically related and decently sized groupings—breast cancer, digestive cancer, genitourinary cancer, respiratory cancer, lymphoma, prostate cancer, secondary cancer of the lymph nodes, secondary cancer of solid tumor, and other cancers.

Model Variables – inpatient admissions

1. Age (continuous)
2. Sex (male)
3. Exposure (Number of hospital OPD chemotherapy treatments during period)
4. Respiratory disorders (CC 110 – 113)
5. Renal disease (CC 132, 134 – 140)
6. Diabetes (CC 17 – 20)
7. Other injuries (CC 174)
8. Metabolic disorder (CC 21-26)
9. Gastrointestinal disorder (CC 27-32; 34; 36-38)

10. Psychiatric disorder (CC 50-69)
11. Neurological conditions (CC 70-81)
12. Cardiovascular disease (CC 82-109)
13. Breast cancer
14. Digestive cancer
15. Respiratory cancer
16. Lymphoma
17. Other cancer
18. Prostate cancer
19. Secondary – lymph
20. Secondary – solid
21. Concurrent Radiotherapy

Model Variables – ED visits

1. Age (continuous)
2. Sex (male)
3. Exposure (Number of hospital OPD chemotherapy treatments during period)
4. Respiratory disorders (CC 110 – 113)
5. Other injuries (CC 174)
6. Gastrointestinal disorder (CC 27-32; 34; 36-38)
7. Psychiatric disorder (CC 50-69)
8. Neurological conditions (CC 70-81)
9. Cardiovascular disease (CC 82-109)
10. Breast cancer
11. Digestive cancer
12. Respiratory cancer
13. Other cancer
14. Secondary – lymph
15. Secondary – solid
16. Concurrent Radiotherapy

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.

Not applicable – this measure is risk-adjusted.

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?

Our approach to risk adjustment was tailored to, and appropriate for, a publicly reported outcome measure as articulated in published scientific guidelines [1,2]. In this section we detail both the initial development of the

risk-adjustment models, and then their subsequent refinement during reevaluation to include concurrent radiotherapy as a risk variable.

Initial Model Development

We detail the sequential method for selecting patient-level risk factors below.

Candidate Risk-Adjustment Variables

Candidate risk-adjustment variables were patient-level risk adjustors that are expected to be predictive of the outcomes based on prior literature, clinical judgment, and empirical analysis. We limited our initial selection of candidate variables for inclusion in our preliminary risk-adjustment model to variables with a strong clinical rationale for inclusion as identified in the literature and through clinical expert input. Identification of these variables is described below. See attached Data Dictionary, sheet “2b4.3 Candidate Variables” for the list of candidate variables.

Demographic variables: In alignment with the specifications of other NQF endorsed claims-based outcome measures, as well as the NQF guidelines at the time of development, we included age and sex as candidate covariates. [Note: Due to changes in NQF policy, additional social risk factors were considered during continued assessment of this measure as described later in this section.]

Comorbidities: The model adjusts for case mix differences based on the comorbidities of the patient at the time of the first outpatient chemotherapy treatment during the performance period. During model development, we defined comorbidities using Condition Categories (CCs) from Version 12 (V12) of the CMS-HCC risk-adjustment model, which are clinically meaningful groupings of more than 15,000 ICD-9-CM diagnosis codes. With a subset of our TEP, we reviewed all 189 CCs to determine the clinical appropriateness and prevalence within the cohort for potential inclusion in the model. (Note: During subsequent years, these CCs were updated to Version 22 (V22) of the CMS-HCC model as part of measure reevaluation.)

Specific considerations included the number of patients in our cohort potentially affected, whether the condition affects admission for one of the ten outcome-qualifying diagnoses, and whether inclusion of the condition in the model would incentivize appropriate treatment, even when that variable is theoretically unrelated to admission for one of the identified reasons. For example, patients with diabetes may have gastric paresis, a condition that slows emptying of the stomach and increases the likelihood of nausea. The CCs selected for inclusion were bundled with other clinically related CCs for empirical assessment of significance within the model. The result was nine bundled CCs—diabetes, metabolic disorders, gastrointestinal (GI) disorders, psychiatric disorders, neurological conditions, cardiovascular disease, respiratory disorders, renal disease, and other injuries.

Indicator of disease severity: We explored cancer type as an indicator of disease severity available in claims data by assessing the distribution of patients across a granular level of cancer diagnoses. In conjunction with a subset of our TEP, we aggregated these granular cancer types into nine clinically related and decently sized groupings—breast cancer, digestive cancer, genitourinary cancer, respiratory cancer, lymphoma, prostate cancer, secondary cancer of the lymph nodes, secondary cancer of solid tumor, and other cancers.

Exposure: We also assessed the number of chemotherapy treatments during the performance period (that is, exposure). The exposure variable is necessary because the measure estimates the risk-adjustment models at the patient level and the number of outpatient chemotherapy treatments varies by patient. Patients with more treatments during the period have an increased probability of experiencing an outcome because the algorithm looks for an outcome after each treatment. The exposure variable is the count of outpatient chemotherapy administrations the patient experienced at the attributed hospital during the performance period.

Interactions: Through discussion with our 2015 Expert Working Group (see Section 2b.1.2, above, and Measure Submission Form, Section Ad.1. for full membership list), we determined the most clinically relevant interactions are likely to be between the age variable and the different cancer types. Based on this input, we tested age-cancer type interaction terms as candidate covariates.

Variable Selection

To select the final variables to include in the risk-adjustment model, we fitted a logistic regression model to predict the outcome with the candidate variable set. To develop a parsimonious model, we then removed non-significant variables from the initial model using a stepwise purposeful selection method described by Hosmer and Lemeshow [3,4]. Our goal was to minimize the number of variables in the model while preserving model performance (as measured by the c-statistic). During this process, for each of the two models, the least significant variable in the model was removed one at a time until only statistically significant ($p < 0.05$, assessed using a likelihood ratio test) variables remained in the model. Interaction terms between age and cancer type were tested and were only retained in the model if significant at a level of $p < 0.01$. The higher threshold for statistical significance of interaction terms was to ensure that only interactions that have a higher likelihood of being true interactions were included.

The attached Data Dictionary, sheet “S.15 Risk Model Coefficients” indicates the final risk variables selected, the codes used to define the risk variables for our statistical model, and their odds ratios and 95% CIs.

Social Risk Factors Conceptual Model

Following the selection process for clinically-relevant risk factors described above, we undertook an assessment of the potential need to incorporate additional social risk factors into our risk-adjustment model. In this section, we describe the conceptual model that guided our work. In Section 1.8, we described the three variables evaluated in our analysis (race, Medicaid dual eligible status, and neighborhood SES factors composited into the AHRQ SES Composite Index).

The potential causal pathways by which social risk factors influence the risk of admission or ED visit following outpatient chemotherapy are varied and complex. The presence of disparities in chemotherapy outcomes are due to multiple complementary causes. To help inform our conceptualization of the pathways by which social risk factors affect admissions and ED visits for patient receiving chemotherapy treatment in a hospital outpatient setting, we performed a literature search. The studies indicated that individuals that identify as a racial minority, with low socioeconomic status (SES), with charity care or self-pay insurance, are women, or are unmarried were more likely to experience a gap in cancer care in the outpatient chemotherapy setting than their counterparts. Please refer to question 1 of the “ChemoMeasure_NQF Appendix_SDS” for more information on the literature review.

The following highlights possible social risk-related conceptual pathways that are important to consider:

1. Relationship of social risk with health. People who face sociodemographic disadvantages usually have worse health status, which in turn leads to worse health outcomes compared to people who do not experience these disadvantages. This means that chemotherapy patients who have lower health literacy, income, education, and no insurance might experience a higher symptom burden or have greater disease severity, and in turn have more ED visits and hospital admissions due to having worse health status in general. This pathway could be accounted for within the existing clinical risk-adjustment variables in the current model.
2. Access to care. Limited access to health care may prevent individuals from early detection of cancer, making them more likely to be diagnosed with late-stage cancer that could have been treated more effectively or cured if diagnosed earlier [5]. Worse access to care also impacts patients’ ability to contact their physicians when they are experiencing cancer-related symptoms or adverse effects from treatment, which may make them more likely to experience ED visits, hospital admissions, ambulance use, and hospital mortalities compared to cancer patients that are diagnosed at an earlier stage [6].
3. Differential care across hospitals. Cancer patients at minority-serving hospitals are less likely to receive adequate pain treatment [7]. Poor and minority patients are also more likely to be seen in safety-net hospitals and these hospitals may lack the financial resources to make certain services available, such as specialized palliative care teams, making these patients more likely to require acute care, such as an ED visit or hospital admission, for symptom management.

The combination of treatment disparities, increase symptom occurrence and severity, and inadequate pain management may place minority cancer patients at greater risk of experiencing a gap in outpatient chemotherapy care, which may increase the likelihood of ED visits and hospital admissions.

2017/2018 Model Reevaluation

Updates to Comorbidities: In 2017 we updated the CCs used to identify the comorbidities in the model to Version 22 (V22) of the CMS-HCC model. There were no changes made to comorbidities used in the model after EWG members confirmed clinical appropriateness and prevalence within the cohort for existing comorbidities in the model.

Updates to Exposure: Following stakeholder feedback that administration of concurrent radiotherapy increases the likelihood of adverse events among patients receiving chemotherapy, we solicited feedback from our 2018 EWG members regarding the clinical appropriateness of this as either a new measure exclusion, or a risk factor. The group ultimately advised us to include this as a variable in the risk-adjustment models as, in the group's opinion, this approach will ultimately incentivize better coordination and management of these cases. The group recommended that concurrent radiotherapy be defined as receipt of radiotherapy on the date of chemotherapy or up to 14 days before administration of chemotherapy. In order to capture cases that meet these criteria we identified qualifying radiation therapy procedure codes using ICD-10 procedure codes, Current Procedure Terminology codes, and Healthcare Common Procedure Coding System codes. This increased the number of risk factors in the inpatient admissions model to 21, and in the ED visits model to 16. The measure score and testing results in the NQF application reflect these updates to the measure's risk models.

References

1. Krumholz HM, Brindis RG, Brush JE, et al. Standards for Statistical Models Used for Public Reporting of Health Outcomes: An American Heart Association Scientific Statement From the Quality of Care and Outcomes Research Interdisciplinary Writing Group: Cosponsored by the Council on Epidemiology and Prevention and the Stroke Council Endorsed by the American College of Cardiology Foundation. *Circulation*. 113 (3): 456-462.
2. Normand S-LT, Shahian DM. Statistical and Clinical Aspects of Hospital Outcomes Profiling. *Stat Sci*. 2007; 22 (2): 206-226
3. "Disparities in Cancer Care." *Journal of Oncology Practice / American Society of Clinical Oncology*, vol. 2, no. 5, 2006, pp. 234-239.
4. Hosmer DW LS. *Applied Survival Analysis: Regression Modeling of Time to Event Data*. New York: Wiley; 1999.
5. National Cancer Institute. "Cancer Health Disparities." Available at: [<http://www.cancer.gov/about-nci/organization/crhd/cancer-health-disparities-fact-sheet>]. Accessed on January 19, 2016.
6. Kotajima, F., K. Kobayashi, H. Sakaguchi, and M. Nemoto. "Lung Cancer Patients Frequently Visit the Emergency Room for Cancer-Related and -Unrelated Issues." *Molecular and Clinical Oncology*, vol. 2, no. 2, 2014, pp. 322-326.
7. Fisch, M.J., J.W. Lee, M. Weiss, L.I. Wagner, V.T. Chang, D. Cella, J.B. Manola, L.M. Minasian, W. McCaskill-Stevens, T.R. Mendoza, and C.S. Cleeland. "Prospective, Observational Study of Pain and Analgesic Prescribing in Medical Oncology Outpatients with Breast, Colorectal, Lung, or Prostate Cancer." *Journal of Clinical Oncology*, vol. 30, no. 16, 2012, pp. 1980–1988.

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?

The final results of our initial model development work resulted in 20 variables in the inpatient admission outcome model and 15 variables in the ED outcome model with p values <0.05. The (age x cancer type) interaction terms were retained if p for interaction was <0.01. For the inpatient admission outcome model, only the interaction of (age x digestive cancer) was significant (p-value for interaction <0.001). However, due to the minimal improvement in model fit [AIC (76245 -> 76238) and c-statistic (0.725 -> 0.725)] and our desire to create the most parsimonious model, we did not include any interaction terms in our final model. No interaction terms met this criterion for the ED visit outcome model.

In addition, the final model did not include SDS variables. See Section 2b4.4b for more information.

The following variables were selected as the final risk-adjustment variables for the inpatient admission outcome model. In addition, Table 3 of the attached Measure Technical Report includes the coefficients, odds ratios, and confidence intervals for each variable included in the initial model, before the addition of concurrent radiotherapy risk variable in 2017-2018.

1. Age (continuous)
2. Sex (male)
3. Exposure (Number of hospital OPD chemotherapy treatments during period)
4. Respiratory disorders (CC 110 – 113)
5. Renal disease (CC 132, 134 – 140)
6. Diabetes (CC 17 – 20)
7. Other injuries (CC 174)
8. Metabolic disorder (CC 21-26)
9. Gastrointestinal disorder (CC 27-32; 34; 36-38)
10. Psychiatric disorder (CC 50-69)
11. Neurological conditions (CC 70-81)
12. Cardiovascular disease (CC 82-109)
13. Breast cancer
14. Digestive cancer
15. Respiratory cancer
16. Lymphoma
17. Other cancer
18. Prostate cancer
19. Secondary – lymph
20. Secondary – solid

The following variables were selected as the final risk-adjustment variables for the ED visit outcome model. In addition, Table 4 of the attached Measure Technical Report includes the coefficients, odds ratios, and confidence intervals for each variable included in the initial model, before the addition of concurrent radiotherapy risk variable in 2017-2018.

1. Age (continuous)
2. Sex (male)
3. Exposure (Number of hospital OPD chemotherapy treatments during period)
4. Respiratory disorders (CC 110 – 113)
5. Other injuries (CC 174)
6. Gastrointestinal disorder (CC 27-32; 34; 36-38)

7. Psychiatric disorder (CC 50-69)
8. Neurological conditions (CC 70-81)
9. Cardiovascular disease (CC 82-109)
10. Breast cancer
11. Digestive cancer
12. Respiratory cancer
13. Other cancer
14. Secondary – lymph
15. Secondary – solid

In 2018 we evaluated the impact of adding concurrent radiotherapy into the two existing models, based on stakeholder feedback and clinical input from our 2018 EWG members. We found that concurrent radiotherapy was significant at $p < 0.05$ in all four models (both outcomes, for each facility type) and did not markedly change the coefficients or significance of other included variables. In addition, the model c-statistics remained strong, and were 0.6933 for the RSAR and 0.6470 for the RSEDR at PCHs, and 0.7114 for the RSAR and 0.6504 for the RSEDR at non-cancer hospitals. As a result, we added this risk factor to our models, resulting in 21 risk factors for the admissions model and 16 risk factors for the ED visits model. With this revision, the list of variables included in the final risk-adjustment models are:

Model Variables – inpatient admissions

1. Age (continuous)
2. Sex (male)
3. Exposure (Number of hospital OPD chemotherapy treatments during period)
4. Respiratory disorders (CC 110 – 113)
5. Renal disease (CC 132, 134 – 140)
6. Diabetes (CC 17 – 20)
7. Other injuries (CC 174)
8. Metabolic disorder (CC 21-26)
9. Gastrointestinal disorder (CC 27-32; 34; 36-38)
10. Psychiatric disorder (CC 50-69)
11. Neurological conditions (CC 70-81)
12. Cardiovascular disease (CC 82-109)
13. Breast cancer
14. Digestive cancer
15. Respiratory cancer
16. Lymphoma
17. Other cancer
18. Prostate cancer
19. Secondary – lymph
20. Secondary – solid
21. Concurrent Radiotherapy

Model Variables – ED visits

1. Age (continuous)
2. Sex (male)

3. Exposure (Number of hospital OPD chemotherapy treatments during period)
4. Respiratory disorders (CC 110 – 113)
5. Other injuries (CC 174)
6. Gastrointestinal disorder (CC 27-32; 34; 36-38)
7. Psychiatric disorder (CC 50-69)
8. Neurological conditions (CC 70-81)
9. Cardiovascular disease (CC 82-109)
10. Breast cancer
11. Digestive cancer
12. Respiratory cancer
13. Other cancer
14. Secondary – lymph
15. Secondary – solid
16. Concurrent Radiotherapy

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.**

This section includes results from analyses conducted during initial model development and more recently during 2018 re-evaluation activities.

Initial Model Development. As described in Section 1.8, during the initial model development process we considered three variables in our social risk analysis: (1) race (black, other), (2) Medicaid dual eligible status, and (3) neighborhood SES factors composited into the AHRQ SES Composite Index. We conducted several analyses, presented below, including (1) variation in patient SDS risk factors across hospitals; (2) the association between social risk factor variables and the outcomes; (3) the impact of including social risk factor variables as part of risk-adjustment on model performance; and (4) the impact of including social risk factor variables as part of risk-adjustment on hospital rankings. Key findings and our conclusion are described below. A complete summary of our initial social risk factor assessment, including analysis tables, can be found on page 31 in the measure technical report: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/HospitalQualityInits/Downloads/Admissions-and-Emergency-Department-Visits-for-Patients-Receiving-Outpatient-Chemotherapy-Measure-Technical-Report.zip>

Variation in Prevalence of Social Risk Factors across Hospitals

There is substantial variation in the prevalence of black, Medicaid dual-eligible, and low SES patients (scores below 43.27 on the AHRQ SES Composite Index) in the measure cohort across hospitals. For the measure, the percentage of patients who are black ranges from 0% to 100% across hospitals, with a median of 0.7% (interquartile range [IQR] 0%-10.6%). The percentage of patients who are Medicaid dual eligible ranges from 0% to 100% across hospitals, with a median of 18.1% (IQR 9.0% - 30.7%). The percentage of patients with low SES ranges from 0% to 100% across hospitals, with a median of 19.0% (IQR 2.2% - 52.5%).

Association between SDS variables and the outcomes

At the patient-level, our analysis shows that “high social risk” patients (as characterized by three individual indicators: Medicaid dual-eligibility, race as black, and low SES) receiving hospital-based outpatient chemotherapy are more likely to have an inpatient admission and emergency department (ED) visit within 30 days than “low social risk” patients.

- Dual eligible patients were more likely to have an inpatient admission or ED visit than non-dual eligible patients (13.7 percent of dual eligible vs 9.7 percent of non-dual eligible for inpatient admission, and 6.2 percent of dual eligible vs 3.8 percent of non-dual eligible for ED visits);
- Black patients were more likely to have an inpatient admission or ED visit than non-black patients (12.9 percent of black patients vs 10.0 percent of non-black for inpatient admission, and 5.5 percent of black patients vs 4.0 percent of non-black for ED visits); and
- Patients with low SES were more likely to have an inpatient admission or ED visit than patients with higher SES (11.5 percent of patients with low SES vs 9.4 percent of patients with high SES for inpatient admission, and 4.8 percent of patients with low SES vs 3.6 percent of patients with high SES for ED visits).

At the *hospital-level*, no between-hospital effects were observed for hospital case-mix by Medicaid dual-eligibility, race, or the AHRQ SES Composite Index. Specifically, there was no clear relationship between the median risk-standardized rates and hospitals' case mix by these three social risk factors. In addition, the distributions of risk-standardized rates overlapped significantly across hospitals grouping by these three social risk factors, suggesting that hospitals caring for a greater percentage of high social risk patients have similar rates of inpatient admission and ED visits within 30 days of hospital-based outpatient chemotherapy. For example, the hospitals in the lowest quartile of proportion of black patients had a median risk-adjusted admission rate of 10.2, the second quartile had a rate of 10.6, third quartile had a median rate of 10.1, and the top quartile of hospitals with proportion of black patients had a rate of 10.2. For full presentation of results the measure technical report.

Risk-adjustment model diagnosis and performance with and without social risk factor variables

All three social risk variables exhibit statistically significant association with the outcome, and their directions make sense in explaining their associations with the outcome.

Models exhibit similar performance with and without including social risk variables in the risk adjustment. Specifically,

- C-statistics exhibit very similar model discrimination between risk adjustment using original risk factors and using original risk factors plus social risk variables. For example, for the Validation Split Sample, the inpatient admission measure C-statistics are 0.725 for the model that does not adjust for social risk variables and 0.728 for the model that adjusts for social risk variables. For the ED visit measure, the C-statistics are 0.636 without adjusting for social risk and 0.644 when adjusting for social risk.
- The model calibration results are very similar between risk adjustment using original risk factors and using original risk factors plus social risk variables.
- The results of overfitting indices remained similar with and without adding social risk variables in the risk-adjustment model.

Hospital rankings after considering social risk variables

There is very high agreement of hospital rankings between risk-adjustment models which incorporate social risk variables and those that do not (Spearman rank correlation = 0.988 for the inpatient admission model and 0.984 for the ED visit model), suggesting that accounting for the social risk factors will not have a major impact on hospital rankings.

Initial Model Development: Conclusions

There are clear patient-level effects, but at the hospital level, accounting for patient social risk factors has minimal to no impact on model performance or hospital performance ranking for both the admission or ED measure, indicating that the added risk of being high social risk is captured within current risk variables and arguing against inclusion of patient social risk factors in the chemotherapy measure. Given these findings, we did not include social risk factors in the initial risk-adjustment model for this measure.

2018 Reevaluation. As part of our 2018 reevaluation, we updated our analysis examining the impact of social risk factors on the measure calculation. We evaluated two indicators of social risk: 1) race, specifically African-American or not and 2) the AHRQ SES index. Dual status was not examined due to lack of availability in our reevaluation data.

For these analyses, we used the Medicare FFS FY 2016 Dataset and 2009–2013 data from the Census Bureau’s American Community Survey. These data included 3,562 OPD facilities, 11 PCH facilities, and 289,543 unique patients. Our goal for these analyses were twofold: 1) to examine whether these factors were associated with increased risk in inpatient admissions and ED visits after adjusting for other risk factors and 2) to evaluate the impact of social risk factors on facility-level measure scores. Key findings are detailed below.

Variation in Prevalence of SDS Risk Factors across Hospitals

There is substantial variation in the prevalence of African-American patients and patients with low SES (AHRQ SES Composite Index values below 42.7) across patients in the measure cohort across hospitals in the FY 2016 dataset. For the measure cohort, the facility-level percentage of patients who are African-American ranges from 0% to 100%, with a median of 0% (interquartile range [IQR] 0%-8.1%). The facility-level percentage of patients with low SES ranges from 0% to 100%, with a median of 23.8% (IQR 7.5% - 47.4%).

National Observed Rates

To evaluate the patient-level association of these risk factors with the outcomes, we first quantified the observed rate by each group. We found that African-American patients had higher rates of inpatient admissions, with an observed inpatient admission rate of 14.2% relative to 12.6% for all other patients, as shown in Table 7. This same pattern was true for observed rates of ED visits; the observed rate for African-Americans was 7.6%, whereas it was 5.8% for all others. Patients with low SES also had higher rates of inpatient admissions, with the observed rate of 14.4% relative to 12.4% for patients without low SES, as shown in Table 8. Similarly, patients with a low SES Index value had an observed rate of 7.1% of ED visits relative to 5.7% for non-low SES Index patients. The same pattern held true when results were examined separately for African-American or low SES Index values at PCHs versus non-cancer hospitals (Tables 7 and 8).

Table 7. National Observed Rates for African-American Patients vs. All Others

Hospital Type	Inpatient Admission Observed Rate		ED Visit Observed Rate	
	African-American	All Others	African-American	All Others
All Hospitals	14.2	12.6	7.6	5.8
Cancer Hospitals	14.8	13.9	6.9	6.3
Non-Cancer Hospitals	14.1	12.5	7.7	5.7

Table 8. National Observed Rates for Patients with Low SES (<42.7 AHRQ SES Composite Index values) vs. All Others

Hospital Type	Inpatient Admission Observed Rate		ED Visit Observed Rate	
	Low SES Index	All Others	Low SES	All Others

All Hospitals	14.4	12.4	7.1	5.7
Cancer Hospitals	15.9	13.8	6.7	6.3
Non-Cancer Hospitals	14.3	12.3	7.1	5.7

Association between social risk variables and the outcomes

We then evaluated the patient-level association of these social risk factors with the outcome after adjustment for the age, sex, chemotherapy exposure, concurrent radiotherapy, clinical comorbidities and cancer type variables currently in the inpatient admission and ED visit models. Each factor's effect was quantified using odds ratios (ORs) and testing for significance. In addition, we evaluated the change in the models' predictive ability (c-statistic and range of predictability) when adding SDS factors to the model.

Table 9. Patient-level relationship between social risk factors and the measure outcome

Outcome	Facility Type	Social Risk Factor	p-value	OR (LB, UB)
RSAR	Non-Cancer	African-American	0.412	1.06 (0.92 – 1.23)
RSEDR	Non-Cancer	African-American	0.580	1.06 (0.86 – 1.30)
RSAR	Cancer	African-American	<.0001	1.08 (1.04 – 1.13)
RSEDR	Cancer	African-American	<.0001	1.30 (1.24 – 1.36)
RSAR	Non-Cancer	Low SES Index	0.248	1.07 (0.95 – 1.20)
RSEDR	Non-Cancer	Low SES Index	0.942	0.99 (0.84 – 1.17)
RSAR	Cancer	Low SES Index	0.0003	1.06 (1.03, 1.10)
RSEDR	Cancer	Low SES Index	<.0001	1.16 (1.11 – 1.21)

As shown in Table 9, for **non-cancer hospitals**, African-American race was not statistically significant for either the RSAR and RSEDR models. In the RSAR model, African-American race had a p-value of 0.412, and OR of 1.06, with the 95% confidence interval (CI) for the OR of 0.92 – 1.23. In the RSEDR model, African-American race had a p-value of 0.580, with an OR of 1.06 (95% CI: 0.86, 1.30). The association between Low SES Index and the RSAR and RSDER were similarly non-significant for non-cancer hospitals. In the RSAR model, Low SES Index had a p-value of 0.248, and an OR of 1.07 (95% CI: 0.95, 1.20), while in the RSDER model Low SES Index had a p-value of 0.942, with an OR of 0.99 (95% CI: 0.84, 1.17).

Among the 11 **cancer hospitals**, both African-American and Low SES Index status had a significant association (at $p < 0.001$) with the outcome in the RSAR and RSEDR models. For African American race, the OR was 1.08 (95% CI: 1.04, 1.13) in the RSAR model, and 1.30 (95% CI: 1.24, 1.36) in the RSEDR model. For low SES Index status, the OR was 1.06 (95% CI: 1.03, 1.10) for the RSAR model, and 1.16 (95% CI: 1.11, 1.21) for the RSEDR model.

For **both non-cancer and cancer hospitals**, the addition of either the African-American race or the low SES Index social risk factor had little effect on model c-statistics or predictive ability, as shown in Tables 10 and 11.

Table 10. Comparison of Risk Model Discrimination Statistics with and Without African-American Risk Factor

Model	c-statistic without African-American Risk Factor	c-statistic with African-American Risk Factors	Predictive Ability without African-American Risk Factor	Predictive Ability with African-American Risk Factor
Cancer Hospitals: RSAR	0.711	0.711	3.0 – 31.4%	3.0 – 31.4%
Cancer Hospitals: RSEDR	0.650	0.652	2.5 – 12.5%	2.4 – 12.6%
Non-Cancer Hospitals: RSAR	0.693	0.693	3.2 – 31.4%	3.2 – 31.4%
Non-Cancer Hospitals: RSEDR	0.647	0.647	2.4 – 13.1%	2.4 – 13.1%

Table 11. Comparison of Risk Model Discrimination Statistics with and Without Low SES Index Risk Factor

Model	c-statistic without Low SES Index Risk Factor	c-statistic with Low SES Index Risk Factors	Predictive Ability without Low SES Index Risk Factor	Predictive Ability with Low SES Index Risk Factor
Cancer Hospitals: RSAR	0.711	0.711	3.0 – 31.4%	3.0 – 31.4%
Cancer Hospitals: RSEDR	0.650	0.651	2.5 – 12.5%	2.5 – 12.5%
Non-Cancer Hospitals: RSAR	0.693	0.693	3.2 – 31.4%	3.2 – 31.4%
Non-Cancer Hospitals: RSEDR	0.647	0.647	2.4 – 13.1%	2.4 – 13.1%

Distribution of RSARs and RSEDRs

We further examined the potential impact of these social risk factors on measure scores by comparing RSAR and RSEDR distributions at facilities by proportion of patients with social risk factors (i.e., percent African-American or percent with low SES Index value). Facilities were stratified by the proportion of patients at the facility with each factor, and placed into quartiles based on these proportions. For example, facilities with few African-American patients in their sample would be in the first quartile while facilities seeing high numbers of African-American patients would be in the fourth quartile. We performed a similar analysis for quartiles of the SES Index. These stratified distributions were examined for systematic differences in RSARs and RSEDRs across quartiles. Because a large portion of hospitals with very few (< 25 patients) had no African-American patients, we restricted both the analysis of results for African-American and SES Index quartiles to the 1,535 hospitals with at least 25 patients, which is consistent with public reporting of the measure. In addition, we focus on results for non-cancer hospitals, since there are only 11 cancer hospitals and stratifying by quartile would be comparing only a few hospitals. There are 1,524 non-cancer hospitals with at least 25 patients in the FY 2016 dataset.

As shown in Table 12, facilities with the highest proportion of African-American patients (Q4) had slightly higher RSARs throughout the distribution relative to facilities with the lowest proportion of these patients

(Q1). However, the opposite was true for the RSEDRs, with facilities with the highest proportion of African-American patients experiencing slightly lower rates throughout the distribution, as found in Table 12. With regard to facilities with the highest proportion of low SES Index patients (Q4), both RSAR and RSEDR values were slightly higher relative to facilities with the lowest proportion of low SES Index patients (Q1) (see Tables 12 and 13).

Table 12. Variation in RSARs across Non-Cancer Hospitals by proportion of African-American race and Low SES patients (hospitals with >25 patients; n=1524 hospitals)

	Percent African-American		Percent Low SES Index	
	1 st Quartile (0.00%)	4 th Quartile (≥12.2%)	1 st Quartile (≤47.8%)	4 th Quartile (≥54.3%)
Number of Facilities	383	383	379	383
Number of patients	30,750	74,409	84,698	41,269
Maximum RSAR	17.4	18.5	18.4	18.4
90 th	14.2	15.1	14.7	15.1
75 th	13.2	13.9	13.5	13.9
Median	12.3	12.8	12.5	12.7
25 th	11.6	11.8	11.6	11.8
10 th	11.0	11.2	11.0	11.0
Minimum RSAR	9.3	9.1	8.9	9.1

Table 13. Variation in RSEDRs across Non-Cancer Hospitals by proportion of African-American race and Low SES patients (hospitals with >25 patients; n=1524 hospitals)

	Percent African-American		Percent Low SES Index	
	1 st Quartile (0.00%)	4 th Quartile (≥12.2%)	1 st Quartile (≤47.8%)	4 th Quartile (≥54.3%)
Number of Facilities	383	383	379	383
Number of patients	30,750	74,409	84,698	41,269
Maximum RSEDR	9.4	15.2	9.1	11.5
90 th	7.6	7.4	7.0	7.8
75 th	6.8	6.5	6.1	6.8
Median	6.0	5.8	5.6	6.1
25 th	5.5	5.2	5.1	5.4
10 th	5.1	4.8	4.5	4.9
Minimum RSEDR	3.6	3.3	2.9	3.3

Correlation of RSARs and RSEDRs with Social Risk

Finally, to further understand the relationship between the RSAR and RSEDRs and escalating proportions of patients with high social risk (i.e., higher percentage African-American patients and higher percentage of low SES Index patients), we plotted RSARs and RSEDRs versus the hospital-level proportion of percent African-American and low SES Index patients. We restricted this analysis to non-cancer hospitals with at least 25 patients that were in the highest quartiles for both social risk factors. We then calculated a Pearson correlation statistic to evaluate the relationship at the hospital-level between the risk-adjusted rates and these social risk factors.

As shown in Figures 1 and 2, there was no association between RSAR or RSEDR values and the facility-level percentage of African-American patients. This was confirmed by the Pearson Correlation coefficient, which was 0.047 for the RSAR (p-value = 0.361) and 0.096 for the RSEDR (p-value = 0.061). Similarly for the facility-level percentage of low SES Index patients, there was no significant association with the RSAR or RSEDR, as shown in Figures 3 and 4. This was supported by the Pearson Correlation coefficient, which was -0.022 for the RSAR (p-value = 0.661) and 0.004 for the RSEDR (p-value = 0.945).

Figure 1. RSAR vs. Percent African American, among Non-Cancer Hospitals with Highest Proportion of African-American Patients (Q4) (hospitals with >25 patients; n=1524 hospitals)

Pearson correlation coefficient: 0.047

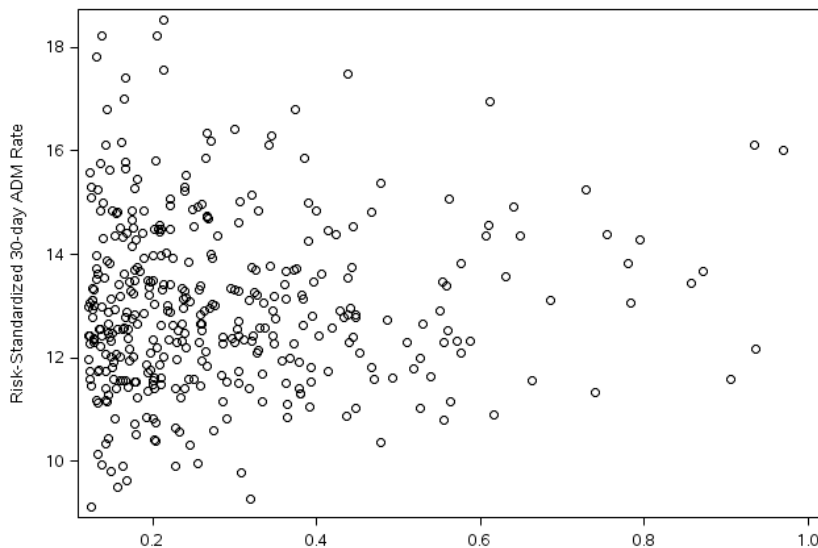


Figure 2. RSEDR vs. Percent African American, among Non-Cancer Hospitals with Highest Proportion of African-American Patients (Q4) (hospitals with >25 patients; n=1524 hospitals)

Pearson correlation coefficient: 0.096

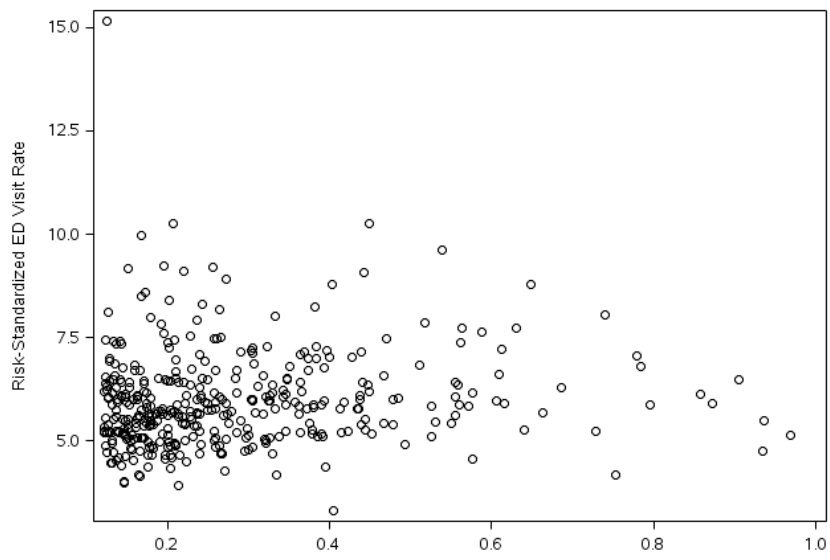


Figure 3. RSAR vs. Percent Low SES Index, among Non-Cancer Hospitals with Highest Proportion of Low SES Index Patients (Q4) (hospitals with >25 patients; n=1524 hospitals) Pearson correlation coefficient: -0.022

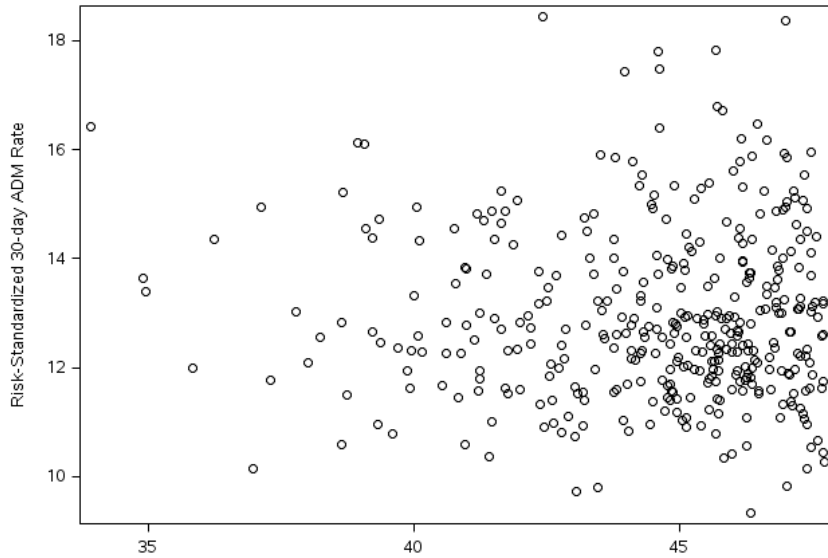
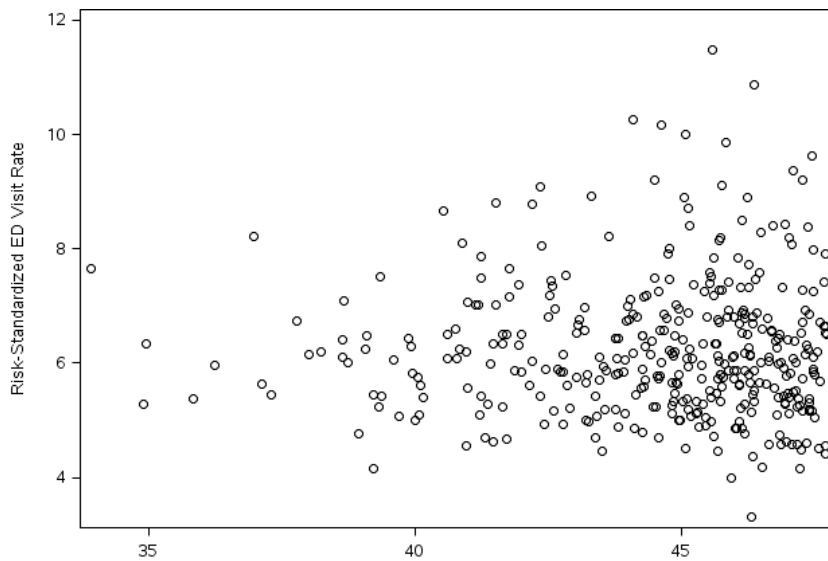


Figure 4. RSEDR vs. Percent Low SES Index, among Non-Cancer Hospitals with Highest Proportion of Low SES Index Patients (Q4) (hospitals with >25 patients; n=1524 hospitals)

Pearson correlation coefficient: 0.004



Impact of SDS Risk-Adjusters on Measure Scores (among Cancer hospitals only)

Because our analysis showed that in **cancer hospitals**, there was a significant patient-level association of both social risk factors (African-American race and low SES Index) with the outcome, we examined the impact of including these variables as risk-adjusters in our model on the hospital-level measure scores. We found that entering these variables into the risk-adjustment model did not substantially change measure scores for cancer hospitals. Correlation coefficients between the measure score with and without adjustment for these factors were 1 or near 1 (see Table 14 below). This indicates that including these social risk factors in hospital-level measure scores for cancer hospitals will result in virtually no differences in hospital-level results after accounting for other risk factors included in the risk model.

Table 14: Correlation between hospital-level measure scores (Pearson Coefficients) with and without social risk factors for cancer hospitals

Outcome	Social Risk Factor	Pearson Coefficient	p-value
RSAR	African-American	0.99982	<.0001
RSEDR	African-American	1.00000	<.0001
RSAR	Low SES Index	0.99898	<.0001
RSEDR	Low SES Index	0.99999	<.0001

2018 Reevaluation: Conclusions

We found that in cancer hospitals, there are clear patient-level effects, as reflected in the significant, relationships between these social risk factors and the two measure outcomes, after adjusting for other risk factors. However, inclusion of social risk factors had no impact on model performance. At the hospital level, the distribution of RSARs and RSEDRs were not consistently higher or lower for facilities with higher proportions of African-American patients, and facilities with fewer low SES Index values had higher values of RSARs and RSEDRs throughout the distribution. There was no obvious statistical relationship between these variables and the measure outcome, as demonstrated by the non-linear, non-significant correlation results. Furthermore, at the hospital level, including the variables in the model did not change hospital-level measure scores. Given these findings, we did not change our original conclusion that SDS factors should not be included in the risk-adjustment models for this measure.

This is consistent with CMS's concern that facilities should not be held to different standards for patients with social risk factors. CMS remains committed to considering options for accounting for social risk factors within individual measures and in the OQR (82 FR 59427) and PCHQR (82 FR 38421) programs.

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)*

We performed a number of tests to evaluate the model performance during initial development, and then re-examined key statistics during subsequent measure reevaluation.

Initial Model Development. We assessed adequacy of the patient-level risk-adjustment models (described above). We evaluated the model performance first in the 2012-2013 Medicare FFS Development Split Sample. We then re-tested the model performance in the 2012-2013 Medicare FFS Validation Split Sample. We did this separately for both the inpatient admission outcome model and the ED visit outcome model.

Using the 2012-2013 Medicare FFS Development Split Sample, we computed three summary statistics for assessing the risk-adjustment model performance: area under the receiver operating characteristic (ROC) curve (c-statistic), predictive ability, and over-fitting indices. We then compared the model performance in the development sample with its performance in the validation sample.

- The c-statistic is a measure of how accurately a statistical model is able to distinguish between a patient with and without a hospital visit. For binary outcomes, the c-statistic is identical to the ROC. A c-statistic of 0.50 indicates random prediction, implying that patient risk factors contribute no additional information. A c-statistic of 1.0 indicates perfect prediction, implying that patients' outcomes can be predicted completely by their risk factors.
- Discrimination in predictive ability measures the ability to distinguish high-risk from low-risk subjects. Good model discrimination is indicated by a wide range between the lowest and highest deciles.
- We assess model calibration by calculating over-fitting indices. Over-fitting refers to the phenomenon in which a model describes the relationship between predictive variables and outcome well in one group of patients, but fails to provide valid predictions in another distinct group of patients. Over-Fitting indices (γ_0 , γ_1) provide evidence of over-fitting and require several steps to calculate. Estimated values of γ_0 far from 0 and estimated values of γ_1 far from 1 provide evidence of over-fitting.

Model Reevaluation. To assess performance of the patient-level risk-adjustment model, the area under the receiver operating characteristic curve, as measured by the c-statistic, was calculated. Observed inpatient admission rates and ED visit rates were compared to predicted inpatient admission and ED visit probabilities across predicted rate deciles to assess calibration, and the range of observed inpatient admission rates and ED visit admission rates between the lowest and highest predicted deciles was also calculated to assess model discrimination.

Several analyses to validate the patient-level risk-adjustment model were performed. First, we compared model performance for the updated model with prior years' models. The c-statistic and model discrimination (predictive ability) were compared. Second, we examined the stability of the risk variable frequencies and regression coefficients between the current model and prior years' models.

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):

Initial Model Development Results

Inpatient admission outcome model

2012-2013 Medicare FFS Development Split Sample results:

- c-statistic=0.73
- Predictive ability (lowest decile %, highest decile %): 2.09-27.70%

2012-2013 Medicare FFS Validation Split Sample results:

- c-statistic=0.72
- Predictive ability (lowest decile %, highest decile %): 2.16-27.98%

ED visit outcome model

2012-2013 Medicare FFS Development Split Sample results:

- c-statistic=0.63
- Predictive ability (lowest decile %, highest decile %): 1.91-8.33%

2012-2013 Medicare FFS Validation Split Sample results:

- c-statistic=0.64
- Predictive ability (lowest decile %, highest decile %): 1.93-8.22%

2018 Model Reevaluation Results

Inpatient admission outcome model

2016 Medicare FFY Dataset, PCHs:

- c-statistic=0.6933
- Predictive ability (lowest decile %, highest decile %): 3.21 – 31.40%

2016 Medicare FFY Dataset, Non-Cancer Hospitals:

- c-statistic=0.7114
- Predictive ability (lowest decile %, highest decile %): 2.98 – 31.43%

ED visit outcome model

2016 Medicare FFY Dataset, PCHs:

- c-statistic=0.6470
- Predictive ability (lowest decile %, highest decile %): 2.35 – 13.1%

2016 Medicare FFY Dataset, Non-Cancer Hospitals:

- c-statistic=0.6504
- Predictive ability (lowest decile %, highest decile %): 2.47 – 12.46%

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

Initial Model Development Results

Inpatient admission outcome model

2012-2013 Medicare FFS Development Split Sample results:

- Calibration: $(\gamma_0, \gamma_1) = (0, 1)$

2012-2013 Medicare FFS Validation Split Sample results:

- Calibration: $(\gamma_0, \gamma_1) = (0.01, 1.00)$

ED visit outcome model

2012-2013 Medicare FFS Development Split Sample results:

- Calibration: $(\gamma_0, \gamma_1) = (0, 1)$

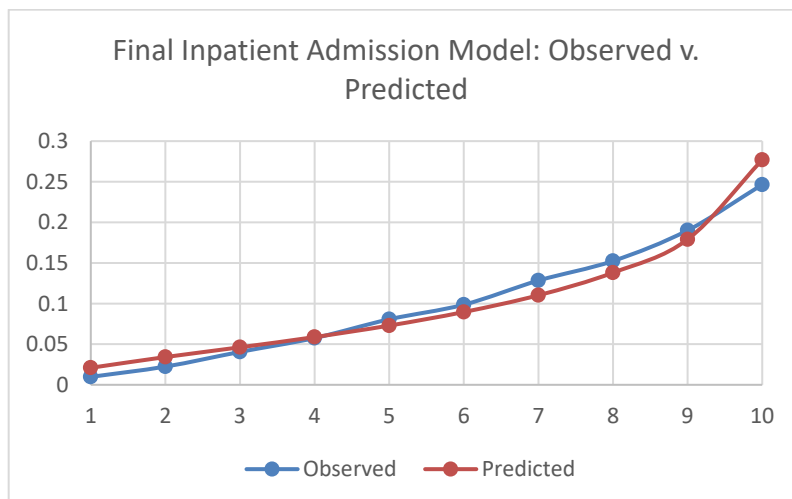
2012-2013 Medicare FFS Validation Split Sample results:

- Calibration: $(\gamma_0, \gamma_1) = (-0.04, 0.99)$

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

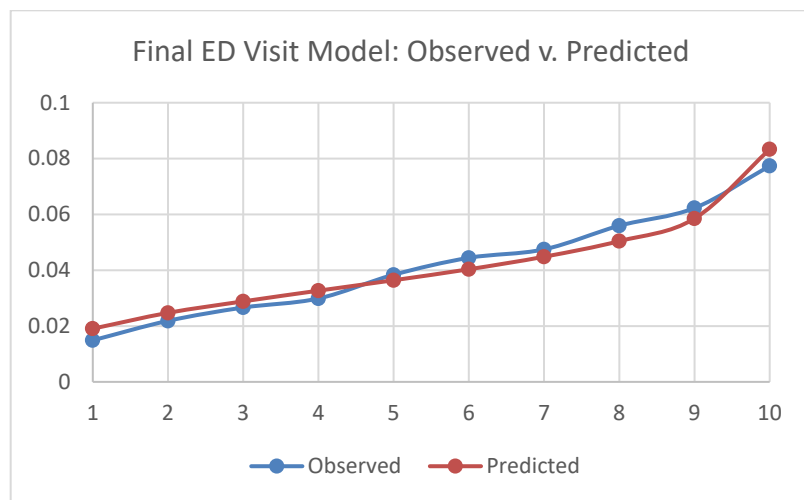
Initial Model Development Results

Figure 5. Inpatient admission outcome model: Plot of observed vs. predicted values for risk deciles (2012-2013 Medicare FFS Development Split Sample)



A second plot using 2012-2013 Medicare FFS Validation Split Sample showed very similar results.

Figure 6. ED visit outcome model: Plot of observed vs. predicted values for risk deciles (2012-2013 Medicare FFS Development Split Sample)



A second plot using 2012-2013 Medicare FFS Validation Split Sample showed very similar results.

2018 Model Reevaluation Results

Figure 7. Inpatient admission outcome model, PCHs: Plot of observed vs. predicted values for risk deciles (2016 Medicare FFS Data)

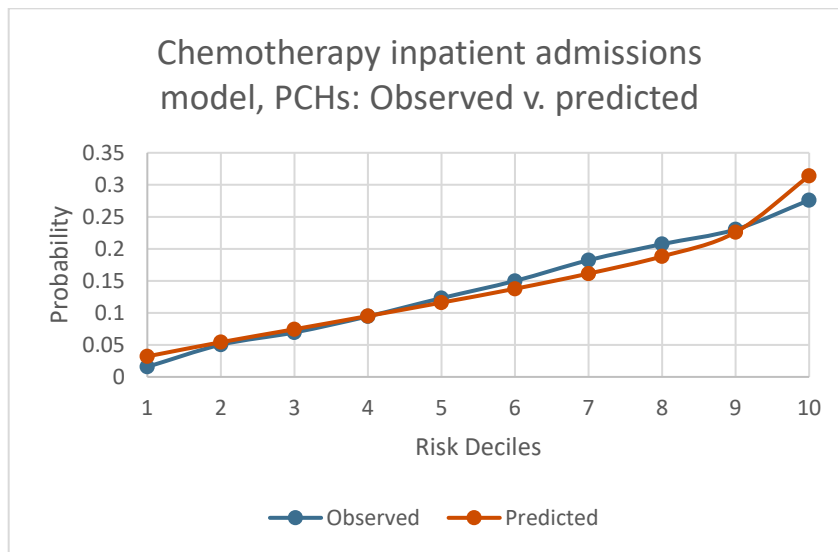


Figure 8. Inpatient admission outcome model, Non-Cancer Hospitals: Plot of observed vs. predicted values for risk deciles (2016 Medicare FFS Data)

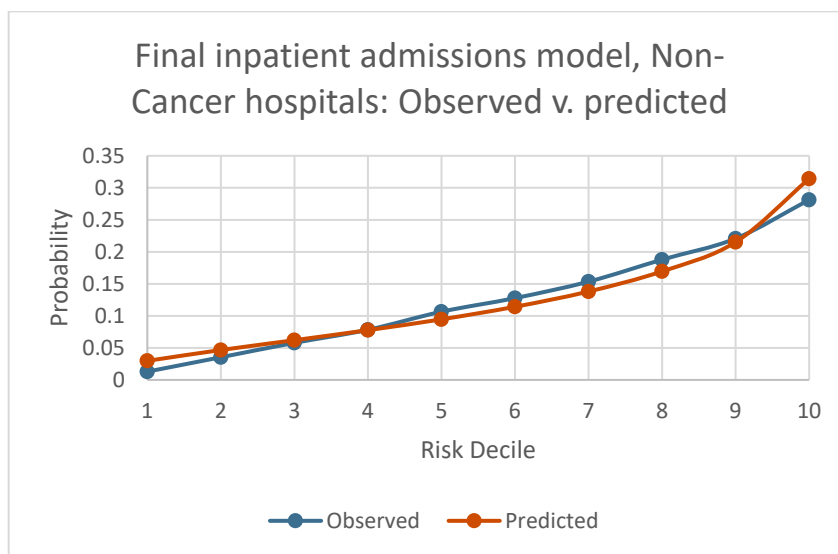


Figure 9. Emergency Department Visits outcome model, PCHs: Plot of observed vs. predicted values for risk deciles (2016 Medicare FFS Data)

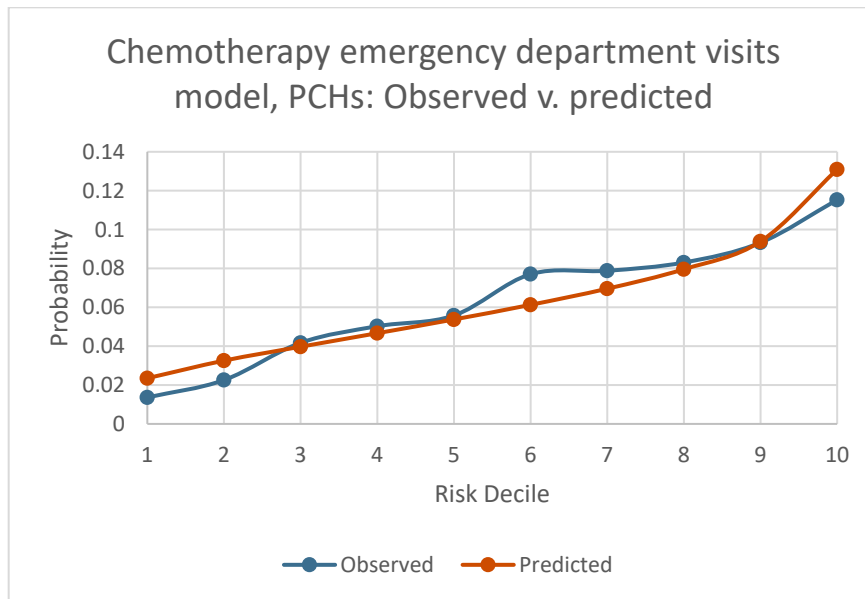
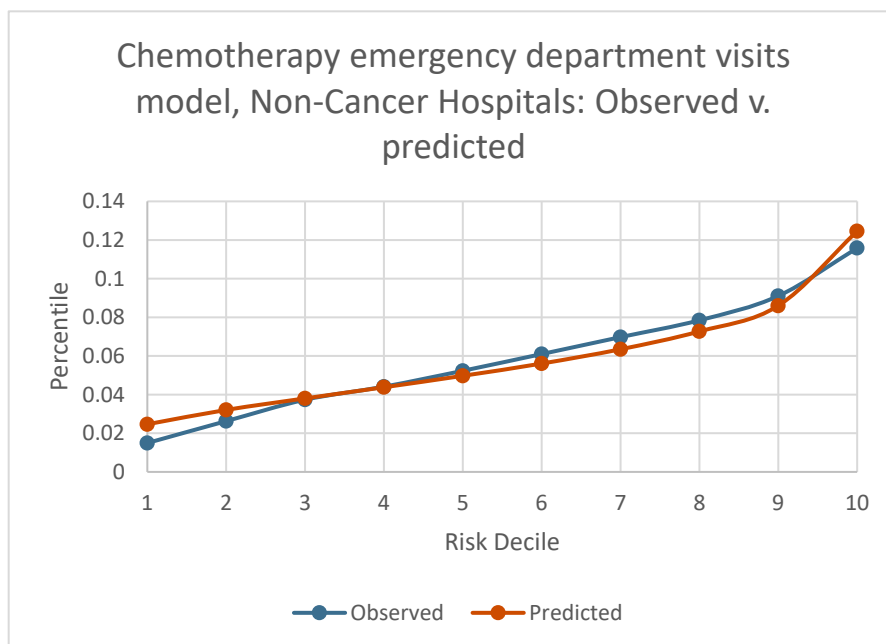


Figure 10. Emergency Department Visits outcome model, Non-Cancer Hospitals: Plot of observed vs. predicted values for risk deciles (2016 Medicare FFS Data)



2b3.9. Results of Risk Stratification Analysis:

Not applicable.

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)

Initial Model Development. For both models, model performance was similar in the development and validation datasets, with strong model discrimination and fit. Predictive ability was also similar across datasets. The c-statistics of 0.73 (inpatient) and 0.63 (ED visit) indicate good model discrimination. The models indicated a wide range in predictive ability between the lowest decile and highest decile, indicating the ability to distinguish high-risk subjects from low-risk subjects. The calibration value of close to 0 and close to 1 indicates

good calibration of the model. Additionally, the risk decile plots show that the model performs similarly in each of the risk deciles across a broad range of risk.

2018 Model Reevaluation. After updating the models to include concurrent radiotherapy and refitting on the newer dataset, we continued to observe strong model discrimination and fit for both outcomes, in both PCHs and non-cancer hospitals. The c-statistics ranged from 0.6470 (RSAR, for PCHs) to 0.7114 (RSEDR, for non-cancer hospitals), indicate good model discrimination. The models continued to show a wide range in predictive ability between the lowest decile and highest decile, indicating the ability to distinguish high-risk subjects from low-risk subjects.

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*)

The measure scores are hospital-level RSAR and RSEDRs, produced separately for PCHs and non-cancer hospital OPDs. The RSAR is calculated as the ratio of the number of predicted qualifying inpatient admissions to the number of expected qualifying inpatient admissions multiplied by the national observed qualifying inpatient admission rate. Similarly, the RSEDR is calculated as the ratio of the number of predicted qualifying ED visits to the number of expected qualifying ED visits multiplied by the national observed qualifying ED visit rate.

For each hospital, the numerator of the RSAR or RSEDR ratio is the number of hospital admissions or ED visits predicted for the hospital's patients, accounting for its observed rate, the age, sex, chemotherapy exposure, radiotherapy exposure, cancer diagnoses and clinical comorbidities. The denominator is the number of hospital visits expected nationally for the hospital's patient population.

To calculate a hospital's predicted-to-expected (P/E) ratio, the measure uses a two-level hierarchical logistic regression model that accounts for the clustering of patients within hospitals and variation in sample size. The log-odds of the outcome for an index chemotherapy procedure is modeled as a function of the patient demographic, exposure, cancer diagnoses, clinical comorbidities, and a random hospital-specific intercept. A ratio greater than one indicates that the hospital's patients have more inpatient admissions or ED visits than expected, compared to an average hospital with similar patient complexity. A ratio less than one indicates that the hospital's patients have fewer inpatient admissions or ED visits than expected, compared to an average hospitals with similar patient complexity.

We characterize the degree of variability in the FY 2016 dataset by:

1. Providing the median odds ratio (MOR) [1]. The median odds ratio represents the median increase in odds of a hospital inpatient admission or visit if a patient received outpatient chemotherapy at a higher risk hospital compared to a lower risk hospital. It is calculated by taking all possible combinations of hospitals, always comparing the higher risk hospital to the lower risk hospital. The MOR is interpreted as a traditional odds ratio would be.
2. Reporting the distribution of the RSAR and RSEDRs.
3. Assessing facility performance by comparing the 95% confidence interval around the RSAR or RSEDR with the program-specific national observed rate, and categorizing the results as follows:

- Better than national rate: If the entire 95% confidence interval of the facility's rate is lower than the national observed rate.
- No different from the national rate: If the 95% confidence interval of the facility's rate includes the national observed rate.
- Worse than national rate: If the entire 95% confidence interval of the facility's rate is higher than the national observed rate.
- Number of cases too small: If a facility does not have at least 25 patients qualifying for the measure, CMS cannot reliably determine how well the facility is performing and therefore does not assign a performance category.

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)

The median odds ratio for PCHs was 1.82 for the RSAR and 2.04 for the RSEDR, and for non-cancer hospital OPDs it was 1.39 for the RSAR and 1.45 for the RSEDR.

Among cancer hospitals, the median RSAR and RSEDR were 13.7% and 6.7%, respectively. The values ranged from 12.3% to 15.2% for RSARs and 3.6% to 9.1% for RSEDRs. For non-cancer hospital OPDs, the median RSAR and RSEDR were 12.5% and 5.6%, respectively. The values ranged from 8.9% to 18.5% for RSARs and 2.9% to 15.2% for RSEDRs. The percentiles of the distribution are shown below in Table 15.

Table 15. Distribution of RSARs and RSEDRs for Cancer and Non-Cancer Hospitals

Percentile	PCHs		Non-Cancer Hospitals	
	RSARs (percent)	RSEDRs (percent)	RSARs (percent)	RSEDRs (percent)
Minimum	12.3	3.6	8.9	2.9
1st	12.3	3.6	10.2	4.2
5th	12.3	3.6	11.1	4.8
10th	13.4	4.1	11.6	5.2
25th	13.4	4.4	12.2	5.6
50th (median)	13.7	6.7	12.5	5.6
75th	14.8	8.9	13.0	6.2
90th	14.8	9.1	13.9	6.8
95th	15.2	9.1	14.8	7.4
99th	15.2	9.1	16.4	8.6
Maximum	15.2	9.1	18.5	15.2

The measure identifies significant differences in performance among both PCHs and non-cancer hospitals. We found that PCHs had a national observed admissions rate of 14.0% and a national observed ED visit rate of 6.3%. Among non-cancer hospitals, the national observed admissions rate was 12.6%, and the national observed ED visit rate was 5.9%. Among the 11 PCHs, 1 was identified as performing significantly better on the RSAR, 3 were identified as performing significantly better on the RSEDR, and 3 were identified as performing significantly worse. For the 3,562 non-cancer hospitals, the measure had additional ability to discriminate

performance, with 13 hospitals performing significantly better on the RSAR, 65 performing significantly worse on the RSAR, 26 hospitals performing significantly better on the RSEDR, and 33 performing significantly worse on the RSEDR.

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?)

The chemotherapy measure produces measure scores that demonstrate meaningful differences in performance across measured entities.

The median odds ratio suggests a meaningful increase in the risk of an inpatient hospital admission or ED visit if chemotherapy was administered at a higher risk facility compared to a lower risk facility. For instance, the PCH inpatient admissions outcome rate has a median odds ratio of 1.82 which indicates that a patient has an 82% increase in the odds of a hospital inpatient admission if the same procedure was performed at higher risk facility compared to a lower risk facility. Median odds ratios across the four models ranged from 1.39–2.04 indicating the impact of quality on the outcome rate is substantial.

The distribution shows a clinically meaningful range of measure scores (RSARs, RSEDRs) for PCHs and non-cancer hospitals (Table 15). In addition, the median measure scores indicate that patients receiving outpatient chemotherapy who are treated at PCHs are expected to experience an inpatient admission on average 13.7% of the time, and an emergency department visit 6.7% of the time; patients treated at non-cancer hospitals are expected to experience an inpatient admission 12.5% of the time, and an emergency department visit 5.6% of the time.

Furthermore, for hospital admission rates (RSARs) the best-performing PCH hospitals (12.3%) are performing 10% better than an average performer, while the worst-performing PCH hospitals (15.2%) are performing 11% worse than an average performer. For ED visit rates (RSEDRs), the best-performing PCH hospitals (3.6%) are performing 46% better than an average performer, while the worst-performing PCH hospitals (9.1%) are performing 36% worse than an average performer. For non-cancer hospitals' admission rates (RSARs), the best-performing hospitals (8.9%) are performing about 30% better than an average performer, while the worst-performing hospitals (18.5%) are performing 48% worse than an average performer. For ED visit rates (RSEDRs), the best-performing non-cancer hospitals (2.9%) are performing 48% percent better than an average performer, while the worst-performing hospitals (15.2%) are performing 1.7 times (or 171%) worse than an average performer.

This variation shows a clear quality gap, as some facilities can achieve substantially lower rates than the average performer, while other facilities are performing meaningfully worse than an average performer. It is important to note that here the average performer refers to a facility with the same case and procedure mix performing at the average.

The significance testing results suggest that the measure has the ability to detect meaningful differences in the quality of care received for adult cancer patients receiving chemotherapy treatment in the hospital outpatient setting. Among non-cancer hospitals, the measure detected 78 outliers for the RSAR (13 significantly better, 65 significantly worse) and 59 outliers for the RSEDR (26 significantly better, 33 significantly worse). Despite there only being 11 PCHs, the measure also detected outliers for these facilities, with 1 outlier identified for the RSAR (significantly better) and 6 outliers identified for the RSEDR (3 significantly better, 3 significantly worse).

Overall, our results suggest that there is substantial need to both reduce the expected rate and the variation in rates across facilities, and that this improvement goal is achievable.

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)

Not applicable, this measure has only one set of specifications.

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)

Not applicable, this measure has only one set of specifications.

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)

Not applicable, this measure has only one set of specifications.

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)

CMS claims and enrollment data are routinely validated for completeness, and we examined the extent of missing data for key variables during measure calculation. No patients or observations were excluded due to missing data.

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)

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)

Not applicable, there was no missing data.

3. Feasibility

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

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.

Coded by someone other than person obtaining original information (e.g., DRG, ICD-9 codes on claims)

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**.

ALL data elements are in defined fields in electronic claims

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).

Not applicable.

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.

Measure development, testing, and 2017 national dry run for implementation in the PCHQR and OQR programs showed that the measure cohort can be defined and outcomes reported using routinely collected Medicare claims and enrollment data. The measure is primarily based on key fields in the claims data that are used for payment and, therefore, have a high level of completeness across claims and are considered reliable.

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).

The measure relies on ICD-10, CPT, UB-04, and HCPCS codes to identify the measure cohort, measure outcomes, and risk factors. There are no licensing requirements or fees for use of ICD-10 and HCPCS data. While the CPT and UB-04 data are readily available on the CMS claims, we note two copyrights:

The American Medical Association (AMA) holds a copyright to the CPT codes utilized in the measure specifications. The AMA assumes no liability for the data contained herein. Applicable FARS/DFARS restrictions apply to government use.

The American Hospital Association (AHA) holds a copyright to the Uniform Bill Codes (“UB”) utilized in the measure specifications. Anyone desiring to use the UB Codes in a commercial product to generate measure results, or for any other commercial use, must obtain a commercial use license directly from the AHA. To inquire about licensing, please contact ub04@healthforum.com.

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 Outpatient Quality Reporting Program (first public reporting planned in January 2020) https://www.qualitynet.org/dcs/ContentServer?cid=1228776191711&pagename=QnetPublic%2FPage%2FQnetTier3&%20c=Page PPS-Exempt Cancer Hospital Quality Reporting Program (first confidential reporting planned in January 2019) https://www.qualitynet.org/dcs/ContentServer?c=Page&pagename=QnetPublic%2FPage%2FQnetTier3&cid=1228776203560

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

The measure is currently used in two CMS programs, the Hospital Outpatient Quality Reporting (OQR) Program and PPS-Exempt Cancer Hospital Quality Reporting (PCHQR) Program.

The Hospital OQR program is a pay-for-reporting program implemented by CMS for outpatient hospital services. The Hospital OQR Program promotes higher quality, more efficient health care for Medicare beneficiaries through measurement. All acute care hospitals paid by Medicare and subject to the Outpatient Prospective Payment System (OPPS) are included; during the 2017 dry run 3,571 hospitals were eligible for the OQR program.

The PCHQR program is a public reporting program implemented by CMS for the 11 PPS-Exempt Cancer Hospitals. It is intended to equip consumers with quality-of-care information to make informed decisions about healthcare options. It is also intended to encourage hospitals and clinicians to improve the quality of care provided to Medicare beneficiaries by ensuring that providers are aware of and reporting on best practices for their respective facilities and type of care.

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?)
Not applicable.

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.)
Not applicable.

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.

Prior to the measure's first public reporting in the Hospital OQR and PCHQR programs, CMS held a confidential, national dry run with all eligible facilities (OQR: 3,571 facilities; PCHQR: 11 facilities) during August/September 2017. During this period, all facilities had the opportunity to ask questions about the measure specifications, and their non-public, facility-level results and detailed patient-level data for the FY 2016 data period provided in Facility-Specific Reports (FSRs). CMS provided reports to the 11 PCH hospitals and 3,571 non-PCH hospitals during the dry run. We received and responded to 216 questions during this period, 3 from PCHs and 213 from non-cancer hospitals.

CMS adopted the measure for public reporting in the Hospital OQR program beginning in CY 2020 (81 FR 79764) and in the PCHQR program for confidential reporting, beginning in FY2019, and future public reporting (81 FR 57190). Prior to publicly reporting measure results on Hospital Compare, CMS will release annual Facility-Specific Reports (FSRs) to facilities in both the OQR and PCHQR programs which provide the facility with a summary of their performance on the measure, national performance on the measure, patient data, and characteristics of the patient population at the facility and in the nation. Facilities in the OQR program also receive semi-annual claims-detail reports (CDRs) which provide them with patient data for their facility so that they can see how they are performing ahead of release of the annual FSR.

Facilities wishing to ask questions regarding the measure are able to do so using the question and answer tool on QualityNet. Additionally, each program's QualityNet site includes a measure page for this measure. The page includes measure methodology, fact sheet, frequently asked questions, and archived information from the measure dry run.

Facility-level results are then published on CMS's Hospital Compare website, where they are available to the general public.

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.

Facilities in the PCHQR and OQR programs receive confidential FSRs once per calendar year. The FSR provides the facility with a summary of their performance on the measure, national performance on the measure, detailed patient data, and characteristics of the patient population at the facility and in the nation. In addition to the FSR, facilities in the OQR program receive two CDRs per calendar year that provide interim detailed patient level data for their facility, prior to the annual FSR.

The first distribution of confidential FSRs occurred in August 2017, as part of the measure's national dry run. As part of the dry run, CMS held a 45-day public comment period from August 15 through September 29, 2017. During this period, facilities participating in the PCHQR and OQR programs had the opportunity to ask questions about the measure specifications, and their non-public, facility-level results for the FY 2016 data period. We received 216 questions during this period, 3 from PCHs and 213 from non-cancer hospitals. In

addition, CMS hosted a national provider call on August 23, 2017 to review the measure specifications, share national results, and answer stakeholder questions.

Facilities wishing to ask questions or looking for information regarding the measure are able to do so using the question and answer tool on QualityNet (www.qualitynet.org). Additionally, each program's QualityNet site includes a measure page for this measure. The page includes measure methodology, fact sheet, frequently asked questions, and archived information from the measure dry run. These materials are updated prior to every confidential or public reporting period for the measure.

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.

During the measure's national dry run, CMS held a 45-day public comment period from August 15 through September 29, 2017. During this period, facilities participating in the PCHQR and OQR programs had the opportunity to ask questions/comment about the measure specifications, and their non-public, facility-level results for the FY 2016 data period. We received 216 questions during this period, 3 from PCHs and 213 from non-cancer hospitals. (See 4a2.2 for types of feedback received.)

We used the feedback from all of these sources to refine the measure specifications during the initial development phase and then during reevaluation. They served as a source of ongoing face validity review on key aspects of the measure, including the codes and logic used to define the cohort, outcomes, exclusions, and risk-adjustment models.

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

The majority of feedback from those being measured came during the 2017 national dry run of the measure. During the dry run, the most common feedback we received from facilities involved the following three topics:

- 1.) Patients were included in the measure cohort who were receiving chemotherapy treatment for an autoimmune disease and not cancer;
- 2.) Concern over patients being included in the outcome who were admitted for planned procedures (e.g., for stem cell transplantation); and,
- 3.) Concern over patients being included in the cohort who had Leukemia in remission.

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

During the measure's first NQF endorsement review in 2016, members of the NQF Cancer Committee expressed concern over inclusion of patients in the measure receiving concurrent chemotherapy and radiotherapy, noting that these patients are at higher risk for an outcome due to increased exposure to toxins. In response to this feedback, the 2018 EWG recommended revising the risk-adjustment model to ensure that facilities treating a higher proportion of patients receiving concurrent chemotherapy and radiotherapy were not penalized for providing treatment to higher risk patients.

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.

In order to address the comments received from facilities being measured and other users, we implemented a number of updates to our measure specifications ahead of implementation of the measure into the OQR program (note: due to timing issues, these changes in specifications were not included in the first year of PCHQR confidential reporting but will be included in subsequent years). Specifically, we:

- 1.) Implemented a new case-level exclusion in which patients receiving chemotherapy to treat a qualifying autoimmune condition rather than cancer are excluded from the measure. Cases qualifying for this exclusion are identified by the presence of a chemotherapy code and an autoimmune diagnosis, and the absence of a cancer diagnosis code;

2.) Implemented new logic into the measure that identifies and excludes outcomes identified as “always planned.” The measure considers inpatient hospital admissions with the following AHRQ Clinical Classification Software (CCS) procedures or diagnoses as always planned, and they do not qualify as an outcome for the chemotherapy measure:

Procedures

- AHRQ CCS 64 – Bone marrow transplant
- AHRQ CCS 105 – Kidney transplant
- AHRQ CCS 176 – Other organ transplantation (other than bone marrow corneal or kidney)

Diagnoses

- AHRQ CCS 45 – Maintenance chemotherapy; radiotherapy

3.) Reviewed and revised the code set for exclusion of patients with leukemia to also exclude patients with leukemia in remission; and,

4.) Added a new risk-adjustment variable to the risk models for both outcomes that assesses whether a patient is receiving concurrent radiotherapy and chemotherapy. We define concurrent treatment, based on recommendations from the measure’s expert work group, as receipt of radiotherapy on the date of chemotherapy or up to 14 days before administration of chemotherapy [1].

Citations

1.) Church, D.N., Flubacger, M., Cameron, A., et al. “Toxicity of concurrent radiotherapy with CMF chemotherapy in the E-CMF adjuvant breast carcinoma regimen.” *Journal of Clinical Oncology* 25, no. 18_suppl (June 20 2007) 582-582.

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.

The measure has been adopted for public reporting in the Hospital OQR program beginning in CY 2020 (81 FR 79764) and for confidential reporting, beginning in FY2019, and future public reporting in the PCHQR program (81 FR 57190). [1] In preparation for the first year of public reporting in these programs, the measure underwent a confidential, national dry run in August/September 2017, using the FY 2016 dataset described in Sections 1.2 and 1.7 of the Testing Attachment. However, subsequent years of results are not yet available for comparison.

As described above, the Hospital OQR and PCHQR programs promote quality improvement through the public reporting of measure results. We expect there to be improvement in annual measure scores over time as public reporting of chemotherapy measure results through these two programs identifies and illuminates opportunities for improvement in outpatient chemotherapy care to providers, patients and other stakeholders.

In addition to hospital performance being publicly reported, each participating hospital receives patient-level data outlining details of cases and outcomes attributed to their facility. Low performing hospitals will be able to use this data to make informed decisions on how to improve current protocols or develop new interventions aimed at improving quality.

Footnote

1. This measure's testing form notes that the measure will be publicly reported in the PCHQR program beginning in FY2019, however, we were informed on September 25, 2018 that the measure will be confidentially reported to facilities in FY2019, with public reporting planned for a future year.

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.

We did not identify any unintended consequences during measure development, model testing, or confidential reporting of the measure during its national dry run. However, during the NQF Measure Applications Partnership (MAP) review of this measure in December 2015, the MAP expressed concerns about a possible unintended consequence related to treatment decisions and underuse of appropriate care. The MAP's concern was that the measure might indirectly discourage more aggressive treatment plans that would have had clinical benefits. However, the purpose of the measure is to open lines of communication between the patient and provider on risks and preventative actions that can be taken for each type of treatment, and set the expectations for the patient so they can make more informed decisions on healthcare utilization as well [1].

Furthermore, the measure is risk-adjusted to help account for the variation in patient mix and aggressiveness of treatment. For example, aggressiveness of chemotherapy regimens can range by cancer type and patient age, which are accounted for in our models. We also adjust for the number of treatments and whether or not the patient is receiving radiotherapy concurrently, both of which may also be indicators of aggressiveness of treatment. Lastly, the measure rate is not intended to be zero and CMS recognizes that not all admissions and ED visits are avoidable. To this end, CMS only categorizes hospitals with rates significantly higher or lower than the national rate as performing either "worse" or "better," as described in more detail in Section 2b4 of the Testing Attachment. Improving patient/provider communication and appropriately adjusting the model mitigates the risk of the unintended consequences.

We are committed to monitoring this measure's use and assessing potential unintended consequences over time.

Citation

1. Aprile, G., F.E. Pisa, A. Follador, L. Foltran, F. De Pauli, M. Mazzer, S. Lutrino, C.S. Sacco, M. Mansutti, and G. Fasola. "Unplanned Presentations of Cancer Outpatients: A Retrospective Cohort Study." *Supportive Care in Cancer*, vol. 21, no. 2, 2013, pp. 397–404.

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

Not applicable. There were no unexpected findings identified during testing of this measure and the measure has not yet been publicly reported.

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)

0383: Oncology: Plan of Care for Pain – Medical Oncology and Radiation Oncology (paired with 0384)

0384e: Oncology: Medical and Radiation - Pain Intensity Quantified

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?

No

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

We identified three related NQF-endorsed measures. All three measures (NQF 0383, NQF 0384e, and NQF 1628) focus on cancer patients receiving outpatient chemotherapy; however, there are some key differences in measure scope and measure type. Measure scope: Each of the three related measures (NQF 0383, NQF 0384e, and NQF 1628) narrowly focuses on pain management and/or fatigue/anemia. The proposed measure does not target a specific symptom, but rather assesses the overall management of 10 important symptoms and complications that were more frequently cited in literature as reasons for ED visits and inpatient admissions following outpatient chemotherapy. Measure type: The three related measures (NQF 0383, NQF 0384e, and NQF 1628) are all process measures encouraging the use of screening and care plans to improve care. The proposed measure is an outcome measure not encouraging or measuring specific processes to detect and treat these conditions, but rather assessing the outcomes of the care being provided. The three process measures, which are not risk-adjusted, support the intent of the measure by reinforcing that those providing outpatient care should screen for and manage symptoms such as pain.

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.)

Not applicable.

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.

Attachment **Attachment:** ChemoMeasure_NQF_Appendix_SDS_and_Technical_Report-636771081790630717.pdf

Contact Information

Co.1 Measure Steward (Intellectual Property Owner): Centers for Medicare and Medicaid Services (CMS)

Co.2 Point of Contact: Vinitha, Meyyur, Vinitha.Meyyur@cms.hhs.gov, 410-786-8819-

Co.3 Measure Developer if different from Measure Steward: Mathematica Policy Research

Co.4 Point of Contact: Jessica, Ross, JRoss@mathematica-mpr.com, 617-674-8384-

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.

The 2018 Expert Workgroup (EWG) advised on refinements of the technical specifications during 2018 measure reevaluation, and provided an assessment of the measure's face validity. Prior EWGs similarly provided guidance either during the measure's initial development or subsequent revisions.

2018 EWG members

1. Robert Daly, MD, MBA – Memorial Sloan-Kettering Cancer Center (Staff Physician, Medical Oncology)
2. Stephen Edge, MD* – Roswell Park Memorial Institute (Vice President, Healthcare Outcomes and Policy, Professor of Oncology and Surgical Oncology)
3. Michael Hassett, MD, MPH* – Dana Farber Cancer Center (Attending Physician, Medical Oncology; Assistant Professor, Medicine, Harvard Medical School)
4. Scott Huntington, MD, MPH – Yale New Haven Hospital (Attending Physician, Hematology)
5. Denise Morse, MBA – City of Hope Cancer Treatment and Research Center (Senior Manager, Quality Analytics)
6. Joseph Ross, MD, MHS – Yale University School of Medicine (Associate Professor of General Medicine and of Public Health)
7. Weijing Sun, MD – University of Kansas Cancer Center (Director of Medical Oncology and Associate Director of University of Kansas Cancer Center)
8. Allison Snyderman, Ph.D.* - Memorial Sloan-Kettering Cancer Center (Outcomes Researcher)

*Also served as member of TEP and 2014 PPS-Exempt Cancer Workgroup

2017 EWG members

1. Susan Armstrong—City of Hope Cancer Treatment and Research Center (Senior Manager, Coding and Data Quality)
2. Arnold Chen, MD, MPH – Mathematica Policy Research (Clinician, Senior Researcher)
3. Michael J. Hassett, MD, MPH—Dana-Farber Cancer Institute, Outpatient Clinic and Treatment Center (Staff Physician, Breast Oncology)

4. Joseph Ross, MD, MHS – Yale University School of Medicine (Associate Professor of General Medicine and of Public Health)

5. Tracy Spinks, CPH—MD Anderson Hospital, Cancer Care Delivery (Program Director)

6. Denise Stone, RN, MBA – Mathematica Policy Research (Clinician, Lead Program Analyst)

2016 EWG members

1. Peter B. Bach, MD, MAPP—Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Director)

2. Stephen Edge, MD—Baptist Cancer Center (Cancer Center Director)

3. Michael J. Hassett, MD, MPH—Dana-Farber Cancer Institute, Outpatient Clinic and Treatment Center (Staff Physician, Breast Oncology)

4. Karl Lorenz, MD, MSHS—UCLA, Department of Veterans Affairs (Associate Professor); Rand Health

5. Allison Snyderman, PhD —Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Researcher)

6. Tracy Spinks, CPH—MD Anderson Hospital, Cancer Care Delivery (Program Director)

2015 EWG members

1. Peter B. Bach, MD, MAPP—Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Director)

2. Stephen Edge, MD—Baptist Cancer Center (Cancer Center Director)

3. Michael J. Hassett, MD, MPH—Dana-Farber Cancer Institute, Outpatient Clinic and Treatment Center (Staff Physician, Breast Oncology)

4. Karl Lorenz, MD, MSHS—UCLA, Department of Veterans Affairs (Associate Professor); Rand Health

5. Allison Snyderman, PhD —Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Researcher)

6. Tracy Spinks, CPH—MD Anderson Hospital, Cancer Care Delivery (Program Director)

PPS-Exempt Cancer hospital workgroup members

1. J. Robert Beck, MD—American Oncologic Hospital (Fox Chase) (Senior Vice President and Chief Academic Officer)

2. Joe Jacobson, MD—Dana Farber Cancer Institute (Chief Quality Officer)

3. Barbara Jagels, MHA, RN, OCN—Seattle Cancer Care Alliance (Fred Hutchinson Cancer Research Center) (Director of Nursing and Clinical Excellence)

4. Dana Jenkins—Roswell Park Memorial Institute (Vice President of Organizational Improvement)

5. Tricia Kassab, RN, MS, CPHQ, HACP—City of Hope National Medical Center (Vice President of Quality and Patient Safety)

6. Jeremy Miransky, PhD—Memorial Hospital for Cancer and Allied Disease (MSKCC) (Quality Analytics Manager)

7. Shyroll Morris, MBA, MPH—University of Miami Hospital and Clinics

8. Thomas Ross, MS—H. Lee Moffitt Cancer and Research Institute Hospital, Inc. (Director of Quality and Safety)

9. Anthony Senagore, MD—University of Southern California Kenneth Norris Jr. Cancer Hospital (Chief of Colorectal Surgery)

10. Ron Walters, MD, MHA, MBA—The University of Texas MD Anderson Cancer Center (Associate Vice President of Medical Operations and Informatics)

11. Saul Weingart, MD, PhD —Dana Farber Cancer Institute (Vice President for Quality Improvement and Patient Safety)

TEP members

1. Peter B. Bach, MD, MAPP—Memorial Sloan-Kettering Cancer Center, Center for Health Policy and Outcomes (Director)
2. Stephen Edge, MD—Baptist Cancer Center (Cancer Center Director)
3. Andrew Glass, MD—Kaiser Permanente Northwest, Center for Health Research (Senior Investigator)
4. Mark Gorman—Independent Consultant (Patient Advocate)
5. Michael J. Hassett, MD, MPH—Dana-Farber Cancer Institute, Outpatient Clinic and Treatment Center (Staff Physician, Breast Oncology)
6. Karl Lorenz, MD, MSHS —UCLA, Department of Veterans Affairs (Associate Professor); Rand Health
7. Joan McClure, MS—National Comprehensive Cancer Network, Clinical Information and Publications (Senior Vice President)
8. Bruce Minsky, MD—MD Anderson Hospital, Department of Radiation Oncology (Professor and Director of Clinical Research)
9. Shirley Stagner, MSN, ONP, AOCNP—Lawrence Hospital Center, Cancer Survivorship Program (Nurse Practitioner)
10. Janet H. Van Cleave, PhD, MSN, AOCNP—New York University College of Nursing (Assistant Professor)
11. Sandra L. Wong, MD MS—University of Michigan Health System, Division of Surgical Oncology

Measure Developer/Steward Updates and Ongoing Maintenance

Ad.2 Year the measure was first released: 2017

Ad.3 Month and Year of most recent revision: 08, 2018

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? 07, 2019

Ad.6 Copyright statement: Not applicable.

Ad.7 Disclaimers: Not applicable.

Ad.8 Additional Information/Comments: Not applicable.

Methods Panel Scientific Acceptability: Combined Preliminary Analysis Form

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.**

- **PANEL MEMBER 1:** None.
- **PANEL MEMBER 2:** I am a bit confused as why this is considered one measure, when it actually measures two mutually exclusive outcomes. Would feel "cleaner" if it were one measure. In addition, the reliability and validity of the two measures could be different, so that makes it difficult to assess them as one measure.
- **PANEL MEMBER 3:** My understanding is that there are two rates reported: 1) inpatient admission and 2) ED visit, and that these outcomes are mutually exclusive (i.e. for a given case the outcome may include one or the other, but not both). I do wonder whether there might be an inverse relationship between the two outcomes (facilities with a higher inpatient rate have a lower ED rate) in the same manner than readmissions and mortality rates often have an inverse relationship.
- **PANEL MEMBER 4:** The numerator (section S.4) is described as being composed of "two mutually exclusive outcomes" – one or more hospital admissions and one or more ED visits. But these are not mutually exclusive; one can have both an ED visit and be admitted. Indeed, I suspect many admissions start out as ED visits. Reading further, and seeing that there are separate risk adjustment models for each outcome, it seems that there are really two measures: (1) admissions and (2) ED visit rates.
 - In addition, the Brief description (De.3) refers to "cancer patients =18 years of age"; presumably this is a typo, meaning "18 years of age or older." But more substantively, since this only refers to Medicare patients, nearly all will be 65 years of age or greater. It might be better to focus on this group, and to say so.
- **PANEL MEMBER 5:** No concerns.

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

- **PANEL MEMBER 1:** Both split sample reliability (for overall measure reliability) and facility-level reliability were assessed. The rationales and the specific steps for implementing both these approaches are well explained.
- **PANEL MEMBER 2:** Used split-sample and signal-to-noise analyses to assess the reliability of the measure. From what I understand from their description, split-sample was relay test-retest.
- **PANEL MEMBER 3:** The developer uses the Beta-binomial model (Adams 2009) to estimate signal-to-noise
- **PANEL MEMBER 4:** Appropriate use of both split-sample and facility-level reliability.
- **PANEL MEMBER 5:** Split-sample (test-retest)

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

Submission document: Testing attachment, section 2a2.3

- **PANEL MEMBER 1:** As documented in testing document, the proposed measure score has sufficient overall reliability. However, for non-cancer hospitals, reliability of the measure score generally fair
- **PANEL MEMBER 2:** Concerns with the relatively low ICC for non-cancer hospitals, [RSAR = 0.4134; RSEDR = 0.3585]
- **PANEL MEMBER 3:** The developer reports median reliability of 0.7848 for the RSAR, and 0.9808 for the RSEDR, which is moderate to high in addition to distribution by volume
- **PANEL MEMBER 4:** Reliability seems sufficient.
- **PANEL MEMBER 5:** High-reliability in cancer hospitals; however, low-reliability in non-cancer hospitals that make use of this measure in non-cancer hospitals suspect. What is the rationale for the poor reliability in non-cancer hospitals to consider modifications to the measure to improve it for these types of facilities?

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.
- **PANEL MEMBER 1:** My ranking of “moderate” is based on the moderate to fair Intraclass Correlation Coefficient (ICC) scores in Risk-Standardized Admission Rate and Risk-Standardized Emergency Department Visit Rate for non-cancer hospitals.
 - **PANEL MEMBER 2:** Really two measures, so each has their own reliability statistic. I do have concerns with the reliability for non-cancer hospitals
 - **PANEL MEMBER 3:** The submission is close to a best practice in the reporting of reliability.
 - **PANEL MEMBER 4:** The reliability results suggested, as the developers say, “that there is sufficient reliability in the measure score.”
 - **PANEL MEMBER 5:** Moderate because it is high for one type of facility (cancer) and low for other facilities (non-cancer).

VALIDITY: ASSESSMENT OF THREATS TO VALIDITY

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

Submission document: Testing attachment, section 2b2.

- **PANEL MEMBER 3:** None
- **PANEL MEMBER 4:** No concerns.
- **PANEL MEMBER 5:** No concerns.

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

Submission document: Testing attachment, section 2b4.

- **PANEL MEMBER 2:** No concerns.
- **PANEL MEMBER 3:** None
- **PANEL MEMBER 4:** There is insufficient information on which to judge.
- **PANEL MEMBER 5:** The distribution of measure scores is broad for cancer and non-cancer hospitals indicating room for improvement; however, with the measure not performing reliably for non-cancer hospitals, it is difficult to ascribe valid meaning to this type of facility.

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.

- **PANEL MEMBER 1:** N/A
- **PANEL MEMBER 2:** Not applicable.
- **PANEL MEMBER 3:** None
- **PANEL MEMBER 4:** NA
- **PANEL MEMBER 5:** No concerns.

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

Submission document: Testing attachment, section 2b6.

- **PANEL MEMBER 2:** Not applicable.
- **PANEL MEMBER 3:** None.
- **PANEL MEMBER 5:** No concerns.

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

- **PANEL MEMBER 1:** I agree with the risk-adjustment approaches. For each of the two outcomes measure (hospital admission rate and ED visit rate), a two-level hierarchical logistic regression model was used to estimate risk-adjusted outcomes rates. The approach accounts for patient mix, clustering of patients within hospitals and variation in sample size.
- **PANEL MEMBER 2:** From what I can tell, “concurrent radiology” would not be consider a risk factor present at the start of care. One concern might be the conflation of complications from radiology with complications from chemotherapy?
- **PANEL MEMBER 3:** Acceptable
- **PANEL MEMBER 4:** Appropriate
- **PANEL MEMBER 5:** No concerns.

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

- **PANEL MEMBER 1:** As detailed in Section 2b1.2 of the testing document, the face validity of the measure has been established through application measure development guidelines, through assessment by external groups (Technical Expert Panel, Expert Working Groups), and through incorporation of public comments.
- **PANEL MEMBER 2:** From what I understand, it sounds like a panel of 8 persons reviewed the measure and voted on the measure’s usefulness in establishing quality of care. It sounded like all 8 were involved in helping develop the measure. I am a bit concerned about the potential bias that those individuals may have for believing the measure is a “good” measure.
- **PANEL MEMBER 3:** No empirical validity testing of measure score

- **PANEL MEMBER 4:** The developers used a very complex method to get input from many sources to assess face validity.
- **PANEL MEMBER 5:** Face validity appropriately assessed through use of measure guidelines, technical expert panels, work groups and public comments.

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

Submission document: Testing attachment, section 2b2.3

- **PANEL MEMBER 1:** Based on face validity testing, the measure had at least moderate
- **PANEL MEMBER 2:** 100% the panel members thought the measure had strong face validity
- **PANEL MEMBER 3:** No empirical validity testing of measure score
- **PANEL MEMBER 4:** After this intensive process, there was strong agreement among experts that the measures have face validity.

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.**

- **PANEL MEMBER 1:** The testing results demonstrated that the measure has moderate reliability and validity.
- **PANEL MEMBER 2:** As stated above, I have concerns that using a panel to assess the face validity of a measure where all of the panel members were part of the measure development process. A stronger test of face validity would be test it with those who are not involved in the measure's development.
- **PANEL MEMBER 3:** No empirical validity testing of measure score or data elements. If this were a maintenance measure the rating would be insufficient.

- **PANEL MEMBER 4:** Although the develops used an extensive formal process, at the end of the day it's simply face validity.
- **PANEL MEMBER 5:** The measure only received face validity testing, which is insufficient, given it is an outcome measure seeking continued endorsement. Empirical validity testing is expected at this point in the measure's life cycle.

25. **Briefly explain rationale for rating of EMPIRICAL ANALYSES TO SUPPORT COMPOSITE CONSTRUCTION**

PANEL MEMBER 4: Note: I did not regard this as a composite, but rather two separate rates for hospital admissions and ED visits.

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.**

PANEL MEMBER 4: None.