ASCO® Measures Methodology Manual

Application: Applies to ASCO and its affiliates

History: Approved by the ASCO Board of Directors on 8/20/2024

Table of Contents

1. Introduction

The American Society of Clinical Oncology (ASCO) Measure Development Methodology Manual communicates the methods by which ASCO develops, maintains, and implements oncology measures, and is updated every three years (approximately). The ASCO measures program falls under the purview of the ASCO Measures Steering Group (MSG), a subgroup of the Care and Quality Improvement Committee (CQIC). The MSG oversees measure topic prioritization, measure development, the formation and progress of technical expert panels (TEPs) and is the review and approval body for all clinical quality measure products **(Appendix I).**

Measures inform stakeholders how the health care system is performing. They also help identify weaknesses, prioritize opportunities, and can be used to identify mechanisms to drive quality improvement. Measures can prevent the over, under, or misuse of health care services and can identify disparities in care delivery and patient outcomes.

1.1 History of Quality Measures

Measuring quality in healthcare dates to the 18th century when the Pennsylvania Hospital collected patient outcomes data organized by diagnoses codes.[1] In the 1960s, quality measurement was embedded into Centers for Medicare and Medicaid Services (CMS) programs with the start of utilization review committees which, to prevent fraud and abuse, required hospitals to review whether services were necessary.[1] In 1966, Donabedian's Framework of structure, process, and outcome, laid the groundwork for evaluating the quality of medical care. In 1991, the National Committee for Quality Assurance (NCQA) developed the Healthcare Effectiveness Data and Information Set (HEDIS) as a tool for measuring the performance of health plans, and in the early 2000s, CMS began developing hospital reporting measures[1]. The Affordable Care Act (ACA) mandated that CMS have payment schemes that embedded cost and quality metrics for value-based payments.

Although initially envisioned as metrics to inform physician-led quality improvement efforts, measures have become the cornerstone of accountability and performance-based reimbursement. The emerging value-based market has required increasingly complex and sophisticated measures, including the need for more comprehensive and difficult to capture metrics that involve patient-centeredness and outcomes across the continuum of care. ASCO develops and maintains quality measures through a rigorous process that aligns with strict standards and requirements set by CMS and the Partnership for Quality Measurement (PQM), which drive the uptake of measures in federal and private payer quality and accountability programs. Wide implementation of ASCO measures, where appropriate, is achieved by several means, including pursuit of Consensus-Based Entity (CBE) endorsement, submission to the CMS Measures Under Consideration (MUC) list for potential use in federal programs, evaluation for inclusion in internal ASCO quality programs and initiatives, and inclusion in ASCO's measure library where measures are available to license.

1.2 General and Technical Principles for Measure Development

The principles outlined in **Appendix II** are used throughout the measure development process, especially when identifying concepts for de novo measures. These principles serve as ASCO's overarching strategy

for measure development that meets the standards and rigor expected for a meaningful, valid, and useful measure.

2. Measure Use

At their core, quality measures show performance of an action or outcome of interest, and these data can be used in a variety of ways, including quality improvement, accountability, and research. Appropriate use of a measure varies depending on the level of evidence supporting the measure. Quality improvement measures based on lower-level evidence may not be appropriate for use in programs tied to accountability and/or payment. Accountability measures generally require a higher level of evidence, such as systematic reviews of the evidence, clinical practice guidelines with descriptions of the level of evidence and strength of recommendations, and/or demonstrated utility.

2.1 Quality Improvement

Quality measures can be used for both quality improvements internal to an institution or system of care, or externally across institutions of systems of care. Internal quality improvement involves three basic steps: identifying gaps in care or opportunities for improvement, selecting appropriate measures to address the gap, and obtaining a baseline assessment of current practice before remeasuring to assess the effect of improvement efforts on measure performance.

External quality improvement measures may be used in programs operated by state, regional, or national entities or organizations, accreditation and quality improvement organizations, or professional societies. The end users of the results of quality measurement are the participating institutions and providers therein.

2.2 Public Reporting & Accountability

Measures used in public reporting and accountability programs are not only used to drive improvement, but also to hold clinicians accountable for data reporting and measure performance. Both public and private payers use measures for various accountability purposes, including public reporting, pay-for-reporting, and pay-for-performance programs. For example, CMS created the Five-Star Quality Rating System to help consumers compare nursing homes more easily by providing ratings for health inspections, staffing, and quality measures. The CMS Quality Payment Program (QPP) Merit-based Incentive Payment System (MIPS) was created to reward high-value, high-quality Medicare clinicians with payment increases, while reducing payments to clinicians who fail to meet performance thresholds. MIPS measures are also regularly included in CMS Alternative Payment Models.

3. Measure Classification & Designation

Measures may be classified in many ways, including measure type, data source, care setting, conditions, or level of analysis.

3.1 Measure Types

Individual quality measures reveal aspects of quality care that, taken together, provide a more comprehensive picture of healthcare quality.[2] Definitions in **Table 1** are based on CMS' Measures Management System (MMS) and examples are sources from the MIPS program.[3]

Table 1: Measure Types

Measure Type	Definition	Example
Process	A process measure focuses on steps that should be followed to prove good quality care. There should be a scientific basis for believing that the process, when executed well, will increase the probability of achieving a desired outcome. Most healthcare quality measures currently used for public reporting are process measures.	Percentage of female patients aged 18 to 70 with stage I (T1c)-III HER2 positive breast cancer for whom appropriate treatment is initiated.
Outcome	An outcome measure focuses on the health status or change in health status of a patient – either desirable or adverse – resulting from healthcare. Outcome measurement may be supplemented by risk adjustment, a statistical method for controlling factors that are known to influence the relationship between the predictor and outcome (e.g., controlling for hospital resources or patient conditions present upon arrival in a measure of surgical site infection). Stratification is another tactic, which may help identify specific patient populations for whom measure performance varies based on influence from external factors.	Percentage of patients aged 18 years and older who had a surgical site infection (SSI).
Intermediate Outcome	An intermediate outcome measure assesses the change produced by a healthcare intervention that leads to a long-term outcome.	The percentage of current smokers who abstain from cigarettes prior to anesthesia on the day of elective surgery or procedure.
Patient-Reported Outcome-Based Performance Measure (PRO-PM)	A patient-reported outcome is the status of a patient's health condition or behavior reported directly by the patient. A patient-reported outcome-based performance measure (PRO-PM) is based on patient-reported outcome measure (PROM) data aggregated for an accountable healthcare entity. These data are collected directly from the patient using a PROM tool, which may be an instrument, scale, or single-item measure.	Percentage of cancer patients currently receiving chemotherapy or radiation therapy who report significant pain improvement (high to moderate, moderate to low, or high to low) within 30 days.
Structure	A structure measure assesses features of a healthcare organization or clinic relevant to its capacity to provide good quality healthcare, such as the environment of care or administrative processes and policies. Accreditation and certification programs commonly use structure measures to assess compliance with accreditation	Percentage of patients, regardless of age, with a current diagnosis of melanoma or a history of melanoma whose information was entered, at least once within a 12-month period, into a recall system that

	standards or certification requirements. Structure measures typically rely on organizational information rather than patient-level data.	includes: A target date for the next complete physical skin exam, AND A process to follow up with patients who either did not make an appointment within the specified timeframe or who missed a scheduled appointment.
Efficiency	An efficiency measure assesses the cost of care (i.e., inputs to the health system in the form of expenditures or other resources) associated with a specified level of health outcome.	Percentage of Stage IV colon/rectal cancer patients receiving any white cell growth factors with chemotherapy
Composite	A composite measure contains two or more individual measures or quality actions but results in a single performance score.	Severe Sepsis and Septic Shock: Management Bundle - Consistent with Surviving Sepsis Campaign guidelines, it assesses measurement of lactate, obtaining blood cultures, administering broad spectrum antibiotics, fluid resuscitation, vasopressor administration, reassessment of volume status and tissue perfusion, and repeat lactate measurement. The first three interventions should occur within three hours of presentation of severe sepsis, while the remaining interventions are expected to occur within six hours of presentation of septic shock.
Paired	Paired measures are two or more individual measures that are endorsed for use together as a unit, while still resulting in individual measure scores.	Percentage of patient visits, regardless of patient age, with a diagnosis of cancer currently receiving chemotherapy or radiation therapy in which pain intensity is quantified. Percentage of visits for patients, regardless of age, with a diagnosis of cancer currently receiving chemotherapy or radiation therapy who report having pain with a documented plan of care to address pain.

3.2 Data Source

The data source is the origin of the data obtained for measurement. Measures rely on several types of data sources, each of which has an impact on the scope, purpose, and generalizability of the measures using the data. Data source refers to the type of data used to calculate the measure and consideration must be given to the data source to determine how the measure will be specified. Several data sources in the CMS Measures Management System (MMS) include:

- **Administrative Data:** Includes demographic information about the patient. Non-patient data, such as staffing information (i.e., payroll) or organizational policies, may also be included.
- **Claims Data:** Includes healthcare reimbursement or payment information which comes from claims or from the provider's billing system. Includes admission and discharge dates, diagnoses, procedures, and source of care. Claims data is utilized by MIPS in two different measure types:
    - Administrative Claims Measure: This measure type utilizes only data naturally found in medical and/or pharmacy claims without the need for providers to include specific quality codes.
    - Part B Claims Measure: This measure type utilizes specific quality codes, reported by physicians and other providers, to represent measure outcomes. This measure type is currently exclusive to small practices.
- **Registry:** Includes collection of clinical data for assessing clinical performance quality of care. May be regional or national level across multiple clinicians and institutions and is typically standardized.
- **Paper Patient Medical Records:** Includes data from the clinical laboratory, imaging services, health records, and pharmacy.
- **Electronic Patient Medical Record:** Includes digital sources for data from the clinical laboratory, imaging services, health records, and pharmacy.
- **Electronic Clinical Record:** Includes patient-level information that can be extracted in a format that can be used in a measure, such as data from personal health devices, which may be uploaded to the electronic health record (EHR).
- **Standardized Patient Assessment**: Includes data collected from validated instruments and question sets. This data collection must be validated and tested.
- **Patient-Reported Data and Survey:** Includes data collected via survey, questionnaire, or instrument. Patient/family or caregiver-completed surveys provide a person's perspective on individual experiences and feelings such as symptom management, quality of life, and functional status.

In recent years, measurement programs and developers have prioritized the use of electronic data in quality measurement, either through claims data, electronic health records, or a registry. Understanding what types of data are available from the intended data source, as well as how those data are captured, is essential to developing quality measures that can be feasibly implemented. A measure developer must also consider whether the necessary data elements could be seamlessly captured within a clinical workflow and in the routine course of care.

3.3 Care Setting

The care setting is the place in which the clinical action or outcome of interest occurs, and where the measure is applied, and performance is assessed. Care settings may include ambulatory care, inpatient/hospital, clinician offices, home care, hospice, behavioral health, or emergency departments. The care setting must be established early in the development process to determine whether data elements are feasibly captured within the chosen setting.

3.4 Level of Analysis

The level of analysis is the level at which the measurement is assessed, and determines the clinician, group of clinicians, or entity whose performance is being measured. Measures may be assessed at various levels of analysis, including:

- Clinician (individual level)
- Clinician (group/practice level)
- Facility (e.g., hospital, nursing home, home health agency)
- Health Plan (i.e., insurer for an enrolled population)
- Integrated delivery system
- Population (i.e., community, county/city, regional or state, national)

Many quality measures are intended to assess the performance of individual providers, while other measures address the performance of a facility, health plan, or health system. It is important to align the level of measurement with the appropriate care setting.

4. Measure Specifications

Measures comprise components required to calculate the measure and evaluate performance, as detailed in **Table 2.**[3] Measures are expressed as a fraction and include numerator and denominator statements and any applicable exclusions or exception.

Table 2: Measure Specification Components

Measure Component	Definition	Example
Numerator	The numerator statement describes the target process, condition, event, or outcome. It describes the action that satisfies the conditions of the measure. The numerator is also often called the "measure focus."	Patient visits in which pain intensity is quantified.
Denominator	The denominator statement describes the population evaluated by the individual measure. It may include parameters such as age range, setting, diagnosis, procedures, and/or time interval.	All female breast cancer patients aged 18 to 70 with stage I (T1c) – III HER2 positive breast cancer.
Denominator Exclusion	Denominator exclusions remove certain patients or cases from the denominator before calculating the numerator. A denominator exclusion means the	Women who had a bilateral mastectomy or who have a history of a bilateral

	numerator event is not applicable to patients addressed by the denominator exclusion.	mastectomy or for whom there is evidence of a right and a left unilateral mastectomy (for a measure on breast cancer screening mammography).
Denominator Exception	A denominator exception removes a patient, procedure, or unit of measurement from the denominator of the performance rate only if the numerator criteria are not met. Exceptions allow for the measured entity to avoid penalty if there is an appropriate reason the numerator action is not met.	Reason for not administering adjuvant treatment course including both chemotherapy and HER2-targeted therapy (e.g., poor performance status (ECOG 3-4; Karnofsky ≤50), cardiac contraindications, insufficient renal function, insufficient hepatic function, other active or secondary cancer diagnoses, other medical contraindications, patients who died during initial treatment course or transferred during or after initial treatment course.
Numerator Exclusion	Numerator exclusion defines elements excluded from the numerator data.	Infections caused by specific bacterium that should not be included in the total number of central line bloodstream infections per 1,000 catheter days.

4.1 Narrative Specifications

The measure narrative refers to the narrative description of the measure specifications, including the description, numerator, denominator, denominator exceptions, denominator exclusions, and other vital components and information about the measure, such as the clinical rationale and supporting evidence. Please refer to **Appendix III** for an example of a measure narrative.

4.2 Technical Specifications

Measure specification development is an iterative process and continues to evolve throughout the measurement lifecycle. Narrative specifications inform the development of the logic and technical specifications, which provide instructions for building and calculating a measure, adding increasing amounts of detail, including precisely defined data elements and attributes, timing parameters and intervals, value sets, and measure logic. These components ensure data collection and implementation are consistent, reliable, and effective. Measure specifications require explicitly defined data elements with accompanying analysis to identify constraints and criteria for the measure. Measure specifications often become more detailed and precise after testing.

During technical specification development, the measure developer may consider the following process, as outlined on the CMS' Measures Management System (MMS) Blueprint[4]:

- Considers the data elements necessary for the proposed measure and conducts preliminary feasibility assessments.
- May request preliminary input from standards subject matter experts (SMEs) regarding data model, terminology, data elements and content, Clinical Quality Language (CQL) expression, and impact on clinician workflow.
- Drafts initial specifications, which the TEP and other stakeholders, such as work groups, SMEs, and other measure developers, will review, and may suggest changes. Technical specifications at this stage likely include high-level numerator and denominator statements and initial information on potential denominator and numerator exclusions, if applicable.
- Continues to detail specifications and refine them throughout the development process.

For measures based on electronic, administrative, or claims-based data, the measure developer may provide draft technical specifications to test sites to evaluate feasibility and father feedback. For measures based on chart abstraction, the measure developer creates and tests data collection tools. Please refer to **Appendix IV** for examples of tools and resources available to assist measure developers in producing measure specifications.

4.2.1 Codes, Coding Systems, and Value Sets

Measures rely on the use of various standardized codes or code systems for classifying provided care. Code system concepts are used to represent clinical information. All codes and the code system and version are vital to the accuracy of the measure, and the code source and instructions for their use must be explicitly stated.

Value sets are lists of specific codes and corresponding terms from one or more code systems (e.g., ICD-10-CM, CPT, SNOMED-CT, RxNorm, LOINC) that define clinical concepts and represent data elements used within a measure. Value set development entails gathering input from SMEs who are knowledgeable with the clinical and/or administrative need, as well as terminology experts who are familiar with the code systems used. The National Library of Medicine's Value Set Authority Center (VSAC) is a value set authoring and repository tool. The VSAC allows users to search existing value sets, create new value sets, and maintain value set content consistent with current versions of the code systems they use. Use and access of the VSAC requires a free Unified Medical Language System (UMLS) license.

4.2.2 Data Protocol

A data protocol explicitly defines the types of data and how to aggregate or link these data so that the measure calculation can be reliable and valid. A data protocol must define key terms, data elements, codes, and code systems; describe the level of measurement and analysis; detail the sampling; determine risk adjustment; establish time intervals; describe how the measure results are scored and reported; and detail the calculation algorithm.

4.2.3 Calculation Algorithm

The calculation algorithm, also referred to as performance calculation, measure logic, or measure flow, is a depiction of the path from the raw data to the result of a measure's performance score. The calculation algorithm must be consistent with the measure text, as the calculation algorithm will serve as the basis for development of computer programming to produce the measure results. The calculation

algorithm should ensure there is a logical end point for all applicable clinical scenarios. Alpha testing and preliminary feasibility assessments assist in testing each scenario.

5. Measure Development Lifecycle

The completion of measure development yields a precisely specified, valid, reliable, and clinically significant measure that will be widely used to provide value in oncology. Although this manual describes the phases of the measure lifecycle in a linear, sequential fashion, measure developers have flexibility to adjust the sequence or carry out steps concurrently and iteratively.

5.1 Measure Conceptualization

Measure conceptualization refers to the initial phase in the measure development process. The key components of measure conceptualization are information gathering, business case development, and assessment of measure needs.

The measure conceptualization phase begins by identifying a measure concept and considering whether it meets the characteristics of a meaningful measure to improve the quality of patient care and positively affect patient outcomes (**Table 3**).

The development of any clinical quality measure may not be indicated if a measurement topic does not meet all the required characteristics. While ASCO conducts measure development and maintenance according to strict standards and requirements set by CMS and PQM, measures used solely to support internal ASCO programs may be developed with a modified methodology or may not meet all the considerations detailed in this section.

Table 3: Characteristics of a meaningful clinical quality measure

These are the required characteristics that must be in place prior to beginning work on any proposed measure development concept	
Evidence Base	One or more ASCO or other evidence-based clinical practice guidelines, standards, and/or systematic reviews of existing evidence. Guideline recommendations, standards, and/or systematic reviews may not be available to directly support an outcome that is not amenable to research and high-level evidence; measure developers may need to rely on other types of evidence, including expert consensus.
Performance Gaps and Disparities in Care	Documented evidence of deviation (or observed patterns of deviation) from clinically recommended care. Gaps in care may be manifested by the inappropriate use of health services (i.e., underutilization or overutilization of health services) across providers and/or disparities in healthcare across patient populations.
High Impact	Clinical condition with high prevalence, a significant burden of illness, high cost, or a nationally identified clinical priority area is addressed (e.g., CMS, National Academy of Medicine, National Priority Partners)
Measure Gap	Absence of an existing measure that evaluates the same concept or is otherwise, duplicative.

5.2 Information Gathering & Business Case Development

The creation of a business case for the development of de novo measures is a vital step to assess the anticipated benefits of a new measure against the resources and costs required for development, testing and implementation. ASCO's measure development team collates the business case information during the information gathering phase, which supports our measure prioritization process. This information is developed in conjunction with the MSG and provided to the CQIC to determine if the measure is a strategic fit in ASCO's measure library, its value to the public, and its alignment with ASCO's strategic plan and the needs of ASCO's quality programs, e.g., ASCO Certified. The MSG will also determine if the healthcare system has the capacity to respond to the quality action defined by the measure and its affordability and achievability, with regard to quality improvement and measurement.

Information gathering includes developing a broad-based strategy that includes an environmental scan (e.g., review of the literature, search for clinical practice guidelines and existing measures), review of the regulatory and economic environments, and stakeholder needs[4]. A strong, comprehensive information gathering strategy will improve the likelihood of the success of a quality measure. Measure developers conduct information gathering by completing an environmental scan of existing measures, as well as executing a comprehensive literature review (white and grey) and searching for relevant recommendations among published clinical practice guidelines (**Appendix V**). Information gathering may also include a review of legislation and regulations and their implications on measurement (e.g., the Medicare Access and CHIP Reauthorization Act of 2015, MACRA), conducting empirical data analyses, and collecting expert and stakeholder input (such as the TEP or other experts, and all relevant stakeholders – including patients).

5.3 Measure Harmonization

Differences in measure specifications limit comparability across settings. Multiple measures with the same clinical focus and target population create burden and confusion in choosing measures to implement and when interpreting and comparing the measure results. Measure developers are expected to consider harmonization as one of the core measure evaluation criteria that are applied throughout the measure lifecycle. PQM also requires consideration of measure harmonization with related measures as part of its endorsement processes.

Measure harmonization is defined as standardizing specifications for related measures when they have the same measure focus (i.e., numerator criteria); target population (i.e., denominator criteria); or when components apply to many measures (e.g., age designation for children), as detailed in **Table 4**. Harmonized measure specifications are standardized unless differences are supported by the evidence.

Table 4: Measure Harmonization Scenarios

Scenario	Harmonization Issue	Resolution

Numerator: Same measure focus Denominator: Same target population	Competing measures	• Use existing measure or justify development of an additional measure ○ Note: different data sources require new specifications that are harmonized (e.g., respecified)
Numerator: Same measure focus	Related measures	• Harmonize specifications to unify measures' focus • Justify differences • Respecify existing measure by expanding the target population
Numerator: Different measure focus Denominator: Same target population	Related measures	• Harmonize on the target population • Justify differences
Numerator: Different measure focus Denominator: Different target population	Not competing or related	• No resolution needed, proceed with measure development

5.4 Measure Prioritization

ASCO strives to offer a comprehensive portfolio of meaningful oncology measures to meet the needs of its members, internal ASCO programs, and the wider clinical oncology community. ASCO conducts an annual solicitation for potential new measure topics, which are evaluated through a measure prioritization process involving measure development staff and MSG members.

5.4.1 Measure Development Staff Evaluation

ASCO staff shall issue an annual call for measure concepts from ASCO members, relevant ASCO committees, steering groups, and task forces. This call for measures may be targeted to certain cancers, measure types, or domains, and may change each year to reflect the evolving needs of ASCO programs and measure stakeholders. The ASCO Measure Topic Submission Form is available during topic solicitation for ASCO members and external stakeholders to submit measure concepts for development consideration. Measure concepts received for consideration are thoroughly evaluated and scored by ASCO measure development staff according to established criteria, such as evidence, feasibility or implementability, performance gap or variation in care, applicability or use, and strategic alignment (**Appendix VI**). ASCO staff also perform an environmental scan to identify related or competing measures and will work with those external measure stewards to harmonize measures where possible, avoiding duplicative measure development work. Staff shall review scores with MSG leadership, who may accept or modify staff findings and concept scores prior to concept prioritization. Measures that fail to meet minimum requirements may be removed from consideration, as confirmed by MSG leadership.

5.4.2 MSG Measure Concept Prioritization

Measure concepts that successfully progress through the evaluation process are presented to the MSG annually, and as needed, for consideration. MSG members shall review the measure findings and scores,

and rank-order measure concepts for review and approval by the CQIC. Measures undergoing maintenance that require reconvening a TEP shall be included in the prioritization process.

5.5 Expert & Stakeholder Input

Ensuring stakeholder input is critical throughout the measure development process and is typically accomplished through convening a TEP to guide the development of the measure(s) and holding an open comment period (**Section 5.6**) to invite additional input and diverse stakeholder perspectives.

Technical Expert Panel composition is approved by the MSG leadership, and ASCO staff endeavor to balance TEP membership according to clinical expertise, career stage, and demographic factors, such as practice setting, location, and provider gender. In addition to its clinician membership, the composition of each TEP may vary and require additional expertise, such as payers, EHR vendor representatives, or practice administrators. Additionally, patient, family, and/or caregiver perspectives on the TEP are vital and ASCO promotes person and family engagement in the measure development process. Prospective TEP members receive an invitation to join the TEP, along with the TEP Responsibilities and Authorities document (**Appendix VII**), describing the roles of the chair, TEP member, and ASCO staff.

Measure TEPs are assembled in accordance with **ASCO's Conflict of Interest Policy Implementation for Quality Measures** and the **CMSS Code for Interactions with Companies**. ASCO requires disclosure by individuals involved in drafting, reviewing, and approving measures, and sets limits on the financial relationships that panel members and reviewers can have with Companies that could reasonably be affected by care delivered in accordance with a measure. To carry out this policy, potential panel members must complete a conflict-of-interest (COI) disclosure form prior to a formal invitation to serve on the panel.

5.6 Open Comment

Open comment allows for key stakeholders to critically review and communicate feedback, implementation barriers, unintended consequences, and identify potential errors or gaps in a measure. ASCO holds open comment periods both prior to the finalization and implementation of a newly developed measure, and prior to measure maintenance to gather post-hoc feedback. Open comment periods allow greater transparency in the ASCO measure development process, and adhere to best practices for measure development, enabling engagement with a wide range of interested stakeholders (especially patients and advocacy groups), and facilitating implementation and dissemination efforts.

ASCO measures are available for open comment for a two- to three-week period. Reviewers will submit their comments through a survey form. Prior to viewing the specifications, reviewers must electronically sign a non-disclosure and confidentiality agreement and must identify themselves by name and affiliation; anonymous comments will not be accepted. Measure development staff will review, summarize, and bring relevant comments to the TEP chair, and to the entire panel if necessary. Any changes made from the open comment process will be reviewed by the entire panel prior to MSG approval of the measure. Comments received are advisory in nature only and ASCO is not bound to make

any changes based on comments received. ASCO does not respond directly to open comment participants or post public responses to comments.

5.7 Measure Testing

Testing refers to all the data collection and analysis activities that contribute to the evaluation of the measure specifications. Testing assesses the suitability of the technical specifications and acquires the empirical evidence to help assess the strengths and challenges of the measure with respect to the performance evaluation criteria, especially scientific acceptability (reliability and validity) and feasibility. Testing also provides the opportunity to support the measure's importance (e.g., illustrate variation in current practice) and usability.

5.7.1 Face Validity

Validity is the degree to which a quality measure assesses what it claims to measure. Face validity is a specific type of validity that refers to how a measure appears on the surface: Does it seem to ask all the needed questions? Does it use the appropriate language and definitions to do so? Face validity does not rely on empirical evidence for support. Instead, SMEs and other stakeholders provide their input on whether a measure captures what it intends to capture and can be used to distinguish between good vs. poor quality care. SMEs and stakeholders provide feedback on the measure through open comment, targeted surveys, and group interviews.

5.7.2 Data Element Validity

Data element validity testing establishes that the measure's data elements are valid by analyzing the agreement between the electronically captured measure data and data collected by a human abstractor. The validity of the measure's data elements is then determined by calculating the degree of agreement between electronically and manually abstracted data, resulting in percent agreement, Kappa (chance-adjusted agreement), sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) data points.

Performance score validity testing demonstrates that the measure score correctly reflects the quality of care provided by satisfactorily identifying differences in quality.

Performance score validity is assessed through the following methods:

- Correlation of the measure to other, conceptually related measure(s). The merits of the hypothesized relationship, the method used to assess the relationship, and the assessment results must all be thoroughly evaluated to create a persuasive argument for the validity of the measure being tested.
- Correlation of the measure with another, comparable quality indicator(s) for the topic under investigation, such as outcomes. The merits of the hypothesized relationship, the method used to assess the relationship, and the assessment results must all be thoroughly evaluated to create a persuasive argument for the validity of the measure being tested.

Testing hypotheses that the measure scores indicate the quality of care (e.g., measure scores are different for groups known to have differences in quality from disseminated research).

5.7.3 Feasibility

Feasibility testing analyzes the extent to which the measures' technical specifications, including measure logic and required data elements, are readily available for extraction from an organization's electronic

medical record software. Feasibility testing also determines whether the measure's computed scores are comparable across different organizations. Once feasibility testing is performed, it will pinpoint specific causes of variability. Additionally, comprehensive feasibility testing will assess the measure's impact on clinical workflows and if a measure can be captured without undue burden to the staff at the organization being measured, which is critical to the measure's overall evaluation.

5.7.4 Reliability

Data element reliability testing assesses the reproducibility of the measure's data elements, testing to see whether data elements produce the same result a high proportion of the time when applied to similar populations during the same time. Data element reliability testing is not required for most measures if data element validity has already been demonstrated. Therefore, data element reliability testing is rarely performed.

Performance score reliability testing determines the precision of the measure. Complete performance score reliability testing will demonstrate the measure's ability to distinguish differences between providers due to quality of care rather than chance. Signal-to-noise ratio analysis is commonly used to assess a measure's performance score reliability. The signal is the proportion of variability that can be explained by actual differences in performance. Noise is related to the total variability in measured performance usually due to chance or attributable to measurement errors. Comparison between the signal and the noise will estimate the reliability of a measure. Other methods, such as point estimates and confidence intervals, can also demonstrate reliability if shown for all providers.

5.8 CBE Measure Endorsement

The Partnership for Quality Measurement, staffed by experts in health care quality improvement and supported by Battelle, uses a consensus-based process to ensure informed and thoughtful endorsement reviews of candidate measures. PQM's consensus-based process engages a variety of experts, including clinicians, patients, measure experts, and health information technology specialists.

Quality measures submitted by measure developers to PQM for endorsement and re-endorsement considerations are evaluated based on the following characteristics: importance, feasibility, reliability, validity, risk adjustment, equity, and use and usability. PQM endorsement and re-endorsement of quality measures is not currently required by the federal government (i.e., CMS), nor by many private sector entities for quality measures utilized in programs, however, there is often a preference for endorsed measures, given the rigorous and consensus driven PQM measure evaluation process.

5.9 Measure Maintenance

To ensure viability for use, ASCO measures are evaluated regularly and updated as needed to reflect current evidence, guidelines, and standards.

5.9.1 Annual Maintenance

Measures used in federal quality reporting programs are required to undergo an annual update. This process ensures updates to clinical terminologies such as ICD-10-CM, CPT, SNOMED-CT and LOINC are reflected in the measure specifications. This is also when updates such as changes to the narrative specifications (per changes in supporting evidence and gaps in care) and subsequent technical specifications, or changes to support measure harmonization can be made. If substantive updates to the measure are required, for example, those that affect the measure's original concept or logic, the

measure will have to repeat the rulemaking process for continued inclusion in the federal program. In many cases measure specifications remain the same throughout annual maintenance reviews. All ASCO measures included in a federal program shall be reviewed annually by ASCO staff. Minor updates will be proposed and approved by the MSG, while major updates may necessitate convening a TEP as needed to comprehensively evaluate the measure. As a result of annual maintenance, staff will present any changes to the evidence, gaps, technical and narrative specifications, and testing results to the TEP for their review and approval.

5.9.2 Ad Hoc Maintenance

An ad hoc review is a formal measure evaluation conducted outside of the scheduled annual maintenance process. An ad hoc review is triggered by a material change in a measure's clinical evidence base which could necessitate modification of the measure specifications, and which significantly affects the measure performance, such as:

- Changes to the population being measured (e.g., changes in age inclusions, diagnoses, or other inclusion criteria, or excluded populations),
- Changes to what is being measured (e.g., changes in target values like blood pressure or lipid values),
- Inclusion of new data source(s), or
- Expansion of the level of analysis or care setting.

5.9.3 Maintenance of PQM Endorsement

ASCO measures that are PQM endorsed must undergo a full re-evaluation every three years to ensure currency, relevancy, and maintain endorsement. Full maintenance includes assessment of the measure importance, scientific acceptability, and measure use and usefulness including impact and unintended consequences. The PQM measure endorsement process is further detailed in **section 5.8**.

6. Challenges & Future Measurement Trends

Challenges in the development of oncology quality measures persist, driven by a lack of data standardization which impacts the feasibility of capturing important health equity insights, and hinders advancement of digital measures.

6.1 Data Standardization

A lack of data standardization and inconsistent applications of existing standards within EHRs hinder stakeholder ability to systematically capture important data elements and measure certain aspects of care. Oncology data is granular and complex, and data such as cancer staging, biomarker testing and results, and disease progression are often not captured in EHRs, but reside in unstructured documents such as clinical notes.[6, 7]

The Minimal Common Data Elements (mCODE) initiative aims to improve this issue by establishing a set of basic oncology-related data elements that would be available in all EHRs. mCODE is a collaboration between MITRE, NCI, ASCO, and others. ASCO's Measures Team has been working on a pilot with several other stakeholders to develop a use case for quality measures using FHIR and mCODE.

6.2 Health Equity

There are promising new efforts being made to standardize medical codes related to social drivers of health, which are defined by the World Health Organization as the conditions in which people are born, live, work, and age. New in 2023 is a CMS-stewarded measure titled "Screening for Social Drivers of Health," which measures the percent of adults screened for food insecurity, housing instability, transportation needs, utility difficulties, and interpersonal safety. In addition, HL7's Gravity Project is a public collaborative striving to advance health and social data standardization to achieve health equity. They are seeking to standardize codes and facilitate the use of these data in patient care, care coordination between the health and human services sectors, population health management, value-based payment, and clinical research.

This emphasis on health equity will continue to apply to measures; for example, at the federal level, CMS has updated its meaningful measures framework to include a focus on equity-related measures. The recently announced Enhancing Oncology Model, which is housed within the Innovation Center at CMS, states that it will focus on value-based, patient-centered care for cancer patients undergoing chemotherapy with a specific focus on health equity. The MAP Rural Health Advisory group provides input on issues typically encountered by rural residents and clinicians, when reviewing measures under consideration for use. The MAP Health Equity Advisory Group, created during the 2021-2022 cycle, provides input on measures from a health equity and disparities lens, with the overarching goal of reducing health differences related to SDOH.

At present, in addition to the lack of SDOH-related data standardization, many patients decline to answer questions related to SDOH as there is an underlying belief that doing so will result in negative consequences.[8] Efforts such as The Gravity Project, NCQA SDOH Resource Guide, and AMA STEPS Forward, however, are in place to guide clinicians in building trust to increase patient comfort in providing SDOH-related data needed to ensure they are receiving the best healthcare.

6.3 Digital Measures

CMS has set the goal of transitioning all clinical quality measures used in its report programs to digital quality measures (dQMs) over the next several years.[9] CMS defines dQMs as, "quality measures, organized as self-contained measure specifications and code packages, that use one or more sources of health information that is captured and can be transmitted electronically via interoperable systems. Data sources for dQMs may include administrative systems, electronically submitted clinical assessment data, case management systems, electronic health records, laboratory systems, prescription drug monitoring programs, instruments (for example, medical devices and wearable devices), patient portals or applications (for example, for collection of patient-generated data such as a home blood pressure monitor, or patient-reported health data), health information exchanges, or registries, and other sources".[10] When developing measures, claims data often capture the target population via ICD-10-CMdiagnosis codes and CPT procedure codes, but the granularity needed to evaluate the clinical action or outcome of interest is found in the EHR or registry. Capturing data from multiple sources allows for a more granular measure and accurate depiction of cancer care. Additional benefits of digital quality measures include more accurate attribution and advancing team-based measurement, as the additional data sources provide a better picture of all providers involved in treating a cancer patient.

Appendix I: Measures Steering Group Roles and Authorities

AMERICAN SOCIETY OF CLINICAL ONCOLOGY
STEERING GROUP DESCRIPTION

GROUP:	**Measures Steering Group**
REPORTS TO:	**Care and Quality Improvement Committee**
DEPARTMENT:	**Care Delivery**
DEPARTMENT STAFF:	**Measure Development Team**

Purpose

The Measures Steering Group ("the Steering Group") exists to oversee quality measurement prioritization, development, calculation methodology, testing, measure maintenance, and approval of measures for inclusion in the ASCO Measures Library. The ASCO Measures Library is a repository of developed and maintained measures available for stakeholder use, including programs such as ASCO Certified, International Quality, and the QOPI Certification Program (QCP), as well as federal and private programs. The Steering Group also oversees any collaborative measure development projects and the review of externally developed measures.

The Steering Group reports to the Care and Quality Improvement Committee (CQIC). By addressing critical clinical gaps in care, supporting evidence-based medicine, promoting coordinated care, and encouraging the reduction of healthcare disparities, the Steering Group hopes to enhance the quality, effectiveness, and appropriateness of healthcare services throughout the cancer care continuum.

Composition and Appointment Process

The Measures Steering Group comprises ASCO members in good standing, representing a diversity of practice types, settings, and sizes. The Steering Group may include members with diverse expertise in medical, surgical, or radiation oncology, hematology, pharmacology, biostatistics, bioinformatics, quality of life, supportive care, survivorship, or other related fields. Steering Group members may serve as liaisons to other ASCO volunteer groups, as needed.

The Measures Steering Group will convene Technical Expert Panels (TEPs) to address specific measurement activities and are formed at the discretion of the Steering Group Chair and in consultation with the CQIC.

Steering Group Member Terms

The Measures Steering Group members shall serve a three-year term. Members can serve additional terms as determined by the ASCO Board.

Steering Group Chair Term

The Measures Steering Group Chair shall serve one-year consecutive terms as Chair-elect, Chair, and Immediate Past Chair. The Chair, Chair-Elect and Immediate Past Chair will be appointed to the CQIC as members and serve simultaneously on the CQIC during his/her term on the Steering Group.

Steering Group Responsibilities and Authorities

- Oversee ASCO measure development and maintenance, including prioritization, specification, calculation methodology, and testing.
- Follow Board-approved procedures for review and approval of measures and other related projects as appropriate.
- Review and prioritize measure concepts or domains annually for de novo development or maintenance.
- Solicit collaborators (e.g., other societies) for measure development, as appropriate.
- Consider requests for collaboration and/or nominate ASCO representatives to measure panels convened by other organizations.
- Review externally developed, cancer-relevant measures and prepare comments, as necessary.
- Perform regular measure maintenance.
- Address external questions related to measure intent.

Steering Group Chair Responsibilities and Authorities

- Attend MSG leadership calls, MSG meetings, and CQIC meetings.
- Participate as a member of the CQIC.
- Solicit and discuss measure priorities with leadership of the QCP Steering Group, ASCO Certification Program Steering Group, Health Equity and Outcomes Committee, Evidence Based Medicine Committee, International Quality Steering Group, and Care and Quality Improvement Committee.
- Disclose outside relationships as requested and comply with applicable ASCO conflicts of interest policies.
- Oversee the delegation of responsibility for measure development, and other related projects as appropriate, to TEPs.
- In consultation with the Chair-Elect and Immediate Past Chair, approve composition of TEPs charged with developing measures and other related projects, as appropriate.
- In consultation with the Chair-Elect and Immediate Past Chair, identify and approve ASCO representatives appointed to the measure panels of other organizations or appointments for other similar initiatives.
- Identify and promote new volunteer leadership within the Measures Steering Group.
- Represent ASCO at professional society meetings.
- Provide regular updates to the CQIC.

Steering Group Chair-Elect and Past-Chair Responsibilities and Authorities

- Attend MSG leadership calls, MSG meetings, and CQIC meetings.
- Participate as a member of the CQIC.
- Disclose outside relationships as requested and comply with applicable ASCO conflicts of interest policies.
- In Chair's absence, serve as Chair at Steering Group meetings.
- Assist the Chair in carrying out the mission and the objectives of the Steering Group.

- With the Chair, approve composition of TEPs charged with developing measures and other related projects, as appropriate.

Steering Group Member Responsibilities and Authorities

- Disclose outside relationships as requested and comply with applicable ASCO conflicts of interest policies.
- Volunteer to lead measure development TEPs that correspond to clinical expertise and interest.
- Participate in assigned workgroup and panel calls/meetings.

Steering Group Staff Responsibilities and Authorities

- Monitor relevant policy and policy-influencing organizations.
- Contribute to preparation of measure concept prioritization.
- Lead the drafting of measure narrative, calculation methodology, testing, and authoring (e-specifying).
- Maintain accurate records of the ASCO Measures Library and measure concept prioritization pipeline.
- Coordinate and lead measure testing projects.
- Coordinate and lead measure maintenance projects.
- Coordinate and lead measure authorship utilizing current standards (i.e., FHIR).
- Conduct legal reviews and prepare legal documents, as required.
- Manage vendor relationships, as relevant.
- Prepare presentations, reports, and manuscripts, as needed.
- Oversee day-to-day implementation and coordinate meetings, conference calls and follow-up activities.
- Conduct outreach to other professional societies on workgroup-related issues and respond to requests for partnership.
- Coordinate with Care Delivery and Guidelines staff on the development and maintenance of measures, as needed.
- Disclose outside relationships as requested and comply with applicable ASCO conflicts of interest policies.

Meetings Calendar

The Measures Steering Group shall meet at least one to two times per year, either in-person at ASCO Headquarters in Alexandria, VA, virtually, or as a hybrid meeting.

Appendix II: General and Technical Principles for Measure Development

General Measure Development Principles

ASCO's measures are developed in accordance with the following principles. ASCO measures are:

- Independently developed through a transparent process.
- Evidence-based and derived from published guidelines where a guideline is available; and address a performance gap where there is known variation in performance.
- Continually monitored for unintended consequences of measure implementation, including overuse and underuse of care.
- Routinely reviewed and updated to reflect changes in evidence or practice, as applicable.

ASCO measure development:

- Strives to reduce clinician reporting burden by operationalizing eCQMs or dQMs whenever possible.
- Focuses on outcomes, safety, patient experience, care coordination, appropriate use/efficiency, and cost of care.
- Aligns and harmonizes similar measures across data sources to the greatest extent possible.
- Aligns and harmonizes measures across payers, including Medicare, other federal partners, and private payers, to the greatest extent possible.
- Follows regulations for patient privacy and human research protection in development and validation of measures.
- Strives to focus on what is most meaningful to patients, caregivers, and providers (e.g., including patient or caregiver representatives on measure development TEPs).
- Engages appropriate stakeholders early and often in the measure development process.
- Works to reduce disparities in cancer care, which disproportionately affect underserved and vulnerable populations, and advocates for measures that support the delivery of high-quality, equitable care.

Technical Principles for Measure Development

As defined by the Blueprint for Measure Development,[4] the following principles should be applied when developing measures for consideration for quality reporting and value-based purchasing programs:

- Develop a rigorous business case for an evidence-based measure concept.
- Prioritize electronic clinical data sources (e.g., electronic health records [EHRs] and registries), where appropriate, and reduce dependency on data from chart abstraction whenever possible.
- Maintain a focus on iterative testing using both real and synthetic data.
- Consider approaches to aggregate multiple data sources (e.g., hybrid measures) to achieve the most accurate assessment of quality until universal interoperability can be achieved.
- Define outcomes, risk factors, cohorts, and inclusion/exclusion criteria based on clinical and empirical evidence.
- Judiciously select exclusions to capture as broad a patient population as possible and appropriate; consider developing a paired measure to capture and measure the care received for the excluded patients if a significant number of patients are excluded.

- Develop risk adjustment models to distinguish performance between providers rather than predict patient outcomes.
- Include measure stratification and risk adjustment approaches to patient demographic characteristics that promote equitable quality comparisons.
- Harmonize measure methodologies, data elements, and specifications, when applicable and feasible.
- Develop each measure with sufficient statistical power to detect and report statistically significant differences in provider performance.
- Consider strategies to enable clinicians that have smaller practices and low-volume facilities to reliably report a measure.
- Strive to develop measures that can progress to multi-payer applicability using all-payer databases where available.
- Consider the clinical workflow needed in the electronic record for electronic clinical quality measures (eCQMs).

Note: While ASCO conducts measure development and maintenance according to strict standards and requirements set by CMS and PQM, measures used solely to support internal ASCO programs may be developed with a modified methodology.

Appendix III: Measure Narrative Example

Measure Title: Appropriate intervention of immune-related diarrhea and/or colitis in patients treated with immune checkpoint inhibitors	
Measure Description:	Percentage of patients, aged 18 years and older, with a diagnosis of cancer, on immune checkpoint inhibitor therapy, and grade 2 or above diarrhea and/or grade 2 or above colitis, who have immune checkpoint inhibitor therapy held and corticosteroids or immunosuppressants prescribed or administered.

Initial Population:	Patients, 18 years and older, with a diagnosis of cancer and on immune checkpoint inhibitors and who have grade 2 or above diarrhea and/or grade 2 or above colitis.
	
	<u>Initial Patient Population Guidance:</u>
	• Immune checkpoint inhibitors-class of medications that prevent tumors from "hiding" or "evading" the body's natural immune system. This is a form of cancer immunotherapy. Immune checkpoint inhibitor medications include PD-1 inhibitor drugs, PD-L1 inhibitor drugs, and CTLA-4 inhibitor drug.
	  ○ PD-1 inhibitors drugs include: Pembrolizumab, Nivolumab, Cemiplimab
	  ○ PD-L1 inhibitors drugs include: Atezolizumab, Avelumab, Durvalumab
	  ○ CTLA-4 inhibitor drug includes: Ipilimumab
	• Grade 2 Diarrhea - 4-6 bowel movements above baseline per day. Moderate increase in ostomy output compared to baseline; limiting instrumental ADL
	• Grade 3 Diarrhea - increase of >=7 stools per day over baseline; hospitalization indicated; severe increase in ostomy output compared to baseline; limiting self-care ADL
	• Grade 4 Diarrhea - Life-threatening consequences; urgent intervention indicated
	• Grade 2 Colitis - Abdominal pain, mucus or blood in stool
	• Grade 3 Colitis – Severe abdominal pain; peritoneal signs
	• Grade 4 Colitis – Life-threatening consequences; urgent intervention indicated
	*Grading for GI toxicity by Common Terminology Criteria for Adverse Events (CTCAE) v5.0

Denominator:	Equals Initial Population
Denominator Exclusions:	Patients with pre-existing inflammatory bowel disease (IBD) (e.g., ulcerative colitis, Crohn's disease).
Numerator:	Patients with immune checkpoint inhibitor therapy held and corticosteroids or immunosuppressants prescribed or administered. <u>Numerator Guidance:</u> • Immune checkpoint inhibitors should be held for patients who have grade 2 or above diarrhea and/or grade 2 or above colitis.

	• Corticosteroids examples include but are not limited to methylprednisolone, prednisone, or dexamethasone. Route of administration may be oral or intravenous dependent on agent. • Immunosuppressants include but are not limited to vedolizumab or anti-TNF agent such as infliximab. Route of administration may vary dependent on agent.
Numerator Exclusions:	None
Denominator Exceptions:	Documentation of medical reason(s) for not prescribing or administering corticosteroid or immunosuppressant treatment (e.g., allergy, intolerance, infectious etiology, pancreatic insufficiency, hyperthyroidism, prior bowel surgical interventions, celiac disease, receiving other medication, awaiting diagnostic workup results, other medical reasons/contraindication). <u>Denominator Exceptions Guidance:</u> • Diarrhea is not attributed to immune checkpoint inhibitor mucosal inflammation. Examples include but are not limited to infection, pancreatic insufficiency, hyperthyroidism, prior bowel surgical interventions, and celiac disease. • The clinician did not yet prescribe or administer corticosteroid or immunosuppressant due to awaiting diagnostic workup or results for alternative etiologies.
Stratification/ Calculation:	None
Measurement Period:	Calendar Year
Clinical Recommendations:	<u>NCCN Clinical Practice Guidelines in Oncology: Management of Immunotherapy-Related Toxicities. 2020.</u> Recommendation: For moderate diarrhea/colitis (G2), hold immunotherapy and administer prednisone/methylprednisolone (1mg/kg/day). If no improvement is noted within 2 to 3 days, increase corticosteroid dose to 2mg/kg/day and consider adding infliximab. <u>AGA Clinical Practice Update on Diagnosis and Management of Immune Checkpoint Inhibitor (ICI) Colitis and Hepatitis: Expert Review. 2020.</u> Recommendation for >= Grade 2 Colitis or Diarrhea (suspected immune-mediated): Withhold ICI therapy. Best Practice Advice (BPA) 6-ICI colitis typically responds to high dose systemic glucocorticoids, given in doses of 0.5-2 mg/kg prednisone equivalent daily with a taper of 4-6 weeks, although these doses and schedules have not been rigorously examined. Infliximab and vedolizumab are reasonable options for treatment of glucocorticoid refractory colitis. <u>Chemotherapy and Immunotherapy Guidelines and Recommendations for Practice. ONS. 2019.</u>

	Recommendation for Grade 2 Diarrhea: Hold immunotherapy. Administer IV methylprednisolone (1 mg/kg/day). If no response in 2-3 days: -increase dose to 2mg/kg/day -consider infliximab -if refractory to infliximab, consider vedolizumab Recommendation for Grade 2 Colitis: Hold checkpoint inhibitor therapy and continue treatment with antidiarrheal. If symptoms persist up to one week, it is recommended to initiate corticosteroids. American Society of Clinical Oncology Clinical Practice Guideline. Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy. Journal of Clinical Oncology. 2018. Recommendation: ICPi therapy may be suspended for most grade 2 toxicities, with consideration of resuming when symptoms revert to grade 1 or less. Corticosteroids may be administered. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. 2017. Recommendation: In grade 2 diarrhea, ICPi should be interrupted, and the patient should start with corticosteroids depending on the severity and other symptoms (either budesonide or oral corticosteroids 1 mg/kg). In the case of no improvement within 3–5 days, colonoscopy should be carried out and, in the case of colitis, infliximab 5 mg/kg should be administered.
Evidence Strength:	NCCN Clinical Practice Guidelines in Oncology: Management of Immunotherapy-Related Toxicities. 2020. • Category 2A-Based upon lower-level evidence, there is uniform NCCN consensus that the intervention is appropriate. AGA Clinical Practice Update on Diagnosis and Management of Immune Checkpoint Inhibitor (ICI) Colitis and Hepatitis: Expert Review. 2020. • No evidence strength provided. Chemotherapy and Immunotherapy Guidelines and Recommendations for Practice. ONS. 2019. • No evidence strength provided. American Society of Clinical Oncology Clinical Practice Guideline. Management of immune-related adverse events in patients treated with immune checkpoint inhibitor therapy. Journal of Clinical Oncology. 2018. • All recommendations are expert consensus based, with benefits outweighing harms, and strength of recommendations is moderate. Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. 2017.

	Levels of evidence and grades of recommendation (adapted from the Infectious Diseases Society of America-United States Public Health Service Grading Systema) • Level of evidence: IV Retrospective cohort studies or case–control studies V Studies without control group, case reports, expert opinions • Grades of recommendation: B - Strong or moderate evidence for efficacy but with a limited clinical benefit, generally recommended
Rationale:	The occurrence of diarrhea and colitis can be a normal and treatable toxicity (and is many times not immune-related), but if it is immune-related, it can become life-threatening if not addressed in a timely manner (Acharya et al 2013). Diarrhea and colitis are the second-most commonly reported AEs with ICIs, and symptoms typically develop within 6 to 8 weeks of starting treatment (NCCN Guidelines 2020). Preventing diarrhea includes early recognition of symptoms. Proper grading of diarrhea is essential for proper management. Regardless of immunotherapy agent used, effective colitis and diarrhea management is accomplished by early intervention. Colitis related mortality with immunotherapy agents has been associated with delayed reporting, nonadherence with antidiarrheal regimen, and failure to hold the immunotherapy agent. With early intervention, colitis is reversible (ONS 2019). Incidence of diarrhea is higher among patients taking combination anti-CTLA-4/anti-PD-1 therapy (44%) than those receiving anti-CTLA4 (23–33%) or anti-PD-1 (≤19%) monotherapy. The combinatorial approach is also associated with increased risk of grade 3/4 symptoms compared with monotherapy, and the proportion of patients experiencing high-grade symptoms is greater with ipilimumab than anti-PD-1 or antiPD-L1 agents. Diarrhea and/or colitis may recur months after discontinuation of immunotherapy and can mimic chronic inflammatory bowel disease (IBD) (Puzanov et al 2017). The hallmark symptom of ipilimumab-associated colitis is 3-20 loose bowel movements per day with possible associated hematochezia (ONS 2019). For patients on pembrolizumab, 16% of patients have diarrhea of any grade, while 1% of patients will have grade 3-4 diarrhea. (ONS 2019).
Opportunity for Improvement/ Performance Gap:	One study found that only 49% of health care professionals are comfortable with recognizing and managing immune related adverse events. (Schwartzberg et al. 2018) In 2017, a survey conducted by the Association of Community Cancer Centers (ACCC) reported that only 24% of respondents reported that they had a deep familiarity with checkpoint inhibitors, 32% with monoclonal antibody therapy, and only 17% with combination treatment regiments (ACCC 2018). Association of Community Cancer Centers (ACCC) (2017-2018). *Immuno-Oncology: Transforming the Delivery of Cancer Care in the Community* [White paper]. http://www.informz.net/ACCC/data/images/Attachments/2017%20IO%20White%20Paper.pdf

	Schwartzberg, L.S., & Perloff, T. (2018). *Identifying Gaps in Immunotherapy Education: Beyond the Oncology Team*. Abstract #PS26. https://www.mascc.org/assets/2018_Meeting_Files/Sat30/Strauss_3/1324_Perloff_Strauss%203_Sat.pdf
Level of Analysis:	Clinician: Group/Practice Clinician: Individual
Care Setting:	Outpatient Services
Data Source:	Registry
Type of Measure:	Intermediate Outcome
Interpretation of Score:	Better quality is associated with a higher score
Intended Use:	Accountability/Public Reporting
Testing:	Feasibility: Six sites using different EHRs and of different affiliations were evaluated in feasibility analysis. While data collected showed that the measure was mostly feasible (with only few data elements not always captured), the provider workflow will have to be modified to fully implement this measure. Overall, full measure implementation presents an average burden to the providers. Validity: This measure demonstrated high face validity with 92% of subject matter experts agreeing on the denominator, 73% of subject matter experts agreeing on denominator exclusions, 88% of subject matter experts agreeing on denominator exceptions, and 83% of subject matter experts agreeing on the numerator. Moreover, an average of 92% of subject matter experts agreed that the measure is meaningful, addresses a gap in care, will improve care, and addresses a serious ailment with dangerous consequences. Reliability: Measure performance scores on seventy-five patients across seven practices showed high reliability as indicated by an adjusted split-sample correlation coefficient of 0.8952. For more details, please reach out to measurement@asco.org
Risk Adjustment:	None
Telehealth:	Telehealth visits are appropriate to include in the denominator.
Risks to Development/ Implementation:	None
Copyright:	**QPP MIP CQMs ("Registry Measures") and CBE Copyright Language** COPYRIGHT: The Measure is not a clinical guideline, does not establish a standard of medical care, and has not been tested for all potential applications.

	The Measure, while copyrighted, can be reproduced and distributed, without modification, for noncommercial purposes, e.g., use by health care providers in connection with their practices. Commercial use is defined as the sale, license, or distribution of the Measure for commercial gain, or incorporation of the Measure into a product or service that is sold, licensed or distributed for commercial gain. Commercial uses of the Measure require a license agreement between the user and Society for Immunotherapy of Cancer (SITC) and prior written approval of SITC. Contact info@sitcancer.org to license this measure. Neither SITC, the American Society of Clinical Oncology (ASCO), nor its members shall be responsible for any use of the Measure. ASCO's significant efforts and contributions to the development of the Measure are acknowledged. SITC is solely responsible for the review and enhancement ("Maintenance") of the Measure as of June 2021. SITC encourages use of the Measure by other health care professionals, where appropriate. ©2021 Society for Immunotherapy of Cancer (SITC). All Rights Reserved. Limited proprietary coding is contained in the Measure specifications for convenience. Users of the proprietary code sets should obtain all necessary licenses from the owners of these code sets. SITC and its members disclaim all liability for use or accuracy of any Current Procedural Terminology (CPT®), or other coding contained in the specifications. CPT® contained in the Measures specifications is copyright 2004-2021 American Medical Association. ICD-10 is copyright 2021 World Health Organization.
References:	Acharya, Y.H., & Jeter, J.M. (2013). Use of ipilimumab in the treatment of melanoma. *Clinical Pharmacology: Advances and Applications*, 5(Suppl 1), 21–27. doi:10.2147/CPAA.S45884. https://www.researchgate.net/publication/257646995_Use_of_ipilimumab_in_the_treatment_of_melanoma Brahmer, J. R., Lacchetti, C., Schneider, B. J., Atkins, M. B., Brassil, K. J., Caterino, J. M., ..., & Thompson, J. A. (2018). *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. https://ascopubs.org/doi/pdf/10.1200/JCO.2017.77.6385 Common Terminology Criteria for Adverse Events (CTCAE v5), Nov 2017, U.S. DEPARTMENT OF HEALTH AND HUMAN SERVICES, National Institutes of Health, National Cancer Institute. Dougan, M., Wang Y., Rubio-Tapia, A., & Lim, JK. (2020). *AGA Clinical Practice Update on*

	Diagnosis and Management of Immune Checkpoint Inhibitor (ICI) Colitis and Hepatitis: Expert Review. Gastroenterology. https://doi.org/10.1053/j.gastro.2020.08.063. Haanen, J.B.A.G., Carbonnel, F., C. Robert, F., Kerr, K.M., Peters, S., Larkin, J., & Jordan, K. (2017). *Management of toxicities from immunotherapy: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up*. https://www.annalsofoncology.org/article/S0923-7534(19)42153-4/pdf National Comprehensive Cancer Network (2020). *NCCN Clinical Practice Guidelines in Oncology: Management of Immunotherapy-Related Toxicities*. https://www.nccn.org/professionals/physician_gls/pdf/immunotherapy.pdf Oncology Nursing Society (2019). *Chemotherapy and Immunotherapy Guidelines and Recommendations for Practice*. Puzanov, I., Diab, A., Abdallah, K., Bingham, C. O., Brogdon, C., Dadu, R., Hamad, L., ..., Ernstoff, M. S. (2017). Managing toxicities associated with immune checkpoint inhibitors: consensus recommendations from the Society for Immunotherapy of Cancer (SITC) Toxicity Management Working Group. *Journal for ImmunoTherapy of Cancer*, 5(95). https://doi.org/10.1186/s40425-017-0300-z
Additional Information:	Health Equity Focus: N/A Practice Type: All practices
Original Approval Date:	SITC TEP: 5/5/2021
Last Updated:	SITC TEP: 5/5/2021

Appendix IV: Measure Specification Development Tools & Resources

Tool/Resource	Description
CMS Measure Authoring Development Integrated Environment (MADiE)	MADiE allows users to develop and test measures in an integrated environment. MADiE supports QI-Core profile informed authoring and testing of FHIR measures. The tool can be accessed at: https://www.emeasuretool.cms.gov/madie-mvp
Value Set Authority Center (VSAC)	The VSAC is a repository and authoring tool for public value sets created by measure developers. Use and access require a free UMLS license.
Code System Browsers	
SNOMED CT – US Edition	https://browser.ihtsdotools.org/
LOINC	https://loinc.org/ (need to sign up to access)
RxNorm	https://mor.nlm.nih.gov/RxNav/
CVX	https://www2a.cdc.gov/vaccines/iis/iisstandards/vaccines.asp?rpt=cvx
ICD-10-CM	https://icd10cmtool.cdc.gov/

Appendix V: Literature Review for Measure Development

Conducting a literature review is an essential component of the measure development process. As mentioned in Section 5, the level and characterization of available evidence guides both the type of measure as well as a measure's intended use. While less rigorous evidence such as case reports, expert opinion, or consensus documents may be a sufficient evidence base for surveillance, research, or quality improvement measures; qualityperformance measures used in an accountability context ideally should be supported by more rigorous evidence, such as strong recommendations from United States Preventive Services Task Force (USPSTF) or clinical practice guidelines; systematic reviews; and/or well-designed randomized controlled trials.

The process and approach in performing a literature search may vary depending on the scope of the search, and whether a measure concept has already been identified, as articulated in the guidance below.

- *Clinical Practice Guideline Review*
  Existing clinical practice guidelines may or may not be indexed in online databases, such as PubMed, and a manual search for applicable guidelines is essential. Measure developers should identify relevant guideline-developing organizations, perform a manual search to identify existing guidelines applicable to the clinical topic area of interest, and compile a table of recommendations relevant to the clinical topic. Existing guideline recommendations should be reviewed with the TEP Chair for relevance; the literature referenced in existing guideline recommendations may be sufficient to support a given measure concept such that the need for additional literature review is minimal.

In addition to ASCO as a major developer of cancer guidelines, examples of other organizations that develop clinical practice guidelines that may be reviewed for applicable guideline recommendations may be found at the Guidelines International Network (GIN) member directory.

- *Systematic Literature Review*
  - Non-Targeted Search: If tasked with performing a literature search for measure development in a broad topic when no measure concepts have yet been identified or targeted, a measure developer may choose to conduct a full systematic literature review using a broad search strategy. The measure developer should work with the Measure Panel Chair to formulate research/PICOT questions, create a suitable search strategy, and execute the search using available online databases, such as PubMed. In this circumstance, literature review findings must be reviewed by the measure developer and panel Chair and will guide the creation of measure concepts for further development. Once identified for development, measure concepts may need to be supported by additional targeted literature searches.
  - Targeted Search: If tasked with performing a literature search for measure development in a clinical topic when measure concepts have been identified, targeted searches follow the same approach as the above systematic literature

reviews for Non-Targeted Search with regard to formulating research/PICOT questions and executing a search strategy, but may be more limited as the scope of the search is tailored to a specific clinical measure concept, rather than a broad disease area. Targeted searches may also include supplemental evidence identified through online databases that was determined to be relevant to the measure topic either before or independent of the execution of a search strategy (e.g., grey literature, FDA approvals).

The CMS Environmental Scan Support Tool (ESST) is a can be used to compile literature from PubMed, PubMed Central, and CINAHL into a single search process, with results organized by relevancy to existing measures available in the CMS Measures Inventory. An additional component, the De Novo Measure Scan (DNMS) enables measure developers to conduct environmental scans, reducing the time required to complete the information gathering phase.

Appendix VI: Measure Prioritization Evaluation Criteria

Staff Evaluation

Measure Development Staff Evaluation Criteria Note: In general, measures with a rating of 1 in any category are not considered priorities for measure development.		
Criterion	Score	Guidance
Evidence	1	Low: Consensus-derived recommendations; weak recommendations or those with insufficient evidence; consensus-derived standards without additional supporting literature or supported by observational studies or case series
	2	Medium: Applicable evidence-based guideline recommendations with evidence quality: moderate and recommendation strength: moderate (ASCO); Category 2B (NCCN); consensus-derived standards with evidence quality and/or strength of recommendation: moderate, or lack of such ratings
	3	High: Applicable evidence-based guideline recommendation with evidence quality: moderate-high and evidence recommendation: strong (ASCO); Category 1-2A (NCCN); consensus-derived standards with evidence quality and/or strength of recommendation: high
Feasibility/ Implementability	1	Low: Numerator and denominator data is unlikely to be available from a defined data source, is unlikely to be accessible or present in meaningful quantity.
	2	Medium: Numerator and denominator data may be available from a defined data source but may not be easily accessible or robust.
	3	High: Numerator and denominator data is available from a defined data source and easily accessible.
Performance gap/Variation in care	1	Low: Variation or gap in care is undocumented; evidence suggests consistent performance or little variation in care.
	2	Medium: Lower-level studies indicate a variation or gap in care or opportunity for improvement may be present related to this aspect of care.

	3	High: Guidelines or other high-level studies suggest a variation or gap in care or opportunity for improvement related to this aspect of care.
Applicability	1	Low: Measure is relevant to a small number of patients and is unlikely to result in meaningful measurement or sufficient statistical power. Practices are unlikely to see enough relevant patients on an annual basis.
	2	Medium: Measure may be relevant to adequate numbers of patients to make measurement meaningful with sufficient statistical power (number of patients impacted and magnitude of impact). Some practices may see enough relevant patients on an annual basis.
	3	High: Measure is relevant to adequate numbers of patients to make measurement meaningful with sufficient statistical power (number of patients impacted and magnitude of impact). Most practices are likely to see enough relevant patients on an annual basis.
Use/Alignment	1	Low: Measure aligns with none or 1 of the following priorities: future incorporation within an ASCO-provided program (e.g., ASCO Certified), use in federal reporting programs (e.g., MIPS), and use in global quality programs.
	2	Medium: Measure aligns with 2 of the following priorities: future incorporation within an ASCO-provided program, use in federal reporting programs, and use in global quality programs.
	3	High: Measure aligns with 3 of the following priorities: future incorporation within an ASCO-provided program, use in federal reporting programs, and use in global quality programs.

Appendix VII: Measure Technical Expert Panel Responsibilities and Authorities

AMERICAN SOCIETY OF CLINICAL ONCOLOGY
MEASURES PANEL DESCRIPTION

GROUP:	**Measures Development & Maintenance Technical Expert Panels (TEPs)**
REPORTS TO:	**Measures Steering Group**
DEPARTMENT:	**Care Delivery**
DEPARTMENT STAFF:	**Measure Development Team**

Purpose

Technical Expert Panels (TEPs) develop de novo measures and maintain existing measures through narrative and technical specification development, testing, and approval of measures for Measures Steering Group (MSG) review.

The ASCO Measures Library supports programs such as ASCO Certified, International Quality, and the QOPI Certification Program (QCP), as well as federal and private programs. Panel activities include providing ad hoc subject matter expertise as it relates to quality measurement and practice-based quality improvement. These groups report to ASCO's Measures Steering Group (MSG), which reports to the ASCO Care and Quality Improvement Committee (CQIC).

Composition and Appointment of Panel

The panel comprises ASCO members in good standing and/or representatives from relevant medical specialties. Experts in quality measurement and practice-based quality improvement, representing both academic and community practice, may be included with the goal of having an odd number of members for voting purposes. The panel Chair and members will be selected and approved by the MSG Leadership.

Panel Members' Term

TEPs will typically have a two- to four- year charter, as necessary, for measure development, testing, maintenance, and endorsement activities. Members shall serve for the duration of the TEP or until replaced by the MSG leadership. TEPs may be reconvened, as necessary. Members may serve on reconvened TEPs as determined by MSG leadership.

Panel Chair's Term

Chairs shall serve for the duration of the TEP or until replaced by the MSG leadership. TEPs may be reconvened, as necessary. Members may serve on reconvened TEPs as determined by MSG leadership.

The Chair's responsibilities include:

- Participate in email correspondence and/or phone calls for content review, Chair call, and panel discussion.

- Provide guidance on appropriate panel composition to ensure representation by necessary stakeholders.
- Flag any potential conflicts of interest and assist ASCO staff in the implementation of a conflict mitigation strategy, if required.
- Provide clinical guidance to ASCO staff to assess initial measure concepts for clinical importance and appropriateness for development.
- Provide clinical expertise, feedback, and guidance on staff-initiated measure development/maintenance work as described below under TEP member duties.
- Assist staff in bringing panel members to consensus to enable continued progression of work products throughout the measure development lifecycle.

TEP Member Responsibilities

- Participate in email correspondence and/or phone calls for content review and panel discussion.
- Participate in member surveys regarding measure content, as required.
- Provide clinical expertise, feedback, and guidance on staff-initiated work as it relates to the identification, conceptualization, specification, maintenance, and implementation of measure concepts with special consideration given to:
    - Guideline recommendations and strength of evidence,
    - Gaps and variations in care, and opportunities for improvement,
    - Eligible populations for measure denominators including exclusions and/or exceptions,
    - Quality actions, eligible services, or outcomes that should be provided or achieved for the defined population to be captured in the measure numerator,
    - Clinical workflow in practice settings to help ensure real-world feasibility of measure implementation, and
    - Maintenance of ASCO measures.

ASCO Staff Responsibilities

- Provide primary project management and operational support for measure development and maintenance efforts.
- Conduct literature search to support the evidence review of existing measures and measure topics.
- Provide expertise to identify, specify, code, test, implement, and/or maintain ASCO measures in close collaboration with the panel.
- Initiate panel discussions on review of existing measures to include maintenance, revision, consolidation, or retirement from the ASCO Measures Library.
- Guide measure development efforts to align as closely as possible with requirements and preferences of external stakeholders.
- Schedule and manage recurring Chair and panel calls, including creation of meeting minutes and monitoring completion of next action steps.
- Disclose potential conflicts of interest and comply with applicable ASCO conflict of interest policies.
- Conduct legal reviews and prepare legal documents, as required.

References

1. McIntyre, D., Rogers, L., & Heier, E.J. (2001). Overview, History, and Objectives of Performance Measurement. Health care financing review, 22(3), 7-21.
2. CMS. (2022, May). *Quality Measures: How They Are Developed, Used, & Maintained* [Review of *Quality Measures: How They Are Developed, Used, & Maintained*]. MMSHub. https://mmshub.cms.gov/sites/default/files/Guide-Quality-Measures-How-They-Are-Developed-Used-Maintained.pdf
3. *Types of Measures | The Measures Management System.* (n.d.). Mmshub.cms.gov. https://mmshub.cms.gov/about-quality-new-to-measures/types
4. *Information Gathering | The Measures Management System.* (n.d.). Mmshub.cms.gov. Retrieved September 11, 2023, from https://mmshub.cms.gov/measures-lifecycle/measure-conceptualization/information-gathering/overview
5. *Measure Specification | The Measures Management System.* (n.d.). Mmshub.cms.gov. Retrieved September 11, 2023, from https://mmshub.cms.gov/measure-lifecycle/measure-specification/overview
6. Post, A. R., Burningham, Z., & Halwani, A. S. (2022). Electronic Health Record Data in Cancer Learning Health Systems: Challenges and Opportunities. *JCO clinical cancer informatics*, *6*, e2100158. https://doi.org/10.1200/CCI.21.00158
7. Conway, J. R., Warner, J. L., Rubinstein, W. S., & Miller, R. S. (2019). Next-Generation Sequencing and the Clinical Oncology Workflow: Data Challenges, Proposed Solutions, and a Call to Action. *JCO Precision Oncology*, *3*, 1–10. https://doi.org/10.1200/po.19.00232
8. Devereaux, M. (2022, August 26). *CMS new quality measure requirements worry some health systems* [Review of *CMS new quality measure requirements worry some health systems].* Modern Healthcare. https://www.modernhealthcare.com/safety-quality/cms-new-quality-measure-requirements-worry-some-health-systems
9. *dQMs – Digital Quality Measures | eCQM Resource Center.* (2022). Healthit.gov https://ecqi.healthit.gov/dqm?qt-tabs_dqm=0
10. *dQMs – Digital Quality Measures | eCQM Resource Center.* (n.d.). Ecqi.healthit.gov Retrieved September 11, 2023, from https://ecqi.healthit.gov/dqm?qt-tabs_dqm=1