GLOSSARY AND ACRONYMS

Access Measure
An access measure is a measure that focuses on a patient's or enrollee's attainment of timely and appropriate healthcare.

Adopted Measure
An adopted measure is a measure that has the same numerator, denominator, data source, and care setting as its parent measure. The only additional information that the measure developer needs to provide is particular to the measure's implementation use (such as data submission instructions).

Alignment
Alignment, with respect to measures, is defined by the National Quality Forum (NQF) in its *Changes to NQF's Harmonization and Competing Measures Process* as "encouraging the use of similar, standardized performance measures across and within public and private sector efforts." (p. 6) Achievement of alignment is when a set of measures works well across care settings or programs to produce meaningful information without creating extra work for those responsible for the measurement. Alignment includes using the same quality measures in multiple programs when possible. It can also come from consistently measuring important topics across care settings.

Appropriate Use Criteria
The appropriate use criteria are standards that are evidence-based (to the extent feasible) and assist professionals who order and furnish applicable services to make the most appropriate treatment decisions for a specific clinical condition (modified from CMS *Appropriate Use Criteria Program*).

Attribution
Attribution is the assignment of the results of a measure to an individual, group, or organization responsible for the decisions, costs, and outcomes (Krumholz et al., 2008).

Audit
An audit is a systematic inspection of records or accounts to verify their accuracy.

Bootstrapping
Bootstrap analysis (bootstrapping), as used in risk adjustment models, generally refers to estimating properties of a model estimate or the stability of an estimate by sampling from an approximating distribution. This is often accomplished by constructing many resamples of equal size from the observed dataset (e.g., the development sample), where the resamples are smaller than the observed dataset. This technique allows estimation of the sample distribution of a statistic. It can also be used to construct hypothesis tests. In the case of a regression or logistic regression risk adjustment model, it can be used to provide additional guidance regarding the inclusion of risk factors in the model.

Business Case
A business case is a justification for a proposed project or undertaking on the basis of its expected commercial benefit. It exists if the entity realizes a financial return on its investment in a reasonable time frame. This may be realized as profit, reduction in losses, or avoided costs. A business case may

also exist if the investor believes that a positive indirect effect on organizational function and sustainability will accrue within a reasonable time frame (Leatherman et al., 2003). The business case for a process measure relies on the financial return on the investment necessary to implement the intervention advocated by the measure. The business case for other types of measures relies on the financial return resulting from improving the quality of care indicated by the measure.

Calculation Algorithm
A calculation algorithm is an ordered sequence of data element retrieval and aggregation through which numerator and denominator events or continuous variable values are identified by a measure. Also referred to as the performance calculation.

Clinical Practice Guidelines
Clinical practice guidelines are systematically developed statements to support practitioner and patient decisions about appropriate healthcare for specific clinical circumstances.

Clinical Quality Language (CQL)
CQL is a Health Level Seven International® (HL7®) Standard for Trial Use (STU). It is part of the effort to harmonize standards between electronic clinical quality measures (eCQMs) and clinical decision support. CQL provides the ability to express logic that is human-readable yet structured enough for processing a query electronically. It replaces the logic expressions previously defined in the Quality Data Model (QDM) and QDM for use with CQL includes only the method for defining the data elements (i.e., data model).

Clinical Quality Measure (CQM)
A clinical quality measure is a mechanism used for assessing the degree to which a provider competently and safely delivers clinical services that are appropriate for the patient in an optimal time frame. CQMs are a subset of the broader category of performance measures.

Code Language
A code language, also known as programming language, is a set of commands, instructions, and other syntax used to create a software program. A high-level language is what a programmer uses to write code. The code is compiled into a low-level language, which is recognized directly by the computer hardware (Christensson, 2011).

Code System
A code system is a managed collection of concepts with each concept represented by at least one internally unique code and a human-readable description (e.g., SNOMED CT).

Collinearity
Collinearity is when two or more variables are exactly correlated which means the regression coefficients are not uniquely determined. Collinearity hurts the interpretability of the model because the regression coefficients are not unique and have influences from other features (Saslow, n.d.).

Competing Measures
Competing measures address the same topic and the same population. This term is used when considering harmonization. Refer to Related Measures.

Composite Measure
A composite measure is a measure that contains two or more individual measures, resulting in a single measure with a single score.

Conceptual Framework
A conceptual framework is a theoretical structure of assumptions, principles, and rules that holds together the ideas comprising a broad concept.

Conflict of Interest
A conflict of interest exists when an individual (or entity) has more than one motivation for trying to achieve an objective. In measure development, this situation arises when an individual has opportunities to affect specifications for quality measures that impact an interest with which the individual has a relationship.

Construct Validity
Construct validity is the extent to which the measure actually measures what it claims to measure. Construct validity evidence often involves empirical and theoretical support for the interpretation of the construct.

Continuous Variable (CV)
A continuous variable is a measure score in which each individual value for the measure can fall anywhere along a continuous scale and can be aggregated using a variety of methods such as the calculation of a mean or median (e.g., mean number of minutes between presentation of chest pain to the time of administration of thrombolytics).

Convergent Validity (concurrent validity)
Convergent validity refers to the degree to which multiple measures of a single concept are correlated.

Cost of Care
The cost of care is the total healthcare spending, including total resource use and unit price, by payer or consumer, for a healthcare service or group of healthcare services associated with a specified patient population, time period, and unit of clinical accountability.

Cost/Resource Use Measure
A cost/resource use measure is a measure of health services counts (in terms of units or dollars) applied to a population or event (including diagnoses, procedures, or encounters). A resource use measure counts the frequency of use of defined health system resources. Some may further apply a dollar amount (e.g., allowable charges, paid amounts, or standardized prices) to each unit of resource use.

Criterion
A criterion is an accepted standard, principle, or rule used to make a decision or to inform an evaluator's judgment.

Critical Data Element
A critical data element is an element that contributes most to the computed measure score. In other words, it is an element that identifies the greatest proportion of the target condition, event, or outcome being measured (numerator); the target population (denominator); population excluded (exclusion); and when applicable, risk factors with the largest contribution to variability in outcome.

C-statistic
The c-statistic is used to assess risk-adjusted models, it indicates the ability of the model to discriminate between one event and the other. If a model discriminates randomly, c = 0.5. If the risk factor modeling predicts the outcome well, then discrimination increases. The higher the c-statistic, the better the predictive power of the model.

Data Aggregation
Data aggregation is the process of combining data from multiple sources to generate performance information.

Data Element Validity (part of the Scientific Acceptability of measure properties validity subcriterion)
Data element validity is the extent to which the information represented by the data element or code used in the measure reflects the actual concept or event intended. For example

- A medication code is used as a proxy for a diagnosis code.
- Data element response categories include all values necessary to provide an accurate response.

Data Element
A data element is a single piece of information used in quality measures to describe part of the clinical care process, including both a clinical entity and its context of use (e.g., diagnosis, active).

Data Sources
Data sources are the primary source document(s) used for data collection (e.g., billing or administrative data, encounter form, enrollment forms, and/or patient medical record- electronic or paper-based).

Denominator
The denominator is a statement that describes the population evaluated by the performance measure and is the lower part of a fraction used to calculate a rate, proportion, or ratio. It can be the same as the target/initial population or a subset of the target/initial population to further constrain the population for the purpose of the measure. CV measures do not have a denominator, but instead define a measure population.

Denominator Exceptions
Denominator exceptions are those conditions that should remove a patient, procedure, or unit of measurement from the denominator of the performance rate only if the numerator criteria are not met. A denominator exception allows for adjustment of the calculated score for those providers with higher risk populations. A denominator exception also provides for the exercise of clinical judgment and should be specifically defined when capturing the information in a structured manner fits the clinical workflow. Only proportion measures use denominator exceptions.

The measured entities remove the denominator exceptions from the denominator. However, they still report the number of patients with valid exceptions. Allowable reasons fall into three general categories:

- medical reasons
- patient reasons
- system reasons

Denominator Exclusions
Denominator exclusions are patients who should be removed from the measure population and denominator before determining whether numerator criteria are met. Proportion and ratio measures use denominator exclusions to help narrow the denominator. For example, patients with bilateral lower extremity amputations would be listed as a denominator exclusion for a measure requiring foot exams. Continuous variable measures may use denominator exclusions, but use the term measure population exclusion instead of denominator exclusion.

De novo Measure
A de novo measure is a new measure that is not based on an existing measure.

Direct Reference Code (DRC)
A direct reference code is a specific code referenced directly in the eCQM logic to describe a data element or one of its attributes. DRC metadata include the description of the code, the code system from which the code is derived, and the version of that code system.

Discriminant Validity
Discriminant validity is the degree to which a test of a concept (a quality measure) is not highly correlated with other tests designed to measure theoretically different concepts. Demonstrate discriminant validity by assessing variation across multiple comparison groups (such as healthcare providers) to show that a performance measure can differentiate between disparate groups that it should theoretically be able to distinguish.

Disparities in Healthcare
Disparities in healthcare are differences in health outcomes and their determinants between segments of the population, as defined by social, demographic, environmental, and geographic attributes (Truman et al., 2011, p. 3).

Dry Run
A dry run is full-scale measure testing involving all providers/practitioners representing the full spectrum of the measured population. The purpose is to finalize all methodologies related to case identification/ selection, data collection, measurement calculation, and to quantify unintended consequences.

Efficiency Measure
An efficiency measure is the cost of care (inputs to the health system in the form of expenditures and other resources) associated with a specified level of health outcome.

Electronic Clinical Quality Measure (eCQM)
An electronic clinical quality measure is a CQM expressed and formatted to use data from electronic health records (EHRs) and/or health information technology (IT) systems to measure healthcare quality, ideally data captured in structured form during the process of patient care. For the eCQM to be reported from an EHR, the Health Quality Measure Format (HQMF) is used to format the eCQM content using CQL to express the logic and QDM to express the data elements needed to evaluate a measured entity's performance.

Electronic Health Record (EHR)
The electronic health record is also known as the electronic patient record, electronic medical record, or computerized patient record. As defined by the International Social Security Association, an EHR is a "longitudinal electronic record of patient health information generated by one or more encounters in any care delivery setting. Included in this information are patient demographics, progress notes, problems, medications, vital signs, past medical history, immunizations, laboratory data, diagnoses and treatment, medications, allergies, immunizations as well as radiology images and laboratory results."

Empirical Evidence
Empirical evidence is the data or information resulting from studies and analyses of the data elements and/or scores for a measure as specified, whether unpublished or published.

Encounter
An encounter, as defined by the ASTM International is: "(1) an instance of direct provider/practitioner to patient interaction, regardless of the setting, between a patient and a practitioner vested with primary responsibility for diagnosing, evaluating or treating the patient's condition, or both, or providing social worker services; and (2) a contact between a patient and a practitioner who has primary responsibility for assessing and treating the patient at a given contact, exercising independent judgment." Encounter serves as a focal point linking clinical, administrative, and financial information. Encounters occur in many settings: ambulatory care, inpatient care, emergency care, home healthcare, field and virtual (telemedicine).

Environmental Scan
An environmental scan is the process of systematically reviewing and interpreting data to identify issues and opportunities that will influence prioritization of current or future plans.

Expert Consensus
Expert consensus is the recommendations formulated by one of several formal consensus development methods such as consensus development conference, Delphi method, and nominal group technique.

Face Validity
Face validity is the extent to which a test appears to cover the concept it purports to measure "at face value." It is a subjective assessment by experts of whether the measure reflects the quality of care (e.g., whether the proportion of patients with blood pressure < 140/90 is a marker of quality.)

Fast Healthcare Interoperability Resources® (FHIR)
FHIR® is an HL7 standard for exchanging healthcare information electronically. FHIR can be used as a stand-alone data exchange standard and can be used in partnership with existing widely used standards. (HL7, n.d.)

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

Gaming
Gaming is when measured entities exploit weaknesses in the measurement system to manipulate the data to make their performance look better than they actually are. Includes limiting access to certain populations, neglecting care, or overuse of medications or services to ensure that the measure results are favorable.

Grey Literature
Grey literature is unpublished or not commercially indexed material that can include documentary materials issued by government, academia, business, and industry such as technical reports, working papers, and conference proceedings. For example, contributors to the New York Academy of Medicine Grey Literature website include the Agency for Healthcare Research and Quality (AHRQ), NQF, Centers for Disease Control and Prevention (CDC), the Department of Health and Human Services (HHS), The Joint Commission, National Academy of Sciences, RAND, and RTI International.

Harmonization
Harmonization is the standardization of specifications for related measures with the same measure focus (e.g., influenza immunization of patients in hospitals or nursing homes); related measures for the same target population (e.g., eye exam and HbA1c for patients with diabetes); or definitions applicable

to many measures (e.g., acceptable range for adult blood pressure) so they are uniform or compatible, unless differences are justified (i.e., dictated by the evidence). The dimensions of harmonization can include numerator, denominator, exclusion, calculation, and data source and collection instructions. The extent of harmonization depends on the relationship of the measures, the evidence for the specific measure focus, and differences in data sources. Value sets used in measures (especially eCQMs) should be harmonized when the intended meaning is the same. Harmonization of logic in eCQMs is beneficial when the data source in the EHR is the same.

Health Information Technology (Health IT)
Per Section 3000 of the HITECH Act, the term 'health information technology' means "hardware, software, integrated technologies or related licenses, intellectual property, upgrades, or packaged solutions sold as services that are designed for or support the use by healthcare entities or patients for the electronic creation, maintenance, access, or exchange of health information."

Health Information Technology for Economic and Clinical Health (HITECH) Act
The Health Information Technology for Economic and Clinical Health (HITECH) Act is a provision within the American Recovery and Reinvestment Act (ARRA) that authorizes incentive payments through Medicare and Medicaid to hospitals and clinicians toward meaningful use of EHRs.

Health Level Seven International® (HL7)
HL7® is a standards-developing organization that provides framework and standards for the exchange, integration, sharing, and retrieval of electronic health information that supports clinical practice and the management, delivery, and evaluation of health services.

Health Quality Measure Format (HQMF)
HQMF is a standards-based representation of quality measures as electronic documents. Individuals should refer to a quality measure expressed in this way as an eCQM.

Hosmer-Lemeshow Test (HL Test)
The HL test is a goodness of fit test for logistic regression, especially for risk prediction models. A goodness-of-fit test tells you how well your data fits the model. Specifically, the HL test calculates if the observed event rates match the expected event rates in population subgroups.

The test is only used for binary response variables (i.e., a variable with two outcomes such as alive or dead, yes or no).

Impact of a Measure (Importance Subcriterion)
The impact of a measure, now called High Priority by NQF, is when the measure topic addresses a specific national health goal or priority, affects a large numbers of patients, is a leading cause of morbidity/mortality, indicates high resource use and severity of patient/societal consequences of poor quality. For patient-reported outcomes (PROs), there is evidence that the target population values the PRO and finds it meaningful.

Importance Criterion
The importance criterion is the extent to which the specific measure focus is important to making significant gains in healthcare quality (e.g., safety, timeliness, effectiveness, efficiency, equity, patient centeredness) and improving health outcomes for a specific high-impact aspect of healthcare where there is variation in or overall poor performance.

Intermediate Outcome
An intermediate outcome is a measure that assesses the change produced by a healthcare intervention that leads to a long-term outcome.

Internal Consistency Reliability Testing
Internal consistency reliability testing is testing a multiple item test or survey to assess the extent to which the items designed to measure a given construct are inter-correlated. Pertains to survey type measures and to the data elements used in measures constructed from patient assessment instruments.

Inter-Rater/Abstractor Reliability Testing
Inter-rater reliability testing assesses the extent to which observations from two or more human observers are congruent with each other.

Intra-Class Correlation
Intra-class correlation (ICC) refers to correlations within a class of data (for example correlations within repeated measurements of weight), rather than to correlations between two different classes of data (for example the correlation between weight and length) (Liljequist, Elfving, & Skavberg Roaldsen, 2019).

Inverse Measures
Inverse measures are measures where a lower performance rate is better. For example, the National Healthcare Safety Network calculates most healthcare-associated infections (HAIs) as a standardized infection ratio (SIR). The SIR compares the actual number of HAIs (i.e., the numerator) with the predicted number based on the baseline U.S. experience (e.g., standard population), adjusting for several risk factors that have been found to be most closely associated with differences in infection rates. The goal is to have the numerator equal to or very close to zero, thereby having an SIR equal to or very close to zero.

Jira
Jira is an Atlassian software application that tracks issues and bugs. It also allows users to quickly search issues that have or are currently being resolved. HHS Groups are using the ONC Project Tracking System (Jira) to track issues with eCQMs and eCQM-related standards and tools.

Kappa Coefficient
The Kappa coefficient is a statistical measure of inter-rater agreement for qualitative (categorical) items. Cohen's kappa can be thought of as a chance-corrected proportional agreement. Possible values range from +1 (perfect agreement), 0 (no agreement above that expected by chance) to -1 (complete disagreement).

Lean
Lean is a system of organization principles of process improvement to maximize value and eliminate waste.

Level of Analysis
The level, or unit, of analysis is a performance measurement level (e.g., clinician, health plan, county populations).

Logic
Logic is the criteria used to define a quality measure and its key components.

Material Change
A material change is one that changes the specifications of an endorsed measure to affect the original measure's concept or logic, the intended meaning of the measure, or the strength of the measure relative to the measure evaluation criteria.

Measure Applications Partnership (MAP)
The Measure Applications Partnership is an NQF-convened, multi-stakeholder group that provides input to HHS on the list of measures for use in a specified program. The MAP consists of multiple working groups, including Clinicians, Post-Acute Care/Long-Term Care, Hospitals, and Dual-Eligible Beneficiaries.

Measure Maintenance
Measure maintenance is the periodic and consistent reviewing, evaluating, and updating of performance measures to ensure continued reliability, validity, feasibility, importance, usability, and alignment with scientific understanding. It also involves comparison to similar measures for potential harmonization.

Measure Score
The measure score is the numeric result that is computed by applying the measure specifications and scoring algorithm. The computed measure score represents an aggregation of all appropriate patient-level data (e.g., proportion of patients who died, average lab value attained) for the measured entity (e.g., hospital, health plan, home health agency, clinician). The measure specifications designate the measured entity and to whom the measure applies.

Measure Set
A measure set is a group of measures related in some way such as measures addressing a specific condition, procedure, or specialty.

Measure Steward
A measure steward is an individual or organization that owns a measure and is responsible for maintaining the measure. Measure stewards are often the same as measure developers, but not always. Measure stewards are also the ongoing point of contact for people interested in a given measure.

Measure Testing
Measure testing is the empirical analysis to demonstrate the reliability and validity of the measure as specified, including analysis of issues that pose threats to the validity of conclusions about quality of care such as exclusions, risk adjustment/stratification for outcome and resource use measures, methods to identify differences in performance, and comparability of data sources/methods.

Measure Validity (part of the Scientific Acceptability of measure properties validity subcriterion)
Measure validity is when the measure accurately represents the evaluated concept and achieves the intended purpose (i.e., to measure quality). For example, the measure

- clearly identifies the evaluated concept (i.e., face validity)
- includes all necessary data elements, codes, and tables to detect a positive occurrence when one exists (i.e., construct validity)
- includes all necessary data sources to detect a positive occurrence when one exists (i.e., construct validity)

Measured Entities
Measured entities are the front-line clinicians, including health information technology professionals, and their organizations, who collect quality measurement data. Measured entities are the implementers of quality measures. The effect of quality measure data collection on clinician workflow can be negative. There may be effects on their payments, positive and negative, with respect to reporting and actual performance on quality measures. Because of these potential effects, measured entities should be involved in all aspects of the Measure Lifecycle.

Measures Under Consideration (MUC)
The Measures under Consideration is a list of quality and efficiency measures HHS is considering adopting, through the federal rulemaking process, for use in the Medicare program. Made publicly available by December 1 each year for categories of measures that are described in section 1890(b)(7)(B)(i)(I) of the Social Security Act (SSA) as amended by Section 3014 of the Patient Protection and Affordable Care Act (ACA).

Medical Record (Data Source)
The medical record is data obtained from the records or documentation maintained on a patient in any healthcare setting (e.g., hospital, home care, long term care, practitioner office). Includes digital and paper medical record systems.

Metadata
Metadata are data that describe data. For example, document tags such as author and date created.

Minor Change
A minor change does not change the process of data collection, aggregation, or calculation, nor does it change the intended meaning of the measure or the strength of the measure in terms of the measure evaluation criteria. For example, the updates to the codes in eCQMs with the Annual Update are minor changes.

Morbidity
Morbidity is the rate of incidence of disease. For example, if a lumbar puncture is improperly performed, significant morbidity may follow. It also can refer to the relative incidence of a particular disease state or symptom.

Mortality
Mortality is the number of deaths in a certain group of people in a given time or place. The term "mortality rate" is used synonymously with death rate.

Multiple Chronic Conditions (MCC)
Multiple chronic conditions are defined by the NQF Multiple Chronic Conditions Measurement Framework as "having two or more concurrent chronic conditions that collectively have an adverse effect on health status, function, or quality of life and that require complex healthcare management, decision-making, or coordination." (pp. 7-8)

Non-parametric Methods
Non-parametric methods are a type of statistical test not involving the estimation of parameters of a statistical function. (Nonparametric, n.d.)

Null Performance Rate
The null performance rate is when all of the denominator eligible instances are attributed to all denominator exceptions. Therefore, the performance rate for satisfactory reporting would be 0/0 (null).

Numerator
The numerator is the upper portion of a fraction used to calculate a rate, proportion, or ratio. Also called the measure focus, it is the target process, condition, event, or outcome. Numerator criteria are the processes or outcomes expected for each patient, procedure, or other unit of measurement defined in the denominator. A numerator statement describes the clinical action that satisfies the conditions of the performance measure.

Numerator Exclusions
Numerator exclusions define instances that should not be included in the numerator data. Use numerator exclusions only in ratio and proportion measures.

Opportunity for Improvement
Opportunity for improvement is when data demonstrate considerable variation or overall, less-than-optimal performance, in the quality of care across providers, and/or there are disparities in care across population groups.

Outcome Measure
An outcome measure is a measure that focuses on the health status of a patient (or change in health status) resulting from healthcare – desirable or adverse.

Overfitting
Overfitting a model is when a statistical model begins to describe the random error in the data rather than the relationships between variables. This occurs when the model is too complex. In regression analysis, overfitting can produce misleading R^2 values, regression coefficients, and p-values (Frost, n.d.).

Paperwork Reduction Act (PRA)
The PRA mandates that all federal government agencies must obtain approval from the Office of Management and Budget (OMB) before collection of information that will impose a burden on the public. Measure developers should be familiar with the PRA before implementing any process that involves the collection of new data.

Parameter Estimates
Parameter estimates (also called coefficients) are the change in the response associated with a one-unit change of the predictor, while all other predictors are held constant. Types of parameter estimates include

- Point estimates, which are the single, most likely value of a parameter. For example, the point estimate of population mean (the parameter) is the sample mean (the parameter estimate).
- Confidence intervals, which are a range of values likely to contain the population parameter.

Parametric Methods
Parametric methods make certain assumptions about a data set, namely that the data are drawn from a population with a normal distribution. Parametric methods generally have high statistical power. (Tyler, 2017)

Patient-Reported Outcome (PRO)
PROs are defined as any report of the status of a patient's health condition or health behavior that comes directly from the patient, without interpretation of the patient's response by a clinician or anyone else. This definition reflects the key domains[1] of

- health-related quality of life (including functional status)
- symptoms and symptom burden (e.g., pain, fatigue)
- health behaviors (e.g., smoking, diet, exercise)

(Adapted from the Food and Drug Administration (FDA) *Guidance for Industry PRO Measures: Use in Medical Product Development to Support Labeling Claims*)

Patient-Reported Outcome Measure (PROM)
PROMs are defined by NQF in PROs in Performance Measurement as an "instrument, scale, or single-item measure used to assess the PRO concept as perceived by the patient, obtained by directly asking the patient to self-report." (p. 27)

Patient-Reported Outcome-based Performance Measure (PRO-PM)
A patient-reported outcome-based performance measure (PRO-PM) is a performance measure that is based on patient-reported outcome measure (PROM) data aggregated for an accountable healthcare entity. The data are collected directly from the patient using the PROM tool, which can be an instrument, scale, or single-item measure.

Population
The population is the total group of people of interest for a quality measure, sometimes called the target/initial population. The measure population is a defined subset appropriate to the measure set not excluded from the individual measure.

Population Health Quality Measure
A population health quality measure is a broadly applicable indicator that reflects the quality of a group's overall health and well-being. Topics include access to care, clinical outcomes, coordination of care and community services, health behaviors, preventive care and screening, and utilization of health services.

Predictive Validity
Predictive validity is the ability of measure scores to predict scores on some other related valid measure. The degree to which the operationalization can predict (or correlate) with other measures of the same measured construct at some time in the future.

Process Measure
A process measure is a measure that focuses on steps that should be followed to provide good care. There should be a scientific basis for believing that the process, when executed well, will increase the probability of achieving a desired outcome.

Proportion
A proportion is a score derived by dividing the number of cases that meet a criterion for quality (i.e., the numerator) by the number of eligible cases within a given time frame (i.e., the denominator) when the numerator cases are a subset of the denominator cases (e.g., percentage of eligible women with a mammogram performed in the last year).

[1] Note that CMS and other HHS agencies define and use the term "domain" differently from one another. Therefore, the Blueprint defines the term "domain" differently in different contexts, depending on the relevant agency within the discussion.

Public Domain
The public domain is the "The realm embracing property rights that belong to the community at large, are unprotected by copyright or patent, and are subject to appropriation by anyone" (Merriam-Webster's Dictionary, n.d.).

Qualified Clinical Data Registry (QCDR)
A QCDR is an entity with clinical expertise in medicine and in quality measurement development that collects medical or clinical data on behalf of a Merit-based Incentive Payment System (MIPS)-eligible clinician for the purpose of patient and disease tracking to foster improvement in the quality of care provided to patients.

Qualified Registry
A Qualified Registry is a vendor that collects clinical data from an individual MIPS-eligible clinician, group, or virtual group and submits it to CMS on their behalf.

Quality Data Model (QDM)
The QDM is an information model that defines relationships between patients and clinical concepts in a standardized format to enable electronic quality performance measurement. The model is the current structure for electronically representing quality measure concepts for stakeholders involved in electronic quality measurement development and reporting. The QDM provides the language that defines the criteria for clinical quality measurement. It allows the electronic definition of a clinical concept via its data elements and provides the vocabulary to relate them to each other. By relating attributes between data elements and using filtering functions, the QDM provides a method to construct complex clinical representations for eCQMs.

Quality Measure
The Patient Protection and Affordable Care Act of 2010 (ACA) defined quality measure as "a standard for measuring the performance and improvement of population health or of health plans, providers of services, and other clinicians in the delivery of health care services" (Pub. L. 111-148, §931).

Quality Reporting Document Architecture (QRDA)
QRDA is a standard document format for the exchange of eCQM data. QRDA documents

- contain data extracted from EHRs and other health IT systems
- are used to exchange eCQM data between systems
- serve as the data submission standards for a variety of quality measurement and reporting initiatives
- were adopted by the Office of the National Coordinator for Health IT (ONC) as the standard to support both QRDA Category I (individual patient) and QRDA Category III (provider's aggregate) data submission

R^2 Statistic
The R^2 statistic values describe how well the outcome of a regression model can be predicted based on the values of the risk factors or predictors. It is frequently used to assess the predictive power of specific types of risk-adjusted models.

Ratio
A ratio is a score that is derived by dividing a count of one type of data by a count of another type of data (e.g., number of patients with central lines who develop infection divided by the number of central line days). The key to the definition of a ratio is that the numerator is not in the denominator.

Receiver-Operating Characteristic (ROC) Curve
The ROC curve is a graph that provides the c-statistic value. The ROC curve graphs the predictive accuracy of a logistic regression model.

Related Measures
Related measures are measures that address either the same topic or the same population. This term is used when considering harmonization. Refer to Competing Measures.

Reliability (Scientific Acceptability of measure properties subcriterion)
Reliability reflects that the measure is well defined and precisely specified so it can be implemented consistently within and across organizations and that it distinguishes differences in performance from measurement to measurement under consistent circumstances.

Reliability Testing
Reliability testing evaluates whether the measure data elements are extracted over time, 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. Often referred to as inter-rater or inter-observer reliability, reliability also applies to abstractors and coders. It can also refer to the amount of error associated with the computed measure scores (e.g., signal vs. noise).

Respecified Measure
A respecified measure is an existing measure changed to fit the current purpose or use. This may mean changing a measure to meet the needs of a different care setting, data source, or population; or, it may mean changes to the numerator, denominator, or adding specifications to fit the current use.

Resource Use Measures
Resource use measures, also called cost and resource use measures, refer to broadly applicable and comparable measures of health services counts (in terms of units or dollars) applied to a population or event (broadly defined to include diagnoses, procedures, or encounters). A resource use measure counts the frequency of defined health system resources. Some measures may monetize the health service by applying a dollar amount such as allowable charges, paid amounts, or standardized prices to each unit of resource use.

Risk Adjustment
Risk adjustment is a statistical process used to identify and adjust for extraneous variables not associated with care delivery that threaten validity because they affect the measured outcome outside of the health system's control. Its purpose is a fairer and more accurate comparison of outcomes of care across healthcare organizations or providers.

Sample
A sample is a subset of a population. The subset should be chosen in such a way that it accurately represents the whole population with respect to some characteristic of interest. A sampling frame lists all eligible cases in the population of interest (i.e., denominator) and how they are selected.

Scientific Acceptability of the Measure Properties
Scientific acceptability of the measure properties is the extent to which the measure, as specified, produces consistent (i.e., reliable) and credible (i.e., valid) results about the quality of care when implemented.

Score (Measure Score)
The score, as defined in the NQF Glossary of Terms, is "the numeric result that is computed by applying the measure specifications and scoring algorithm. The computed measure score represents an aggregation of all the appropriate patient-level data (e.g., proportion of patients who died, average lab value attained) for the entity being measured (e.g., hospital, health plan, home health agency, clinician). The measure specifications designate the entity that is being measured and to whom the measure score applies." (p. 12)

Scoring
Scoring is the method(s) applied to data to generate results/scores. Most quality measures categories produce rates; however, other scoring categories include categorical values, CVs, counts, frequency distributions, non-weighted score/composite/scales, ratios, and weighted score/composite/scales.

Semantic Validation
Semantic validation is the method of testing the validity of an eCQM whereby the formal criteria in an eCQM are compared to a manual computation of the measure from the same test database.

Sensitivity
Sensitivity, as a statistical term, refers to the proportion of correctly identified actual positives (e.g., percentage of people with diabetes correctly identified as having diabetes). Refer to Specificity.

Specifications
Specifications are measure instructions that address data elements, data sources, point of data collection, timing and frequency of data collection and reporting, specific instruments used (if appropriate), and implementation strategies.

Specificity
Specificity, as a statistical term, refers to the proportion of correctly identified negatives (e.g., percentage of healthy people correctly identified as not having the condition). Perfect specificity would mean that the measure recognizes all actual negatives (e.g., all healthy people recognized as healthy). Refer to Sensitivity.

Stratification
Stratification involves dividing a population or resource services into distinct, independent groups of similar data, enabling analysis of the specific subgroups. This type of adjustment can show where disparities exist or where there is a need to expose differences in results.

Structure Measure
A structure measure, also known as a structural measure, is a measure that assesses features of a healthcare organization or clinician relevant to its capacity to provide healthcare.

Synthetic Data
Synthetic data are artificially generated data used to replicate the statistical components of real-world data but do not contain any identifiable information (Macaulay, 2019).

Systematic Literature Review
A systematic literature review uses systematic and explicit methods to identify, select, and critically appraise research relevant to a clearly formulated question. A systematic literature review also collects and analyzes data from studies included in the review. Two sources of systematic literature reviews are the AHRQ Evidence-Based Clinical Information Reports and The Cochrane Library.

Target/Initial Population
The target/initial population refers to all events for evaluation by a specific performance measure involving patients who share a common set of specified characteristics within a specific measurement set to which a given measure belongs. Measured entities should draw all patients counted (e.g., as numerator, as denominator) from the target/initial population.

Test-retest Reliability Testing
Test-retest reliability testing assesses the extent to which a survey or measurement instrument elicits the same response from the same respondent across short intervals of time.

Time Interval
The time interval is the time frame used to determine cases for inclusion in the denominator, numerator, or exclusion. The time interval includes an index event and period of time.

Topped-Out
Topped-out, sometimes referred to as topped off, is when a measure has reached a level when rates can no longer increase, so there is no opportunity for performance improvement.

Underfitting
Underfitting is when a data model is unable to capture the relationship between the input and output variables accurately, generating a high error rate on both the training set and unseen data. Underfitting occurs when a model is too simple, which can be a result of a model needing more training time, more input features, or less regularization (IBM Cloud Education, 2021).

Usability and Use
Usability and Use, as defined by NQF in *Measure Evaluation Criteria and Guidance for Evaluating Measures for Endorsement*, is the "extent to which potential audiences (e.g., consumers, purchasers, providers, policymakers) 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" (p.25).

Validation
Validation is testing to determine whether the measure accurately represents the evaluated concept and achieves the purpose for which it is intended (i.e., to measure quality). Use validation in reference to statistical risk models where model performance metrics are compared between the development and validation data samples.

Validity (Scientific Acceptability of measure properties subcriterion)
Validity includes measure validity (when the measure accurately represents the evaluated concept and achieves the intended purpose, which is to measure quality) and data element validity, which is the extent to which the information represented by the data element or code used in the measure reflects the actual concept or event intended.

Validity Testing
Validity testing is empirical analysis of the measure as specified that demonstrates that data are correct and/or conclusions about quality of care based on the computed measure score are correct. Validity testing focuses on systematic errors and bias.

Validity Threats
Validity threats are measure specifications or data that can affect the validity of conclusions about quality. Potential threats include patients excluded from measurement, differences in patient mix for

outcome and resource use measures, measure scores generated with multiple data sources/methods, and systematic missing or "incorrect" data (unintentional or intentional).

Value Set
A value set is a subset of concepts drawn from one or more code systems, where the concepts included in the subset share a common scope of use (e.g., Anticoagulant Therapy).

ACRONYMS

AARP	American Association of Retired Persons
ACA	Affordable Care Act
ACE	Angiotensin Converting Enzyme
AHRQ	Agency for Healthcare Research and Quality
AMA	American Medical Association
AMI	Acute Myocardial Infarction
ANOVA	Analysis of Variance
APA	American Psychological Association
API	Application Programming Interface
APM	Alternative Payment Model
ARRA	American Recovery and Reinvestment Act of 2009
ASC	Ambulatory Surgical Center
ASCQR	Ambulatory Surgical Center Quality Reporting
ASPE	Office of the Assistant Secretary for Planning and Evaluation
AUC	Area Under the Curve
BPS	Binding Parameter Specification
CABG	Coronary Artery Bypass Graft
CAHPS	Consumer Assessment of Healthcare Providers and Systems
CAU	CQM Annual Update
CBE	Consensus-Based Entity
C-CDA	Consolidated Clinical Document Architecture
CDA	Clinical Document Architecture
CDC	Centers for Disease Control and Prevention
CDS	Clinical Decision Support
CEHRT	Certified EHR Technology
CHIP	Children's Health Insurance Program
CHIPRA	Children's Health Insurance Program Reauthorization Act of 2009
CLABSI	Central Line-Associated Blood Stream Infection
CLD	Content Logical Definition

CMIT	CMS Measures Inventory Tool
CMS	Centers for Medicare & Medicaid Services
COPD	Chronic Obstructive Pulmonary Disease
COR	Contracting Officer's Representative
CPT	Current Procedural Terminology
CPT4	CPT, 4th Edition
CQL	Clinical Quality Language
CQM	Clinical Quality Measure
CQS	Composite Quality Score
CRP	Change Review Process
CSAC	Consensus Standards Approval Committee
CT	Computerized Tomography
CV	Continuous Variable
CY	Calendar Year
DEL	Data Element Library
DENEX	Denominator Exclusion
DENEXCEP	Denominator Exclusion Exception
DENOM	Denominator
DEQM	Data Exchange for Quality Measures
DERep	Data Element Repository
DHA	Defense Health Agency
DRC	Direct Reference Code
DTS	Distributed Terminology System
DUA	Data Use Agreement
eCQI	electronic Clinical Quality Improvement
eCQM	electronic Clinical Quality Measure
ED	Emergency Department
EHR	Electronic Health Record
ELM	Expression Logical Model
ESRD	End-Stage Renal Disease

ESST	Environmental Scanning Support Tool
FACA	Federal Advisory Committee Act
FDA	Food and Drug Administration
FFS	Fee for Service
FHIR	Fast Healthcare Interoperability Resources
FOTO	Focus on Therapeutic Outcomes
FPL	Federal Poverty Level
FVC	Forced Vital Capacity
GPCK	Generic Pack
GRADE	Grading of Recommendation, Assessment, Development, and Evaluation
GUID	Globally Unique Identifier
HAC	Hospital-Acquired Condition
HACRP	Hospital-Acquired Condition Reduction Program
HAI	Healthcare-Associated Infection
HARP	HCQIS Access Roles and Profile
HCPCS	Healthcare Common Procedure Coding System
HCQIS	Healthcare Quality Information Systems
HD	Hemodialysis
HDL	High Density Lipoprotein
HHA	Home Health Agency
HHQR	Home Health Quality Reporting
HHS	Health and Human Services
HIPAA	Health Insurance Portability and Accountability Act
HITECH	Health Information Technology for Economic and Clinical Health Act
HITSC	Health Information Technology Standards Committee
HIV	Human Immunodeficiency Virus
HL	Hosmer-Lemeshow
HL7	Health Level Seven International
HOS	Health Outcomes Survey
HQI	Hospital Quality Incentive

HQMF	Health Quality Measure Format
HQR	Hospital Quality Reporting
HQRP	Hospice Quality Reporting Program
HRRP	Hospital Readmissions Reduction Program
HSLOC	Healthcare Service Location Code
HTML	Hypertext Markup Language
HVBP	Hospital Value-Based Purchasing
ICC	Intraclass Correlation Coefficient
ICD	International Classification of Diseases
ICD-9-CM	International Classification of Diseases, Ninth Revision, Clinical Modification
ICD-10-CM	International Classification of Diseases, Tenth Revision, Clinical Modification
ICD-10-PCS	International Classification of Diseases, Tenth Revision, Procedure Coding System
ICF	International Classification of Functioning, Disability, and Health
ICU	Intensive Care Unit
IDIQ	Indefinite Delivery/Indefinite Quantity
IG	Implementation Guide
IHI	Institute for Healthcare Improvement
IMPACT	Improving Medicare Post-Acute Care Transformation
IN	Specific Ingredient
IOM	Institute of Medicine
IP	Target/Initial Population
IPFQR	Inpatient Psychiatric Facility Quality Reporting
IQR	Inpatient Quality Reporting
IRB	Institutional Review Board
IRF	Inpatient Rehabilitation Facility
ISA	Interoperability Standards Advisory
IS-A	Increasing Specificity
ISO	International Organization for Standardization
IT	Information Technology
JSON	JavaScript Object Notation

LA	LOINC Answers
LDL	Low-density Lipoprotein
LOINC	Logical Observation Identifiers Names and Codes
LTC	Long-term Care
LTCH	Long-Term Care Hospital
MACPAC	Medicaid and CHIP Payment and Access Commission
MACRA	Medicare Access and CHIP Reauthorization Act
MAP	Measure Applications Partnership
MAT	Measure Authoring Tool
MC	Measure Collaboration
MCC	Multiple Chronic Conditions
MDP	Measurement Development Plan
MDS	Minimum Data Set
MedPAC	Medicare Payment Advisory Committee
MIDS	Measure and Instrument Development and Support
MIF	Measure Information Form
MIPS	Merit-based Incentive Payment System
MJF	Measure Justification Form
MMS	Measures Management System
MSS	Medicare Shared Savings
MUC	Measures Under Consideration
MUD	Measures Under Development
NCHS	National Center for Health Statistics
NCQA	National Committee for Quality Assurance
NDC	National Drug Code
NIH	National Institutes of Health
NLM	National Library of Medicine
NLP	Natural Language Processing
NQF	National Quality Forum
NUCC	National Uniform Claim Committee

NUMER	Numerator
NUMEX	Numerator exclusion
O/E	Observed-to-expected
OASIS	Outcome and Assessment Information Set
OECD	Organisation for Economic Co-operation and Development
OID	Object Identifier
OMB	Office of Management and Budget
ONC	Office of the National Coordinator for Health Information Technology
OQR	Outpatient Quality Reporting
P.L.	Public Law
PACE	Program of All-Inclusive Care for the Elderly
PACU	Post Anesthesia Care Unit
PCORI	Patient-Centered Outcomes Research Institute
PCS	Procedure Coding System
PFAC	Patient Family Advisory Council
PFE	Person and Family Engagement
PHIN	Public Health Information Network
PIN	Precise Ingredient
PRA	Paperwork Reduction Act
PRO	Patient-Reported Outcome
PROM	Patient-Reported Outcome Measure
PROMIS	Patient-Reported Outcomes Measurement Information System
PRO-PM	Patient-Reported Outcome-based Performance Measure
PSI	Patient Safety Indicator
QCDR	Qualified Clinical Data Registry
QDC	Quality Data Code
QDM	Quality Data Model
QI-Core	Quality Improvement-Core
QIN-QIO	Quality Innovation Network-Quality Improvement Organizations
QIP	Quality Incentive Program

QMTF	Quality Measures Technical Forum
QPP	Quality Payment Program
QPS	Quality Positioning System
QR	Quality Reporting
QRDA	Quality Reporting Document Architecture
QRS	Quality Rating System
RAI	Resident Assessment Instrument
RCT	Randomized Controlled Trial
REST	Representational State Transfer
ROC	Receiver-operating characteristic
ROI	Return on Investment
RRP	Readmissions Reduction Program
SCD	Semantic Clinical Drug
SCIP	Surgical Care Improvement Project
SDS	Sociodemographic Status
SES	Socioeconomic Status
SIR	Standardized Infection Ratio
SME	Subject Matter Expert
SNF	Skilled Nursing Facility
SNF QRP	Skilled Nursing Facility Quality Reporting Program
SNF VBP	Skilled Nursing Facility Value-Based Purchasing
SR	Systematic Review
SSA	Social Security Act
STS	Society of Thoracic Surgeons
STU	Standard for Trial Use
TEP	Technical Expert Panel
THA	Total Hip Arthroplasty
TKA	Total Knee Arthroplasty
TRN	Technical Release Note
TTY	RxNorm ingredient type

U.S.	United States
UCUM	Unified Code for Units of Measure
UMLS	Unified Medical Language System
URL	Uniform Resource Locator
USCDI	United States Core Data for Interoperability
USCRS	U.S. SNOMED CT Content Request System
USHIK	United States Healthcare Information Knowledgebase
USPSTF	United States Preventive Services Task Force
VADS	Vocabulary Access and Distribution System
VHA	Veterans Health Administration
vMR	Virtual Medical Record
VS	Value Set
VSAC	Value Set Authority Center
VSD	Value Set Definition
XML	eXtensible Markup Language

REFERENCES

Agency for Healthcare Research and Quality. (n.d.). *Guidelines and measures*. Retrieved August 14, 2020, from https://www.ahrq.gov/gam/index.html

American Recovery and Reinvestment Act of 2009, Pub. L. No. 111-5, § 13001, 123 Stat. 227 (2009).

Centers for Medicare & Medicaid Services. (n.d.). *Appropriate use criteria program.* Retrieved August 14, 2020, from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/Appropriate-Use-Criteria-Program/index

Christensson, P. (2011, September 23). *Programming language*. TechTerms. Retrieved April 21, 2021, from https://techterms.com/definition/programming_language

Electronic Clinical Quality Improvement Resource Center. (n.d.-a). *CQL- Clinical quality language.* Retrieved August 14, 2020, from https://ecqi.healthit.gov/

Electronic Clinical Quality Improvement Resource Center. (n.d.-b). *QDM - Quality data model*. Retrieved August 14, 2020, from https://ecqi.healthit.gov/qdm

Electronic Clinical Quality Improvement Resource Center. (n.d.-c). *QRDA - Quality reporting document architecture*. Retrieved August 14, 2020, from https://ecqi.healthit.gov/qrda

Frost, J, (n.d.). Overfitting regression models: Problems, detection, and avoidance. Retrieved May 26, 2021, from https://statisticsbyjim.com/regression/overfitting-regression-models/

Health Level Seven International. (n.d.). *HL7 FHIR release 4*. Retrieved October 22, 2020, from https://www.hl7.org/fhir/overview.html

IBM Cloud Education. (2021, March 23). *Underfitting*. Retrieved May 26, 2021, from https://www.ibm.com/cloud/learn/underfitting

International Social Security Association. (n.d.). *Information and communication technology – Guideline 92. Electronic health record system.* Retrieved August 14, 2020, from https://ww1.issa.int/guidelines/ict/180156

Krumholz, H.M., Keenan, P.S., Brush, J.E., Bufalino, V.J., Chernew, M.E., Epstein, A.J., … Walsh, M.N. (2008). Standards for measures used for public reporting of efficiency in health care. *Journal of the American College of Cardiology, 52*(18), 1518-1526. https://doi.org/10.1016/j.jacc.2008.09.004

Leatherman, S., Berwick, D., Iles, D., Lewin, L. S., Davidoff, F., Nolan, T., & Bisognano, M. (2003). The business case for quality: case studies and an analysis. *Health Affairs, 22*(2). https://doi.org/10.1377/hlthaff.22.2.17

Liljequist, D., Elfving, B., & Skavberg Roaldsen, K. (2019). Intraclass correlation – A discussion and demonstration of basic features. *PLoS ONE 14*(7), e0219854. https://doi.org/10.1371/journal.pone.0219854

Macaulay, T. (2019, October). What is synthetic data and how can it protect privacy? Tech Advisor. https://www.techadvisor.co.uk/news/small-business/what-is-synthetic-data-how-can-it-help-protect-privacy-3789108/

Merriam-Webster. (n.d.-a). Nonparametric. In *Merriam-Webster's dictionary.* Retrieved August 14, 2020, from https://www.merriam-webster.com/dictionary/public%20domain

Merriam-Webster. (n.d.-b). Public domain. In *Merriam-Webster's dictionary.* Retrieved August 14, 2020, from https://www.merriam-webster.com/dictionary/public%20domain

National Quality Forum. (n.d.-a). *Glossary of terms*. Retrieved August 14, 2020, from https://www.qualityforum.org/Measuring_Performance/Submitting_Standards/NQF_Glossary.aspx

National Quality Forum. (n.d.-b). *Phrasebook: a plain language guide to NQF jargon.* Retrieved August 14, 2020, from http://public.qualityforum.org/NQFDocuments/Phrasebook.pdf

National Quality Forum. (2012a). *MCC measurement framework*. http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=71227

National Quality Forum. (2012b). *Patient-reported outcomes in performance measurement.* https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72537

National Quality Forum. (2013). *Changes to NQF's harmonization & competing measures process: Information for measure developers*. https://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72716

National Quality Forum. (2019). *Measure evaluation criteria and guidance for evaluating measures for endorsement*. http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=88439

Paperwork Reduction Act of 1995, Pub. L. No. 104-13, 109 Stat. 163 (1995).

Patient Protection and Affordable Care Act of 2010, Pub. L. No. 111-148, 124 Stat. 119 (2010).

Saslow, E. (2018, July 11). *Collinearity- What it means, why it is bad, and how does it affect other models?* Retrieved May 26, 2021, from https://medium.com/future-vision/collinearity-what-it-means-why-its-bad-and-how-does-it-affect-other-models-94e1db984168

Statistics How To. (2014). *Goodness of fit test.* https://www.statisticshowto.datasciencecentral.com/goodness-of-fit-test/

Statistics How To. (2015). *Binary variable: Definition, examples.* https://www.statisticshowto.datasciencecentral.com/binary-variable-2/

Statistics How To. (2016). *Hosmer-Lemeshow test.* https://www.statisticshowto.datasciencecentral.com/hosmer-lemeshow-test/

Truman, B.I., Smith, C.K. Roy, K., Chen, Z., Moonesinghe, R., Zhu, J., ... Zaza, S. (2011). CDC health disparities and inequalities report — United States, 2011, Rationale for regular reporting on health disparities and inequalities – United States. In *Morbidity and Mortality Weekly Report, 60(Suppl.)*, 3-10. https://www.cdc.gov/mmwr/pdf/other/su6001.pdf

Tyler, J. (2017). *What are parametric and nonparametric tests? Sciencing.* https://sciencing.com/parametric-nonparametric-tests-8574813.html

U.S. Department of Health & Human Services, Food and Drug Administration, Center for Drug Evaluation and Research, Center for Biologics Evaluation and Research, and Center for Devices and

Radiological Health. (2009). *Patient-reported outcome measures: Use in medical product development to support labeling claims. Guidance for industry*. https://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM193282.pdf