BlueCross BlueShield of Texas

PEAQ Physician Efficiency, Appropriateness, & Quality

# Physician Efficiency, Appropriateness & Quality (PEAQ) Program FAQs for Providers

## GENERAL QUESTIONS

**What is PEAQ?**

Blue Cross and Blue Shield of Texas (BCBSTX)'s Physician Efficiency, Appropriateness & Quality (PEAQ) Program evaluates physician performance in a transparent and multidimensional way. Our goal is to measure and maximize physician efficiency, appropriateness, and quality. This includes initiatives designed to improve employee health, increase transparency around provider quality, and promote the efficient use of member and employer dollars by guiding members to high-quality, affordable care.

**How is PEAQ being used?**

PEAQ is being used to assist physicians with transparency around how they are performing in comparison to their peers. We believe that this is very important to physicians. It can also be used to help guide BCBSTX members to high-performing physicians. In the future, it may be used to help BCBSTX define its network of contracted physicians.

**What specialties will be applicable to PEAQ?**

We evaluate 15 specialties broken out between PCP, Medical, and Surgical specialties. Within these specialties we measure a physician's working specialty. The specialties and corresponding working specialties we review are as follows:

| Medical | Surgical | Primary Care |
|---|---|---|
| Cardiology | Cardiothoracic Surgery | Family Medicine |
| Endocrinology | Ophthalmology | Internal Medicine |
| Gastroenterology | Orthopedic Surgery | Pediatrics |
| Nephrology | Urology | |
| Pulmonary | Vascular Surgery | |
| Rheumatology | | |
| Obstetrics/Gynecology | | |

Geriatric Medicine and General Practice specialties are included in Internal Medicine and Family Medicine, respectively. Measured specialties will be re-evaluated on a continual basis.

**Will you add additional specialties? If so, when?**
We will continually evaluate measured specialties based on data availability, clinical relevance, and volume.

**How will these measures be used?**
The program is designed to be informational while creating greater transparency for our providers. The information may be used in the future beyond transparency as allowed, in use cases such as member steerage and network optimization. For example, data can be used to steer members to high-performing providers through tools such as Provider Finder, inform benefit design and tiered networks, and update provider systems about physician practice patterns.

**What requirements must be met for a specialty to be included in PEAQ evaluations?**
Specialties were initially selected because they were part of former performance measurement programs and also common in the marketplace. In the future, our additions will be added based on the similar considerations.

**Does the PEAQ program include a review of my medical records?**
No, we do not review medical records. Our data is either claims-based or obtained from feeds such as labs.

**How can providers obtain their ratings?**
Ratings will be distributed via Availity Payer Spaces. Please see the Availity User Guide on the PEAQ website for more information about generating reports.

**Is there any documentation on PEAQ?**
You can find the methodology document on the PEAQ website.

**How will updates to the program and methodology be communicated?**
Updates and changes to the PEAQ program will be communicated through the BCBSTX Provider website and other Provider communications, such as Blue Review.

**Can a physician opt out of the PEAQ program?**
No. If a physician has enough claim volume in a specialty that is being reviewed, they will receive a PEAQ report.

**Why does my PEAQ report list a TIN that I'm no longer affiliated with?**
In order to ensure credible results, the PEAQ measurement uses historical look-back periods. Therefore, a physician can be rated at a practice where they no longer work if they were there during the measurement period. For the measurement evaluation periods, you may consult the notice that was sent prior to the evaluation period or your PEAQ report.

BlueCross BlueShield of Texas

# NETWORK QUESTIONS

**To what networks does this apply?**
PEAQ performance applies to all networks. Physicians are rated at an NPI level and can participate in multiple networks.

**I practice in multiple locations. Do I have different PEAQ results for my various practice locations?**
PEAQ is presently only at the physician level and is not tied specifically to the practice location. We will only assign one rating for a provider within the same region and working specialty.

**Is there an identifier that providers will have on Provider Finder?**
No, right now there is no additional identifier on Provider Finder. This will be re-evaluated on an annual basis.

# RATING QUESTIONS

**What data will be utilized as benchmark?**
The benchmark will be a provider's peer group defined by working specialty and geographic region.

**How are physicians evaluated?**
Physicians will be evaluated in the three performance categories (efficiency, appropriateness, and quality) and their measures will be consolidated into an overall metric. Please see our detailed methodology document on the PEAQ website for more information on physician evaluation.

**How are physicians' peer groups determined?**
Physicians will be compared to peers within the same working specialty and geographic region. Please see our detailed methodology document on the PEAQ website for more information on peer groups.

**I am a particular type of specialist, but the report lists me as a different type of specialist. How did you come up with this specialty and what should I do to correct this?**
We start with the specialty you provided on your application, which is loaded into our system. Then we evaluate the type of care the physician provides. If the type of care is highly inconsistent with the declared specialty (in extremely rare cases), we modify that specialty to align with the physician's actual practice of care.

**At what level will physicians be measured?**
Physicians will be measured at the individual contracted physician level (MD/DO) using NPIs within working specialties and regions. Future iterations of PEAQ may expand to include other provider types and levels, such as physician groups.

**Are ratings risk-adjusted?**
Yes. Efficiency and quality ratings are adjusted based on multiple factors, such as patient comorbidities, demographics, disease severity, and disease category.

**What has PEAQ done to adjust for patient behaviors during the COVID-19 pandemic?**
We believe that patients within a working specialty and geographic region behave similarly. Since PEAQ compares physicians to their peers within working specialties and geographic regions, patients' behaviors should affect physicians similarly, which should alleviate COVID-19 impacts.

**Why are my peer counts different for my quality and efficiency reports?**
The minimum member, measure, and episode thresholds vary across the quality and efficiency programs. Thus, some physicians measured for quality are not measured for efficiency and vice versa, resulting in different peer groups for quality and efficiency.

**Why didn't I receive a PEAQ report?**
For a variety of reasons, not all physicians are evaluated by the PEAQ program. These reasons may include:

- Specialty not included in current measurement
- Inadequate peer group information
- Non-MD/DO physician
- Limited volume of BCBSTX patients
- Limited number of patients in the denominator of quality measures

**Why don't you rate PAs, NPs, hospitalists, or ER doctors?**
PEAQ is limited to MD/DOs at this time. We are purposely keeping our scope limited as we roll out PEAQ initially. In the future, we may add other clinical designations such as Physician Assistants.

**Will I be able to compare historical PEAQ reports with current reports?**
Only current PEAQ reports are made available. Trending reports are not offered at this time. However, everyone's encouraged to keep reports for future reference.

**My patients are more costly and with more complex underlying conditions than other groups. How can my results be adjusted to reflect that?**
Results are risk-adjusted for that. Efficiency, appropriateness, and quality are adjusted on multiple factors, such as patient comorbidities, demographics, disease severity, and disease category.

**Who else can see my results?**
At this time, only those in your practice that pull the reports and disseminate them can see your results, as long as they have appropriate Availity TIN permissions. We do not share your report with other providers. The data may be used in our Provider Finder and by our network teams to understand the composition of the network.

BlueCross BlueShield of Texas

# APPEALS & RECONSIDERATION QUESTIONS

**How can I challenge an evaluation?**
Physicians may request reconsideration of their PEAQ designation before results are finalized. Physicians will have 45 calendar days following initial notification to submit a reconsideration request. Reconsideration request forms are available on the PEAQ website. Requests should be submitted to PEAQ_inquiries@bcbstx.com. Questions sent to the same mailbox will be answered beyond the reconsideration request period, but scores will already be finalized.

Reconsideration requests will be reviewed by a panel including BCBSTX medical director(s), network representative(s), quality specialist(s) and data scientist(s). The physician will be notified of the response to the request.

**How can I request to update my provider-level data, provider demographic data, or working specialty?**
If the provider or provider admin requests to update any provider-level data, provider demographic data, or working specialty, the Provider Data Update process should be followed, as detailed on the BCBSTX Provider website. *Note: PEAQ results will not factor in updated data until the next PEAQ refresh.*

# QUALITY QUESTIONS

**How are patients attributed to physicians for quality measurement?**
*PCPs*: For Health Maintenance Organization (HMO) products, patients are assigned to the Primary Care Provider they have selected. For Preferred Provider Organization (PPO) products, a patient's primary care attribution is derived from their historic claims data. Attributed patients must have a full year of coverage before they are included in measurement. For a PCP to be rated, they must meet a minimum patient volume threshold of 30 BCBSTX patients and at least 10 patients in the denominator of a quality measure.

*Specialists*: Patients are attributed to specialists based on claims data from diagnosis codes for medical specialties or procedure codes for surgical specialties. Attributed patients must have a full year of coverage before they are included in measurement. For a specialist to be rated, they must meet a minimum patient volume threshold.

**Why are there so few quality measures in my specialty?**
Several specialties have few relevant measures that we can analyze. Many potential measures require clinical data that is not available to us. We are constantly evaluating new potential metrics to improve the breadth of measurement.

**How do I know which quality measures were used for my specialty?**
The quality measures are listed in the PEAQ Methodology document posted on the PEAQ website.

**Where can I go to find out more information about HEDIS measurements?**
Information on HEDIS measurements is available on the National Committee for Quality Assurance (NCQA) website and on the Provider section of the BCBSTX website under the Clinical Resources Quality Improvement tab.

# APPROPRIATENESS QUESTIONS

**How do we define appropriateness? Where can I go to find out more information about how appropriate care is determined?**
Our appropriateness metrics evaluate the extent to which physicians make decisions about patient care that are consistent with current evidence-based guidelines. HCSC has partnered with Motive Medical Intelligence to deliver these measures using the Practicing Wisely Solutions appropriateness of care measurement methodology.

Appropriateness of care measures are determined through a systematic examination of data, evidence, and expert opinion. Data are abstracted from BCBS[IL/TX] claims data as shared by the plan. Evidence is culled by Motive Medical Intelligence from peer-reviewed literature, which is analyzed with quantitative bibliometrics. Input is derived from physicians who are in active clinical practice in the areas being measured, and who are identified by quantitative indices of expertise.

[For Texas providers, appropriateness is shown on reports but is currently informational only.]

**How does a provider know which appropriateness measures were used for their specialty?**
The appropriateness component of the PEAQ program measures physicians within primary care, medical, and surgical specialties. Measured specialties for the appropriateness component of the PEAQ program include:

| Medical | Surgical | Primary Care |
|---|---|---|
| Cardiology | Cardiothoracic Surgery | Family Medicine |
| Endocrinology | Ophthalmology | Internal Medicine |
| Gastroenterology | Orthopedic Surgery | Pediatrics |
| Nephrology | Urology | |
| Pulmonary | Vascular Surgery | |
| Rheumatology | | |
| Obstetrics/Gynecology | | |

The appropriateness measures are listed in the methodology on the PEAQ website.

**How are patients attributed to physicians for appropriateness measurement?**

Measures are attributed to the physician responsible for the care decision. Cases that cannot be definitively attributed to a physician are excluded. Several considerations are involved in proper attribution, depending on the measure:

- Specialty procedures are attributed only to physicians within the specialty of interest.
- The timing of interventions relative to physician visits may be a factor in determining attribution.
- For episodes of care in which the physician rendering the service is responsible for the decision to deliver that service (e.g., cardiac catheterization), the event of interest is attributed to the physician identified on the claim as the rendering NPI (National Provider Identifier).
- For evaluation and management (E&M) measures, the event of interest is attributed to the presumptive ordering physician at the most proximate prior E&M visit instead of the rendering NPI. This approach is used because the physician rendering the service may be different from the physician responsible for the decision to deliver that service.

**Which appropriateness components are risk-adjusted?**

All of the appropriateness measures are adjusted for risk and severity through the use of detailed denominator exclusion criteria within the measure logic to ensure that all cases that are included in the case mix for a given measure are similar in terms of their level of risk or severity. For example, measures that evaluate the use of high-cost resources exclude cases in which the use of those resources is more frequently justified, such as the use of advanced imaging in the context of cancer. By logically embedding risk and severity adjustment in this way, this approach to measurement mirrors clinical decision-making.

**Can you explain the methodology for the measure rates?**

Appropriateness measure rates are formulated as numerator–denominator statements, using a standardized denominator, exclusion, attribution, and numerator methodology. Cases meeting the numerator and denominator inclusion criteria and exclusion criteria requirements are identified within claims datasets to identify patterns of inappropriate decision-making across cases.

A minimum threshold number of cases is established to generate statistically significant analyses, while ensuring that physicians are evaluated based on the care decisions they make regularly.

Although the appropriateness measures look at both overuse and underuse, all individual appropriateness measures are structured as evaluations of *inappropriateness*. This means that lower individual measure rates are always better. Overuse measures evaluate whether services were provided that were inappropriate (e.g., ordering an MRI for low back pain), whereas underuse measures show the absence of appropriate services (e.g., strep test underuse in antibiotic treatment pharyngitis, or underuse of conservative therapy before surgery).

**How do you account for complex cases, including those with secondary and tertiary referrals and patients with increased disease burden? What about extenuating circumstances that require physicians to go against their better judgment, such as school district policies requiring an antibiotic prescription for conjunctivitis?**

While lower rates are always preferred for both overuse and underuse measures, the appropriateness measures account for situations in which it might be reasonable to provide a service that would otherwise be considered inappropriate. The appropriateness measures address these variations in two ways.

First, a range of better practice (ROBP) threshold according to a two-step process. The first step is statistical. It uses standardized testing and rigorous methodology—such as one would encounter in published, peer-reviewed studies—to assess the range of physician performance variation. The second step is expert curation. Motive's 600-member board of subject matter experts curates the data utilizing systematic algorithms and quantitative techniques to account for factors that inform the real-world practice of medicine. The resultant ROBP brings objectivity and reproducibility to assessments that would otherwise be subjective and irreproducible. The ROBP threshold is set generously to ensure that any rate of performance that is over the threshold is clearly inappropriate. Factors considered in setting the ROBP threshold include variations in medical coding practices, gaps in claims data, and the realities of clinical medicine, such as regional resource limitations, reliance on tertiary referral, and individual patient factors.

Second, cases that cannot be attributed to an individual physician with complete certainty are excluded. For example, if multiple physicians of the same specialty are involved in a case—including any instances of secondary or tertiary referral—the case cannot be attributed to a single individual and is disregarded for the purposes of measurement.

**How do you decide what to measure?**

The appropriateness measures must meet strict criteria to be considered for development, and they must pass rigorous testing and clinical validation before they are included in our methodology.

First, all appropriateness measures are specifically focused on low-value care—meaning services that have the potential to lead to both harm to members and the unnecessary use of costly resources, and in which there is great variability in practice patterns.

Second, every measure must meet the American College of Physicians (ACP) criteria for performance measurement—meaning they must address issues that are important and have meaningful clinical impact, that address inappropriate use including both overuse and underuse, that have a clearly defined and well established evidence base, that are clear and transparent in their measure specifications and methodology, and that can be measured fairly and accurately and provide actionable guidance without creating undue burden for data collection.[1]



[1] Maclean CH, Kerr EA, Qaseem A. Time Out — Charting a Path for Improving Performance Measurement. *N Engl J Med*. 2018:378;19(1757-1761).

Third, every measure must pass a battery of tests for statistical validity, including analyses to set and validate minimum denominator thresholds, analyses of measure stability over time, and analyses of reliability across multiple datasets. Every measure is also subjected to extensive scrutiny by multiple subject matter expert physicians with relevant expertise at each step in the measure development process to ensure that the measures are fair and accurate and that they will be acceptable to practicing physicians while providing actionable insights for helping them improve the care they provide.

# EFFICIENCY QUESTIONS

**How is cost attributed to physicians for efficiency measurement?**

Cost is attributed to physicians using diagnostic-based episodes of care. Episodes are created using IBM Watson's Medical Episode Grouper (MEG) algorithm, which clusters completed episodes of care from a patient's medical and pharmacy claims into one of over 500 episode categories. Physicians who treat a patient for an episode are attributed cost-based on the amount of work they do to treat the episode. The amount of work is based on the proportion of RVUs each physician's services comprise. For more information, see Terminology Questions in this FAQ or the methodology on the PEAQ website.

**What risk adjustment methodology is used for efficiency?**

Efficiency uses MEG output groups and other demographic factors within an internally developed risk adjustment methodology. The risk adjustment uses machine learning techniques to accurately adjust for patient population differences related to comorbidities and demographics. The risk model determines a risk factor at a geographic and episodic level by referencing comorbidity data determined by conditions.

The process runs 2 models. The first model determines an accurate cost by using all attributes (member and episode) as features. The second one models cost but uses features of the episode only with the same parameters as the first. This is designed so that the ratio between the two models reflects the impact the member features have on the episode cost, therefore the additional risk.

BlueCross BlueShield of Texas

# TIMELINE QUESTIONS

**How often will physicians be evaluated?**
PEAQ results are updated annually. The evaluation frequency is subject to change. Physicians will be notified through the Blue Review newsletter a month prior to release of new performance evaluations.

**How often will evaluation results be shared?**
Physicians will receive their results annually. The evaluation frequency is subject to change. Notification of the distribution of future PEAQ reports will be posted on the BCBSTX Provider website and in Blue Review.

# TERMINOLOGY QUESTIONS

**What is a working specialty?**
Providers are matched to other providers within the same working specialty. The working specialty represents a provider's specialty and/or sub-specialty and is determined using information from BCBSTX's provider demographics database and claims submitted by the provider. The working specialty may be more specific than a provider's self-declared specialty. For example, working specialty may distinguish an interventional cardiologist from a non-interventional cardiologist based on claims submitted by the physician.

**What is an episode of care?**
Episodes of care are diagnostic-based groupings of claims (one or more claims that correspond to a single patient and a single diagnosis) from MEG. In total, there are over 500 episode categories that are further segmented by severity and disease stage progression and grouped as Acute or Chronic. Only episodes marked as complete are utilized in efficiency results.

**What is standard deviation?**
Standard deviation is a measure of the amount of variation or dispersion of a set of values.