

# Physician Compare Benchmark and Star Ratings Fact Sheet

## Performance Year 2017

### Overview

This fact sheet provides information about the Physician Compare star rating system, including the benchmark methodology and the star attribution methodology. The Centers for Medicare & Medicaid Services (CMS) finalized an item-level benchmark as the basis for the Physician Compare star ratings in the 2016 Physician Fee Schedule (PFS) final rule ([80 FR 71128 through 71129](#)). The first star ratings were publicly reported in late 2017 for a subset of group Physician Quality Reporting System (PQRS) performance information. Physician Compare will continue to publicly report a subset of Quality Payment Program performance information as star ratings ([82 FR 53828](#)).

Star ratings are only publicly reported if the measure data meet the Physician Compare public reporting standards and resonate with users (§414.1395(b)). Physician Compare conducts analyses to ensure that a statistically robust benchmark is established by measure and collection type, and that there are robust star rating cut-offs that reliably categorize groups into each star rating.

Visit the [Physician Compare Initiative page](#) to download the [2017 Physician Compare Group Star Rating Cut-offs](#).

### Select one of the topics below to learn more:

- [Why a benchmark?](#)
- [The ABC™ methodology](#)
  - [How will the benchmark be calculated?](#)
- [What about star ratings?](#)
  - [Equal ranges method](#)
- [More ways to learn](#)

## Why a benchmark?

Benchmarks are important for ensuring that patients and caregivers accurately interpret and understand the performance information on Physician Compare. They allow patients and caregivers to best understand the performance information on Physician Compare by setting a point of comparison and providing context.

*Benchmarks help us **interpret and understand** the quality data on Physician Compare by setting a point of comparison.*

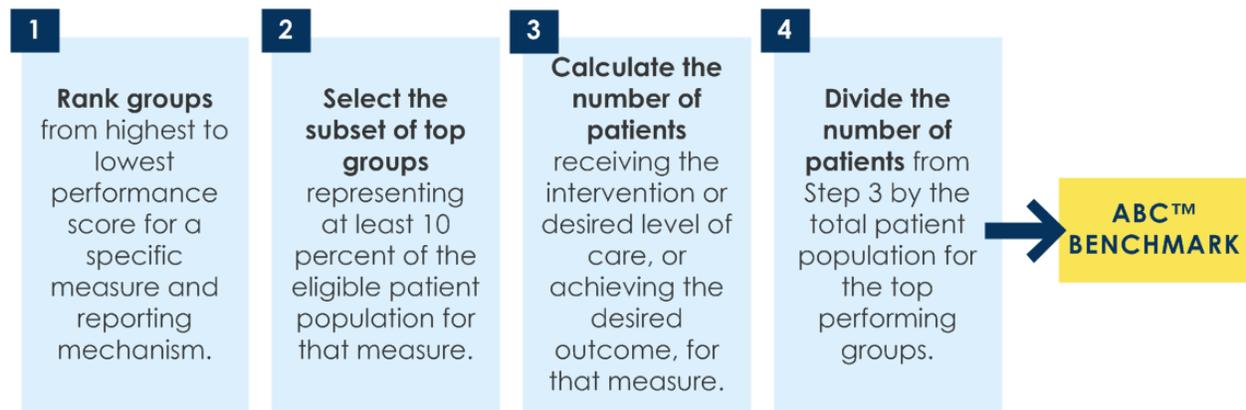
## The ABC™ methodology

Physician Compare uses the Achievable Benchmark of Care (ABC™) methodology to develop the benchmarks that anchor the Physician Compare star rating. The use of the benchmark was first finalized in the CY 2016 Physician Fee Schedule (PFS) final rule ([80 FR 71128 through 71129](#)). ABC™ is a well-tested, data-driven methodology. It represents quality while being realistic and achievable. It also encourages continuous quality improvement and is shown to lead to improved quality of care<sup>123</sup>.

## How will each benchmark be calculated?

An ABC™ is established by measure and collection type. ABC™ starts with the pared-mean. This is the average of the best performers on a measure for at least 10 percent of the patient population – not the population of groups reporting on the measure. Figure 1 provides a step-by-step breakdown of how the benchmark is calculated.

**Figure 1. Benchmark Calculation**



<sup>1</sup> Kiefe CI, Weissman NW, Allison JJ, Farmer R, Weaver M, Williams OD. Identifying achievable benchmarks of care: Concepts and methodology. *International Journal of Quality Health Care*. 1998 Oct; 10(5):443–7.

<sup>2</sup> Kiefe CI, Allison JJ, Williams O, Person SD, Weaver MT, Weissman NW. Improving Quality Improvement Using Achievable Benchmarks for Physician Feedback: A Randomized Controlled Trial. *JAMA*. 2001; 285(22):2871–2879.

<sup>3</sup> Wessell AM, Liszka HA, Nietert PJ, Jenkins RG, Nemeth LS, Ornstein S. Achievable benchmarks of care for primary care quality indicators in a practice-based research network. *American Journal of Medical Quality* 2008 Jan–Feb; 23(1):39–46.

# Physician Compare Benchmark and Star Ratings Fact Sheet

We first rank-order groups from highest to lowest performance score. Next, we include a calculation of a beta binomial model adjustment to account for low denominators. This ensures that very small sample sizes do not over-influence the benchmark but still allows all data to be included in the benchmark calculation.

Then, we create a subset of the groups by selecting the best performers until we have selected enough groups to represent at least 10 percent of all patients relevant for that measure.

We establish a benchmark by calculating mean performance across these top performers. This produces a benchmark that represents the best care provided to at least 10 percent of patients.

For a benchmark to be calculated, the measures must meet our public reporting standards (§414.1395(b)). Each measure must prove to be statistically accurate, valid, and reliable. And, the measure must prove to resonate with patients and caregivers via testing. If these criteria are met, then we calculate the benchmark. The benchmark itself, must also meet our statistical reporting standards.

## What about star ratings?

Physician Compare is committed to moving to star ratings because this is a user-friendly way to share complex information. Star ratings give patients and caregivers more context to best understand the performance scores published on Physician Compare. For example, on some measures, a clinician's raw score of 80% is considered very good relative to other clinicians' performance on that measure at this moment in time. However, without star ratings users may not realize this and assume 80% is just average performance. Star ratings help patients and caregivers accurately evaluate performance scores because these ratings provide a point of comparison.

After we determine the benchmark for a given measure, groups that meet or exceed the benchmark are assigned 5 stars. The next step in moving to star ratings was to decide on a method for assigning 1 to 4 stars. For Physician Compare, we focused on a method that met the following requirements, as requested by stakeholders and encouraged by the many experts consulted:

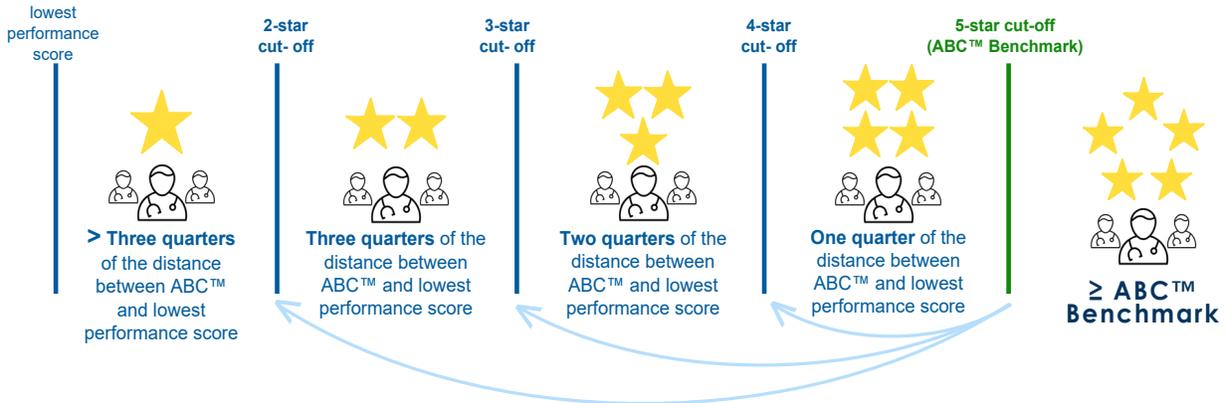
1. Avoids forcing a star-rating distribution;
2. Does not make it hard to achieve a moderate to good rating; and,
3. Reliably assigns groups into a star rating.

As discussed in the CY 2018 Quality Payment Program final rule ([82 FR 53827 through 53829](#)), we conducted extensive statistical analysis, sought the Physician Compare Technical Expert Panel's (TEP) and CMS measure expert input, and reached out to stakeholders including specialty societies and professional groups to help determine the best possible method for assigning 1 to 4 stars. This work led to a decision to use the equal ranges method to assign star ratings, starting with a subset of group-level 2016 Physician Quality Reporting System (PQRS) performance scores that were publicly reported in 2017. Physician Compare continues to use this approach to assign star ratings to a subset of Quality Payment Program performance information.

## Equal ranges method

The equal ranges method is based on the difference between the ABC™ benchmark and the lowest performance score<sup>4</sup> for a given measure. The method uses that range to assign 1 to 4 stars. Groups that meet or exceed the established ABC™ benchmark for a measure will be assigned 5 stars.

**Figure 2. Equal Ranges Method**



To determine the 4-star cut-off, we subtract the lowest performance score from the ABC™ benchmark to get the range of performance scores for that measure, and then divide by 4 to give us quarters. The 4-star cut-off is one quarter of the distance between the ABC™ benchmark and the lowest performance score. Groups that score at or above the 4-star cut-off, but below the benchmark will be assigned 4-stars.

We use the same idea to determine the 3-star cut-off. The 3-star cut-off is two quarters of the distance between the ABC™ benchmark and lowest performance score. Groups that have scores at or above the 3-star cut-off but below the 4-star cut-off are assigned 3-stars.

We follow the same method to get the 2-star cut-off, which is 3 quarters of the distance between the ABC™ benchmark and lowest performance score.

Finally, any scores that are greater than three quarters of the distance between the ABC™ benchmark and the lowest performance score are assigned one star.

## More ways to learn

To learn more about Physician Compare, including more about star ratings and the future of public reporting, visit the [Physician Compare Initiative page](#). Also, sign up for the [Physician Compare eNews](#) to receive updates and latest information about all things Physician Compare.

<sup>4</sup> If the data includes any major outliers, the lowest non-outlier performance score is used to calculate the star rating cut-offs.

# Physician Compare Benchmark and Star Ratings Fact Sheet

If you have questions, please contact the Physician Compare support team at [PhysicianCompare@Westat.com](mailto:PhysicianCompare@Westat.com).