

## 2021 Merit-based Incentive Payment System (MIPS) Quality Measure Benchmarks Overview

**Purpose:** This resource provides an overview of how we establish MIPS quality measure benchmarks, how benchmarks are used for scoring, and the information in the 2021 Quality Benchmarks and 2021 Multi-Performance Rate Measures files.

### Table of Contents

[What Are Quality Measure Benchmarks?](#)

[How Are Benchmarks Established?](#)

[How Are Results Displayed in the Benchmark File?](#)

[What if a Quality Measure Does Not Have a Historical Benchmark?](#)

[Are All Topped Out Measures Capped at 7 Points?](#)

[How Do Benchmarks Work for Multi-Performance Rate Measures?](#)

[Where Can I Find Performance Period Benchmarks?](#)

[Version History](#)

### What Are Quality Measure Benchmarks?

Quality measure benchmarks are the point of comparison we use to score the measures you submit. When you submit measures for the MIPS Quality performance category, your performance on each measure is assessed against its benchmark to determine how many points the measure earns.

- We compare your performance on the measure to its benchmark.
- We assign anywhere from 3 to 10 achievement points for each MIPS measure that meets the data completeness standards and case minimum requirements based on this comparison.
- Measures may also be eligible for bonus points, in addition to these achievement points.

## How Are Benchmarks Established?

We establish benchmarks specific to each collection type: Qualified Clinical Data Registry (QCDR) measures, MIPS clinical quality measures (MIPS CQMs), electronic clinical quality measures (eCQMs), CMS Web Interface measures, the Consumer Assessment of Healthcare Providers and Systems (CAHPS) for MIPS Survey, and Part B claims measures.

### Did you know?

Because benchmarks are specific to collection type, a measure reported as an eCQM will be compared to a different benchmark than the same measure reported as a MIPS CQM.

## eCQMs, MIPS CQMs, QCDR Measures, and Medicare Part B Claims Measures

Whenever possible, we use historical data to establish benchmarks. Historical benchmarks for the 2021 performance period for eCQMs, MIPS CQMs, Medicare Part B claims measures, and QCDR measures are based on actual performance data that was submitted to the Quality Payment Program (QPP) for the 2019 performance period. We won't use data submitted for measures that were suppressed in the 2019 performance period to create historical benchmarks for those measures in the 2021 performance period.

To establish a historical benchmark:

- The 2019 and 2021 measure specifications must be comparable (no significant changes to the measure between 2019 and 2021)
- There must be 20 instances of the measure being reported through the same collection type by individual clinicians, groups and/or virtual groups, AND
  - The clinician, group or virtual group was eligible for MIPS in 2019 (no changes to low-volume threshold for performance year 2021), AND
  - The measure met performance year 2021 data completeness (70%) and case minimum requirements (20 cases), AND
  - The measure had a performance rate greater than 0% (or less than 100% for inverse measures).

We **didn't** finalize our proposal to use performance period benchmarks exclusively for scoring quality measures in the 2021 performance period. Based on our analysis of 2019 submission data, we determined that we have sufficient data to calculate historical benchmarks.

\* We will still attempt to calculate performance period benchmarks for new quality measures, quality measures that lack historical data, or when we do not have comparable data from the baseline period.

## CMS Web Interface Measures

We use benchmarks from the Shared Savings Program to assess and score CMS Web Interface measures. These benchmarks can be found in the [Performance Year 2021 APM Performance Pathway: CMS Web Interface Measure Benchmarks for ACOs](#). (These benchmarks are also used for groups, virtual groups and APM Entities that register to report CMS Web Interface measures for [traditional MIPS](#).)

## CAHPS for MIPS Survey Measure

We established a benchmark for each summary survey measure (SSM) in the CAHPS for MIPS Survey measure. (Refer to the 2021 CAHPS for MIPS Benchmarks file in the [2021 Quality Benchmarks zip file](#).) These benchmarks were calculated using historical data from the 2019 performance period. A range of 3 to 10 points will be assigned to each SSM by comparing performance to the benchmark (similar to other measures). The final CAHPS for MIPS Survey score will be the average number of points across all scored SSMs.

## Administrative Claims Measures

We have added the following 2 new administrative claims measures for the 2021 performance period:

- [\*\*Hospital-Wide, 30-Day, All-Cause Unplanned Readmission Rate for MIPS Groups\*\*](#)
  - *This measure replaces the previous [All-Cause Hospital Readmission Measure](#) that was removed from MIPS beginning with the 2021 performance period.*
- [\*\*Risk-Standardized Complication Rate Following Elective Primary Total Hip Arthroplasty and/or Total Knee Arthroplasty for MIPS\*\*](#)

Because these measures are new to MIPS, historical benchmarks won't be available for the 2021 performance period. Instead, we will attempt to calculate performance period benchmarks.

## How Are Results Displayed in the Benchmark File?

Each benchmark is presented in terms of deciles, with the benchmark file displaying Deciles 3 – 10. [Table 1](#) identifies the range of points generally available for the measure, based on which decile your performance rate falls in.

**Exception:** Measures that are topped out for 2 consecutive years are capped at 7 points, even if your performance rate falls in Deciles 7 - 10.)

### Did you know?

For **inverse measures**, better performance is indicated by a lower performance rate. This is reflected in the benchmark file, where lower performance rates are found in higher deciles.

The 2021 benchmark file also reflects the **flat benchmarks** finalized through previous rulemaking for **Measures 001 and 236**.

- **Measure 001:** Flat benchmarks only apply to the **Medicare Part B Claims measure collection type**. (The MIPS CQM and eCQM collection type did not meet the criteria set forth in the rule for establishing a flat benchmark\*.)
- **Measure 236:** Flat benchmarks apply to **MIPS CQM and Medicare Part B Claims measure collection types**

\*Flat benchmarks are applied to collection types where the top decile for a historical benchmark is greater than 90% (or less than 10% for inverse measures).

**Table 1: Using Data Benchmarks to Determine Achievement Points for Measures that Meet Data Completeness and Case Minimum Requirements**

Decile	Number of Points Assigned for the 2020 MIPS Performance Period
<i>No benchmark (historical or performance period)</i>	3 points
<b><i>Below Decile 3</i></b>	3 points
Decile 3	3-3.9 points
Decile 4	4-4.9 points
Decile 5	5-5.9 points
Decile 6	6-6.9 points
Decile 7	7-7.9 points
Decile 8	8-8.9 points
Decile 9	9-9.9 points
Decile 10	10 points

### Historical Benchmarks with Less Than 10 Deciles

Some benchmarks don't include a range of performance rates for every decile. This occurs when a large percentage of clinicians in the historical benchmark data set have the maximum achievable performance rate. These benchmarks are identifiable when one or more of the

deciles between Decile 3 and Decile 9 display "--" while the Decile 10 is identified at 100% (or 0% for inverse measures). The higher the percentage of individual clinicians, groups, and virtual groups that reach the maximum achievable performance rate, the more deciles that will show a value of "--".

For example, in the benchmark for Measure ID138 (MIPS CQM) presented below, historical benchmarking identified that the top 40% of clinicians performed at the maximum rate. Therefore, clinicians submitting through this collection type that performed above the 6th decile would receive the maximum performance score of 10 points.

**Table 2: Example of a Measure Benchmark with Less than Ten Deciles**

Measure Title	Measure ID	Collection Type	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10
Melanoma: Coordination of Care	138	MIPS CQM	0.2 - 75.5	75.51 - 92.58	92.59 - 98.8	98.81 - 99.99	--	--	--	100

### Did you know?

The **Scoring Examples** tab of the 2021 Benchmark file provides examples for various scoring scenarios.

A	B	C	D	E	F	G	H	I	J
Table 3. Four scoring examples using PY 2021 historical benchmark results.									
Scoring Example 1. Measure 005 ( <i>Heart Failure (HF): ACE Inhibitor or ARB or ARNI Therapy for Left Ventricular Systolic Dysfunction (LVSD)</i> ), collected and reported as MIPS CQM									
Dr. Clark submits data for Measure 005 (MIPS CQM) that results in a performance rate of 96.70% and 5.4 achievement points.									
Why?									
This performance rate falls in Decile 5, which means a measure score of 5.0 - 5.9 points. See formula at right for determining partial points.									
Scoring Example 2. Measure 130 ( <i>Documentation of Current Medications in the Medical Record</i> ), collected and reported as an eCQM									
Dr. Clark submits data for Measure 130 (eCQM) that results in a performance rate of 100% and 7.0 achievement points.									
Why?									
This performance rate falls in Decile 10, which would normally mean a measure score of 10 points. However, it's a topped out measure that is capped at 7 points (see Column R on the MIPS Benchmark Results worksheet).									
Scoring Example 3. Measure 472 ( <i>Appropriate Use of DXA Scans in Women Under 65 Years</i> )									

Scoring Example 1.

Apply the following formula based on the measure performance and decile range:

$$\text{Achievement points} = X + \frac{(q - a)}{(b - a)}$$
$$\text{Achievement points} = 5 + \frac{(96.70 - 95.71)}{(98.01 - 95.71)}$$
$$\text{Achievement points} = 5.4$$

$X$  = decile #  
 $q$  = performance rate  
 $a$  = bottom of decile range  
 $b$  = top of decile range

Note: Partial achievement points are rounded to the tenths digit for partial points between 0.01 to 0.89. Partial achievement points above 0.9 are truncated to 0.9.

$$\frac{(96.70 - 95.71)}{(98.01 - 95.71)} = 0.430434...$$

Which is rounded to 0.4

<	>	Column Descriptions	MIPS Benchmark Results	Scoring Examples	Version History	+	:	<
---	---	---------------------	------------------------	------------------	-----------------	---	---	---

X = decile #  
q = performance rate  
a = bottom of decile range  
b = top of decile range

Note: Partial achievement points are rounded to the tenths digit for partial points between 0.01 to 0.89. Partial achievement points above 0.9 are truncated to 0.9.

You can also check the **Version History** tab for information about changes made to the benchmark file during the performance period.

[Original Posting: 3/18/2021](#)



## What If A Quality Measure Does Not Have A Historical Benchmark?

If a quality measure or collection type doesn't have a historical benchmark, we will attempt to calculate benchmarks based on data submitted for the 2021 performance period. We can establish performance period benchmarks when at least 20 instances of the measure are reported through the same collection type and meet data completeness and case minimum requirements and have a performance rate greater than 0% (or less than 100% for inverse measures).

Performance period benchmarks will be established using data submitted by individual clinicians, groups, and virtual groups that are eligible for MIPS in the 2021 performance period.

- This includes individual clinicians and groups that are opt-in eligible and elect to opt-in to MIPS participation.
- Voluntary submissions are excluded from benchmark data.

F
Measure has a Benchmark
N

If no historical benchmark exists and no performance period benchmark can be calculated, then the measure will receive 3 points as long as data completeness and case minimums have been met.

Measures/collection types without **historical** benchmarks display “N” (for “NO”) in the “Measure has a Benchmark” column (Column F).

## Are All Topped Out Measures Capped At 7 Points?

No. A measure is capped at 7 points when it is topped out through the same collection type for 2 consecutive years. The 7-point cap is applied in the second year the measure is identified as topped out.

**A measure may be topped out without being capped at 7 points.** A “Y” (for “YES”) in the **Seven Point Cap** column (column Q) of the benchmark file indicates the measure is capped at 7 points.

**Example 1. Measure ID 128, Preventive Care and Screening: Body Mass Index (BMI) Screening and Follow-Up Plan (Medicare Part B Claims)**

Topped Out	Seven Point Cap
Y	N

Even though it's topped out, it's not capped at 7 points.

A maximum of 10 achievement points is available for the measure.

# Quality Payment PROGRAM

## **Example 2.** Measure ID 130, **Documentation of Current Medications in the Medical Record** (all collection types)

A maximum of 7 achievement points is available for the measure, even if your performance rate is found in Deciles 7 – 10.

Topped Out	Seven Point Cap
Y	Y

### **Did you know?**

The benchmark file displays the range of performance rates associated with Deciles 7 – 10, even though scoring is capped at 7 points.

## **How Do Benchmarks Work for Multi-Performance Rate Measures?**

Several MIPS quality measures and QCDR measures require the collection and submission of data for multiple populations. This means that there can be multiple performance rates associated with a single measure.

- Historical benchmarks for multi-performance rate measures are created based on an "overall performance rate" (based on a weighted average, simple average or CMS-specified performance rate).
- When you are scored on a multi-performance rate measure, we will compare the "overall performance rate" of your submitted measure to the measure's benchmark which is also based on the "overall performance rate".

The [2021 Multi-Performance Rate Measure file](#) identifies the method used to determine the "overall performance rate" for each multi-performance rate measure.

- It is NOT intended to specify an additional performance rate that must be submitted. Measures should be submitted according to their specification.
- Only multi-performance rate QCDR measures allow for the submission of an "overall performance rate".

This file also provides an example for each of the 3 methods for determining the overall performance rate. (Click the tabs at the bottom of the file.)

2021 Multi-Perf Rate Measures	BACKGROUND	<b>WEIGHTED AVG EXAMPLE</b>	SIMPLE AVG EXAMPLE	SPECIFIC PERF RATE EXAMPLE
-------------------------------	------------	-----------------------------	--------------------	----------------------------

[Original Posting: 3/18/2021](#)



## Where Can I Find Performance Period Benchmarks?

We do not publish performance period benchmarks on the QPP Resource Library because the results are specific to the 2021 performance period. When you report a measure (or measures) without a historical benchmark, you should review your final performance year 2021 performance feedback in July 2022. If a performance period benchmark can be calculated, you can view your score(s) based on the comparison of your performance to the performance period benchmark.

## Where You Can Go for Help

- Contact the Quality Payment Program at 1-866-288-8292, Monday through Friday, 8:00 AM-8:00 PM ET or by e-mail at: [QPP@cms.hhs.gov](mailto:QPP@cms.hhs.gov).
  - Customers who are hearing impaired can dial 711 to be connected to a TRS Communications Assistant.
- Connect with your [local technical assistance organization](#). We provide no-cost technical assistance to small, underserved, and rural practices to help you successfully participate in the Quality Payment Program.
- Visit the Quality Payment Program [website](#) for other [help and support](#) information, to learn more about [MIPS](#), and to check out the resources available in the [Quality Payment Program Resource Library](#).

## Version History

If we need to update this document, changes will be identified here.

Date	Change Description
3/18/2021	Original posting.

[Original Posting: 3/18/2021](#)