Quality Payment PROGRAM

Merit-based Incentive Payment System (MIPS)

2024 Quality Benchmarks User Guide

Table of Contents

How To Use This Guide ---- 3

Overview ---- 5

- What's New with 2024 Benchmarks? ---- 6
- What are Benchmarks? ---- 9
- How are Benchmarks Established? ---- 9

Scoring Measures Against a Benchmark ---- 13

- When Are Measures Scored Against A Benchmark? ---- 14
- What if a Measure Doesn't Have A Historical Benchmark? ---- 15
- How Do I Know If A Measure Doesn't Have A Historical Benchmark? ---- 16
- How Are Measures Scored Against a Benchmark? ---- 17
- Benchmarks with Less than 10 Deciles ---- 18

Scoring Examples ---- 19

Frequently Asked Questions ---- 24

Help and Version History ---- 27

Purpose: This resource focuses on Merit-based Incentive Payment System (MIPS) quality benchmarks, providing high level information and scoring examples for the 2024 MIPS performance period.

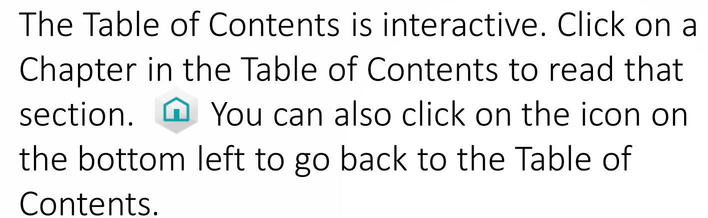
Already know what MIPS is? Skip ahead by clicking the links in the Table of Contents.

How to Use This Guide

How to Use This Guide

Please Note: This guide was prepared for informational purposes only and isn't intended to grant rights or impose obligations. The information provided is only intended to be a general summary. It isn't intended to take the place of the written law, including the regulations. We encourage readers to review the specific statutes, regulations, and other interpretive materials for a full and accurate statement of their contents.

Table of Contents



The Table of Contents is interactive. Click on a Chapter in the Table of Contents to read that section. You can also click on the icon on the bottom left to go back to the Table of Contents.

Hyperlinks

Hyperlinks to the [Quality Payment Program website] are included throughout the guide to direct the reader to more information and resources.

Overview

What's New with 2024 Benchmarks?

- The benchmark file is now accessible through a new webpage on the QPP website instead of stored in a ZIP file: https://qpp.cms.gov/benchmarks.
- When you download the file, it pulls the benchmark data in real time from an Application Programming Interface (API) used for scoring. **The file you download is current as of the date included in the file name.**
- The benchmark file will download as a comma-separated values (CSV) file, which is similar to the Excel format we've used in the past. You can save the benchmark file in the Excel format, and then add filters so you can use the data in the same way you always have. (See next page)
- We've added the CMS ID for electronic clinical quality measures (eCQMs).
- The column will list N/A for any measure that isn't an eCQM.

C
CMS eCQM ID
N/A
N/A
CMS122v12
N/A

- We've added an indicator of the benchmark type so we can add performance period benchmarks to this same file in summer 2025; we believe this will provide a more complete picture of quality measure benchmarks to the QPP community.
- You'll see "--" when no benchmark is available. Measures that show "--" when historical benchmarks are released will be eligible for a performance period benchmark.

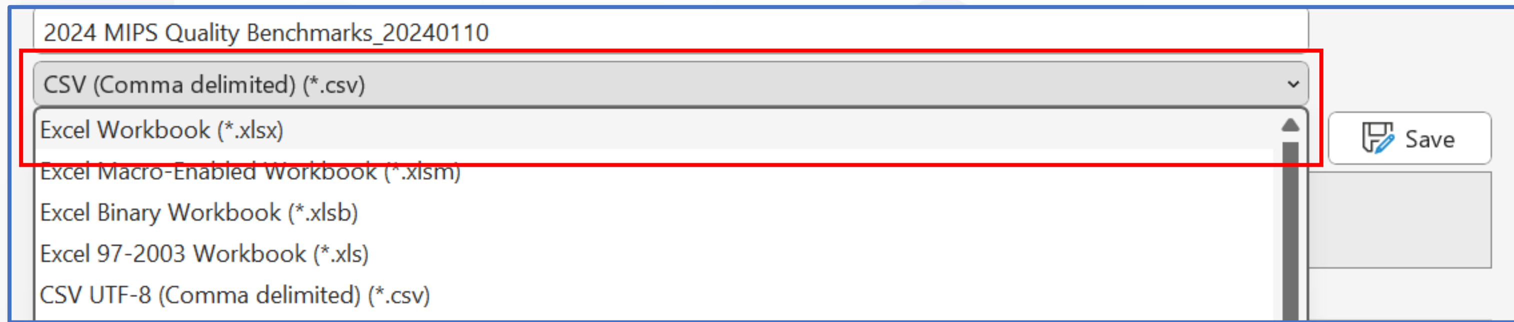
I
Benchmark Type
Historical
Historical
Historical
--
Historical
--
Historical

- The 2024 benchmark file includes benchmarks for the CAHPS for MIPS Summary Survey Measures; these are no longer shared in a separate Excel file.

What's New with 2024 Benchmarks?

How to save a CSV file in Excel

1. In your Excel worksheet, click **File** > **Save as**.
2. Browse for the folder on your computer where you want to save the file.
3. To save as an Excel file, select Excel Workbook (*. xlsx) from the Save as type drop-down menu. To save as a comma-separated file, select CSV (Comma delimited) or CSV UTF-8.
4. Click Save.



What's New with 2024 Benchmarks?

How to add filters to columns in Excel

1. Highlight the top row in the file.
2. Select **Data** > **Filter**.

What Are 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.

Measures that can be scored against a benchmark generally earn between 1 and 10 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 measures, and Medicare Part B claims measures.

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

- Whenever possible, we use historical data to establish benchmarks. Historical benchmarks for the 2024 performance period for eCQMs, MIPS CQMs, QCDR measures, and Medicare Part B claims measures are based on actual performance data that were submitted to the Quality Payment Program (QPP) for the 2022 performance period. We won't use data submitted for measures that were suppressed in the 2022 performance period to create historical benchmarks for those measures in the 2024 performance period.

What Are Benchmarks? (Continued)

eCQMs, MIPS CQMs, QCDR Measures, and Medicare Part B Claims Measures (Continued)

To establish a historical benchmark:

- The 2022 and 2024 measure specifications must be comparable (no significant changes to the measure between 2022 and 2024).
- 20 instances of the measure must have been reported in 2022 through the same collection type by individual clinicians, groups, virtual groups, and/or Alternative Payment Model (APM) Entities AND
  - The clinician, group, or virtual group was eligible for MIPS according to 2024 eligibility criteria (no changes to low-volume threshold from the 2022 performance year), AND
  - The measure met performance year 2024 data completeness (75%) and case minimum requirements (20 cases), AND
  - The measure had a performance rate greater than 0% (or less than 100% for inverse measures).

What Are Benchmarks? (Continued)

CAHPS for MIPS Survey Measure

We establish benchmarks for each scored summary survey measure (SSM) in the CAHPS for MIPS Survey measure.

- These benchmarks are included in the benchmark file that can be downloaded on the Benchmarks page of the QPP website.
- These benchmarks were calculated using historical data from the 2022 performance period.

A range of 1 to 10 points will be assigned to each SSM by comparing performance against the benchmark (similar to other measures).

The final CAHPS for MIPS Survey score will be the average number of points across all scored SSMs.

CMS Web Interface Measures (Medicare Shared Savings Program Only)

We use benchmarks from the Medicare Shared Savings Program (Shared Savings Program) to assess and score CMS Web Interface measures.

These benchmarks are available on the QPP Resource Library.

Reminder: CMS Web Interface measures are available only to Shared Savings Program Accountable Care Organizations (ACOs) reporting the APM Performance Pathway (APP).

What Are Benchmarks? (Continued)

Administrative Claims Measures

There are up to 4 administrative claims measures available in the 2024 performance period depending on your MIPS reporting option.

- Clinician and Clinician Group Risk-standardized Hospital Admission Rates for Patients with Multiple Chronic Conditions
- Hospital-Wide, 30-Day, All-Cause Unplanned Readmission Rate for MIPS Groups
- Risk-Standardized Complication Rate Following Elective Primary Total Hip Arthroplasty and/or Total Knee Arthroplasty for MIPS
- Risk-Standardized Acute Cardiovascular-Related Hospital Admission Rates for Patients with Heart Failure under the Merit-based Incentive Payment System

We score administrative claims measures exclusively against performance period benchmarks.

Scoring Measures Against a Benchmark

When Are Measures Scored Against a Benchmark?

Quality measures are scored against a benchmark when **all 3 criteria** are met:

1. The measure meets data completeness criteria (75% threshold for the 2024 performance period).
2. The measure meets case minimum criteria (generally 20 cases).
3. A benchmark exists for the measure's collection type.

What If a Measure Doesn't Meet Data Completeness Criteria?

Quality measures that don't meet data completeness criteria aren't eligible for scoring against a benchmark, even if a benchmark exists, and will receive 0 points unless:

- The measure is submitted by a small practice; small practices will continue to earn 3 points for measures that don't meet data completeness criteria.

What If a Measure Doesn't Meet Case Minimum Criteria?

Quality measures that don't meet case minimum aren't eligible for scoring against a benchmark, even if a benchmark exists, and will receive 0 points unless:

- The measure is submitted by a small practice; small practices will continue to earn 3 points for measures that don't meet case minimum criteria.
- The measure is in its 1st or 2nd year in the program; these measures will earn 7 and 5 points, respectively, if data completeness is met. (This policy doesn't apply to administrative claims measures.)

What If a Measure Doesn't Have a Historical Benchmark?

If a quality measure's collection type doesn't have a historical benchmark, we'll attempt to calculate a benchmark based on data submitted for the 2024 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 2024 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.

If no historical benchmark exists and no performance period benchmark can be calculated, then the measure will receive 0 points unless:

- The measure is submitted by a small practice; small practices will continue to earn 3 points for measures without a benchmark.
- The measure is in its 1st or 2nd year in the program; these measures will earn 7 and 5 points, respectively, if data completeness is met. (This police doesn't apply to administrative claims measures.)

How Do I Know If A Measure Doesn't Have A Historical Benchmark?

Measures/collection types without historical benchmarks display "No" in the "Measure has a Benchmark" column **(Column H).**

H	I
Measure has a Benchmark	Benchmark Type
Yes	Historical
Yes	Historical
Yes	Historical
No	--
Yes	Historical
No	--

Column V in the benchmark file indicates why there's no historical benchmark for a measure/collection type.

V
Reason for No Benchmarks
--
Insufficient volume of data submitted in PY 2022 to establish historical benchmark.
Measure added in PY 2023; subject to 5-point scoring floor if data completeness is met.
Measure added in PY 2024; subject to 7-point scoring floor if data completeness is met.
Measure was suppressed in PY 2022 (baseline period); data isn't available for historical benchmarking.
N/A
Substantive changes to specification in PY 2023; PY 2024 measure can't be compared to baseline period (PY 2022) measure.
Substantive changes to specification in PY 2024; PY 2024 measure can't be compared to baseline period (PY 2022 measure.

How Are Measures Scored Against a Benchmark?

Each benchmark is presented in terms of deciles; the benchmark file displays Deciles 1 – 10 for each measure. Table 1 identifies the range of points generally available for the measure, based on which decile your performance rate falls into.

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

Your Performance Rate Falls in the Range of This Decile*	Number of Points Assigned for the 2024 Performance Period
No benchmark (historical or performance period)	0 points (small practices will continue to receive 3 points)
Decile 1	1 – 1.9points
Decile 2	2 – 2.9 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

***Exception:** Measures that are topped out for 2 consecutive years are capped at 7 achievement points, even if your performance rate falls in Deciles 7 – 10. The benchmark file still displays values for Deciles 7 – 10 even though the measure can't earn more than 7 achievement points.

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 1 and Decile 9 display “--" while 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 results for the Diabetes: Eye Exam measure (Measure ID 117, MIPS CQM) presented in Table 2, historical benchmarking identified that the top 60% of clinicians performed at the maximum rate. Therefore, clinicians submitting through this collection type who performed above the 4th decile would receive the maximum performance score of 10 points.

Table 2: Example of a Measure Benchmark with Less than 10 Deciles

J	K	L	M	N	O	P	Q	R	S
Decile 1	Decile 2	Decile 3	Decile 4	Decile 5	Decile 6	Decile 7	Decile 8	Decile 9	Decile 10
10.26 - 79.99	80.00 - 96.45	96.46 - 99.20	99.21 - 99.99	--	--	--	--	--	100

Scoring Examples

Example 1.

Measure 009 (CMS128v12) submitted as an eCQM

Dr. Johnson submits data for Measure 009 (eCQM) that results in a performance rate of 75.28% and 3.4 points.

Why?
This performance rate falls in Decile 3, which means a measure score of 3.0 – 3.9 points. See formula below for partial point calculation.

Scoring Example 1.

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

$$\text{Achievement points } = X + \frac{(q-a)}{(b-a)}$$

$$\text{Achievement points } = 3 + \frac{(75.28-74.32)}{(77.03-74.32)}$$

$$\frac{(75.28-74.32)}{(77.03-74.32)} = 0.35424\ldots$$

Which is rounded to 0.4

$$\boldsymbol{Achievement\ points\ = 3.4}$$

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

Note: This scoring example assumes data completeness and case minimum have been met.

Example 2.

Measure 317 submitted as Medicare Part B Claims measure

Dr. Johnson submits data for Measure 317 (Medicare Part B Claims measure) that results in a performance rate of 100% and 7.0 achievement points. This scoring example assumes data completeness and case minimum have been met.

Why?

- This performance rate falls in Decile 10, which would generally mean a measure score of 10 points.
- However, it's a topped out measure that is also capped at 7 points.

Column T identifies measures that are topped out.

Column U identifies measures that are capped at 7 points.

T	U
Topped Out	Seven Point Cap
Yes	Yes
N/A	N/A
Yes	No
No	No

Note: Measures that are topped out for 2 consecutive years for a specific collection type are capped at 7 achievement points, even if your performance rate falls in Deciles 7 – 10. The benchmark file still displays values for Deciles 7 – 10 even though the measure can't earn more than 7 achievement points.

Example 3.

Measure 005 submitted as a MIPS CQM

Dr. Johnson submits data for Measure 005 (MIPS CQM) that results in a performance rate of 40.52% and 1.0 achievement point. This scoring example assumes data completeness and case minimum have been met.

Why?
This performance rate is below Decile 1 and earns the 1-point floor for measures that can be scored against a benchmark.

Example 4.

Measure 238 submitted as a MIPS CQM

Dr. Johnson submits data for Measure 238 (MIPS CQM) that results in a performance rate of 99.99% and 0.0 achievement points at the point of submission. This scoring example assumes data completeness and case minimum have been met.

Why?
There is no historical benchmark for Measure 238 (submitted as a MIPS CQM). See Column V on the 2024 Quality Benchmarks file to identify the reason for no historical benchmark. Measures without a benchmark will earn 0 points – 3 points for a small practice – unless a performance period benchmark can be created for use in PY 2024. We'll attempt to create a performance period benchmark following the data submission period. If we can create one based on submission data, the measure will be eligible for up to 10 points (provided that data completeness and case minimum are met).

Exception: This policy doesn't apply to measures in their 1st or 2nd year in the program.

Example 5.

Measure 498 submitted as a MIPS CQM

Dr. Johnson submits data for Measure 298 (MIPS CQM) that results in a performance rate of 52.99% and 7.0 achievement points at the point of submission. This scoring example assumes data completeness and case minimum have been met.

Why?
There's a 7-point scoring floor for measures added to the program in PY 2024.

If a performance period benchmark can be created, Dr. Johnson would be eligible to earn 7 – 10 points based on their performance in comparison to the performance period benchmark. See Column V is in the 2024 Quality Benchmarks file to identify which measures newly added to the program in 2024.

Example 6.

Measure 289 submitted as a MIPS CQM

Dr. Johnson submits data for Measure 289 (MIPS CQM) that results in a performance rate of 41.99% and 5.0 achievement points at the point of submission. This scoring example assumes data completeness and case minimum have been met.

Why?
There's a 5-point scoring floor for measures in their 2nd year of the program in PY 2024.

If a performance period benchmark can be created, Dr. Johnson would be eligible to earn 5 – 10 points based on their performance in comparison to the performance period benchmark. See Column V is in the 2024 Quality Benchmarks file to identify which measures in their 2nd year of the program in 2024.

Frequently Asked Questions

Which Measures Have Flat Benchmarks?

The 2024 benchmark file also reflects the **flat benchmarks*** finalized through previous rulemaking for **Measures 001 and 236**.

- **Measure 001:** We established **flat benchmarks** for the **MIPS CQM and Medicare Part B claims collection types** .
  - We established a performance-based benchmark for the eCQM collection type.
- **Measure 236:** We established flat benchmarks for the **MIPS CQM and Medicare Part B claims measure collection types.**
  - There's no historical benchmark for the eCQM collection type as this measure was suppressed for the 2022 performance period.

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

How Do Benchmarks Work for Inverse Measures?

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, such as Measure 001.

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 (or more) consecutive years. The 7-point cap is applied in the 2nd year the measure is identified as topped out.

A measure may be topped out without being capped at 7 points. A "Yes" in the Seven Point Cap column (column U) of the benchmark file indicates the measure is capped at 7 points.

Example 1. Measure ID 047, Advance Care Plan (MIPS CQM)

Even though it's topped out, it's not capped at 7 points because it wasn't topped out last year. It's not in its 2nd consecutive year of being topped out.

A maximum of 10 achievement points is available for this measure

T	U
Topped Out	Seven Point Cap
Yes	No

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

This measure has been topped out for at least 2 consecutive years. A maximum of 7 achievement points is available for the measure, even if your performance rate is found in Deciles 7 – 10.

The benchmark file still displays values for Deciles 7 – 10 even though the measure can't earn more than 7 achievement points.

T	U
Topped Out	Seven Point Cap
Yes	Yes

Help and Version History

Where Can You Go for Help?

Contact the Quality Payment Program Service Center by email at QPP@cms.hhs.gov, by creating a QPP Service Center ticket, or by phone at 1-866-288-8292 (Monday through Friday, 8 a.m. - 8 p.m. ET). To receive assistance more quickly, please consider calling during non-peak hours—before 10 a.m. and after 2 p.m. ET.

- People who are deaf or hard of hearing can dial 711 to be connected to a TRS Communications Assistant.

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.

Visit the Small Practices page of the Quality Payment Program website where you can **sign up for the monthly QPP Small Practices Newsletter** and find resources and information relevant for small practices.

Version History

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

Date	Description
01/25/2024	Original Posting.