

Benchmarks and Quality Measures

Posted on July 15, 2025

View this essay online:

<https://academic-master.com/benchmarks-and-quality-measures/>

Introduction

Benchmarks and quality measures are central tools in healthcare improvement because they convert broad goals such as safety, effectiveness, timeliness, equity, and patient-centered care into information that can be tracked and compared. A quality measure defines a specific aspect of care, such as a process completed, an outcome achieved, a patient experience reported, or a structural capability present. A benchmark provides a reference point for judging performance. The distinction matters: a hospital can calculate a readmission rate as a quality measure, but the number becomes more informative when it is compared with prior performance, national data, peer organizations, or an evidence-based target. CMS continues to use quality measures for quality improvement, public reporting, accountability, and payment programs, while AHRQ maintains national tools and benchmark datasets that allow organizations and states to examine trends across hundreds of measures (CMS, 2026a; AHRQ, 2024). The value of benchmarking, however, depends on comparability. Organizations can draw the wrong conclusion when definitions, populations, coding rules, time periods, risk adjustment, or data completeness differ. Quality measurement therefore requires as much attention to methodology as to the final score.

Quality Measures

Healthcare quality measures generally describe structure, process, outcomes, patient experience, or related dimensions of care. Structure measures examine whether an organization has resources or systems associated with high-quality care, such as appropriate staffing, technology, or certified programs. Process measures assess whether recommended actions occurred, for example whether eligible patients received a vaccination or a preventive screening. Outcome measures examine what happened to patients, such as mortality, readmission, infection, functional improvement, or complications. Patient-reported and experience measures capture information that may not be visible in claims or clinical records, including symptom burden, communication, respect, and care coordination. CMS defines quality measures as tools for quantifying healthcare processes, outcomes, patient perceptions, and organizational systems linked to quality goals (CMS, 2026a).

No single measure can represent the quality of an entire hospital, clinic, health plan, or clinician. A low mortality rate may be important but says little about communication, access, medication safety, or whether care was equitable. Likewise, a strong process score does not guarantee a favorable

clinical outcome if the process measure captures only one step in a complex pathway. A balanced measurement program should therefore reflect the purpose of the improvement effort. If the problem is preventable infection, the organization may need process measures for adherence to evidence-based practices, outcome measures for infection rates, and balancing measures for unintended consequences. If the goal is patient access, waiting time and appointment availability may matter more than a general quality score. The measure should follow the clinical question rather than the organization selecting measures merely because data are easy to obtain.

Healthcare Benchmarks

A benchmark can be internal or external. Internal benchmarking compares current performance with the organization's own previous results, which is useful for determining whether an intervention produced improvement. External benchmarking compares performance with other organizations, national averages, high-performing peers, or established targets. CMS describes quality performance benchmarks as standards against which the quality of care can be evaluated, and value-based programs may connect performance against benchmarks with payment (CMS, 2023). AHRQ publishes benchmark tables for its Quality Indicators, including Inpatient Quality Indicators and Patient Safety Indicators, so organizations can compare observed performance with national reference data using defined specifications (AHRQ, 2025a; 2025b).

The best benchmark is not automatically the highest possible score. A useful benchmark should be relevant to the population and purpose. A small rural hospital should be cautious about comparing a specialized outcome directly with a large academic referral center that treats substantially different patients. Similarly, comparing one hospital's unadjusted mortality rate with another institution's risk-adjusted rate would be misleading. Benchmarking should help an organization understand whether performance is unusual and where improvement may be possible, not create a simplistic league table. The strongest comparisons use consistent definitions and similar populations, and they acknowledge uncertainty when case numbers are small.

Measure Specifications

Every quality measure needs a precise numerator, denominator, inclusion criteria, exclusion criteria, time period, data source, and calculation method. Small differences in these elements can change a result substantially. For example, a readmission measure may include only certain diagnoses, exclude planned readmissions, use a defined follow-up period, and apply risk adjustment. If another organization calculates "readmissions" using a different population or time frame, the percentages cannot be interpreted as equivalent. AHRQ recommends using rigorously specified measures when broad comparisons are intended because precision in definitions makes results more credible across entities and locations (AHRQ, 2024b).

Data compatibility is equally important. Healthcare information may come from claims, electronic health records, registries, laboratories, surveys, public-health databases, or manual abstraction. Each source has strengths and limitations. Claims data are often standardized and available across large populations but may lack clinical detail. Electronic health records contain richer clinical information but can vary in coding, documentation, and extraction methods. Patient surveys capture experiences that administrative data cannot measure, yet response rates and sampling methods can influence results. An organization should therefore know exactly how the data were generated before comparing its score with a benchmark. The presence of the same measure name does not guarantee methodological equivalence.

Reliability and Risk Adjustment

Reliable measures produce reasonably consistent results when the same phenomenon is measured under similar conditions. Reliability can be poor when the sample is too small, data are incomplete, or outcomes are rare. Validity asks whether the measure actually represents the quality concept it claims to measure. A measure can be calculated perfectly and still be misleading if the underlying construct is weak. For example, using one documentation code as a proxy for a complex clinical outcome may produce a precise number that does not accurately reflect patient care. Organizations should therefore consider both technical performance and clinical meaning.

Risk adjustment is often necessary for outcome comparisons because patient populations differ in age, disease burden, severity, comorbidities, and other characteristics. Without appropriate adjustment, organizations treating sicker patients may appear to provide worse care even when they perform well. At the same time, risk adjustment should not hide disparities or normalize poor outcomes for disadvantaged groups. AHRQ's National Healthcare Quality and Disparities Report tools allow users to examine performance by population groups, geography, insurance, income, and other characteristics, helping organizations distinguish overall averages from inequities within the population (AHRQ, 2024a). Quality measurement should therefore ask both "How are we performing?" and "For whom are we performing well or poorly?"

Quality Improvement

Measurement is useful only when it leads to action. A healthcare organization should begin with a problem, select measures that represent the problem and desired outcome, establish a baseline, compare performance with an appropriate benchmark, and then investigate the processes that may explain the gap. If performance is worse than the benchmark, the next step is not immediately to blame clinicians. The difference may reflect workflow problems, data quality, patient characteristics, access barriers, staffing, documentation, or a genuine clinical-performance issue. Improvement teams should examine the underlying process and test changes while continuing to monitor the measure over time.

Public reporting can also influence improvement by making comparative performance visible to patients, purchasers, regulators, and competing organizations. The related Academic Master discussion of [public reporting](#) shows why transparent measures can support accountability while also creating risks if data are poorly specified or easily misunderstood. CMS has attempted to reduce measure burden and inconsistency through efforts such as the Core Quality Measures Collaborative and the Cascade of Meaningful Measures, which aim to align measures around important priorities and reduce unnecessary duplication (CMS, 2026b; 2026c). Measure alignment matters because clinicians and organizations can spend substantial time reporting slightly different versions of similar concepts to different payers.

Benchmarking Strategy

A practical benchmarking strategy should start with a limited set of high-value measures rather than collecting every available metric. Each measure should have a clear owner, purpose, data source, calculation method, reporting frequency, and action plan. Organizations should confirm that the comparison benchmark uses compatible specifications and should review whether the benchmark remains current. A national average may be useful for broad context, while a top-decile benchmark may be more appropriate for an improvement target. For some measures, the organization's own best historical performance may provide the most actionable initial goal.

Teams should also monitor balancing measures so that improvement in one area does not create harm elsewhere. Reducing length of stay, for example, may appear efficient but could be problematic if readmissions or post-discharge complications rise. Increasing screening rates may be beneficial but could burden patients if follow-up systems are inadequate. Quality improvement therefore requires interpretation rather than mechanical pursuit of a number. Benchmarks can identify a gap, but clinicians and managers still need to understand the system that produced it.

Digital Quality Measurement

Healthcare quality measurement is moving toward greater use of structured electronic data and interoperable standards. CMS's measure-development infrastructure increasingly favors digital specifications and Fast Healthcare Interoperability Resources (FHIR), with current planning indicating stronger FHIR requirements in coming years (CMS MERIT, 2026). Digital measures could reduce manual abstraction and make reporting more timely, but automation will not solve poor definitions or inconsistent source data. Organizations still need data governance, validation, and clinical review to ensure that extracted information represents actual care.

The future of benchmarking will also depend on improving measure relevance. Too many overlapping measures can create reporting burden without improving care, while narrow measures can encourage organizations to optimize what is counted rather than what matters to patients. A more mature

approach uses a focused set of clinically meaningful measures, combines outcomes with patient experience and equity, and selects benchmarks that reflect the intended population. Technology should make measurement easier, but the central question remains unchanged: does the information help the organization deliver better care?

Conclusion

Benchmarks and quality measures are valuable because they transform healthcare performance into information that can be compared, investigated, and improved. Quality measures define what is being assessed, while benchmarks provide the reference point that gives the result context. Their usefulness depends on methodological discipline: consistent definitions, compatible populations, reliable data, appropriate risk adjustment, adequate sample size, and transparent interpretation. A strong measurement program uses a balanced set of structure, process, outcome, and patient-centered measures rather than treating one score as a complete judgment of quality. It also examines disparities and unintended consequences. Current CMS and AHRQ initiatives reflect a wider movement toward measure alignment, digital reporting, and more meaningful benchmarking. Healthcare organizations should therefore view measurement not as a reporting obligation but as a decision-support system. When measures are carefully chosen and benchmarks are genuinely comparable, the data can identify where performance differs, guide improvement work, and help leaders determine whether changes are producing safer and more effective care.

References

- Agency for Healthcare Research and Quality. (2024a). *National Healthcare Quality and Disparities Report data tools*.
- Agency for Healthcare Research and Quality. (2024b). *Choosing quality measures*.
- Agency for Healthcare Research and Quality. (2025a). *Inpatient Quality Indicators benchmark data tables, v2025*.
- Agency for Healthcare Research and Quality. (2025b). *Patient Safety Indicators benchmark data tables, v2025*.
- Centers for Medicare & Medicaid Services. (2023). *Benchmarking*. CMS Innovation Center.
- Centers for Medicare & Medicaid Services. (2026a). *Quality measures*.
- Centers for Medicare & Medicaid Services. (2026b). *Core measures*.
- Centers for Medicare & Medicaid Services. (2026c). *Cascade of Meaningful Measures*.

Centers for Medicare & Medicaid Services MERIT. (2026). *Measures Under Consideration Entry/Review Information Tool guidance*.