

January 29, 2020

The Honorable Seema Verma
Administrator
Centers for Medicare & Medicaid Services
U.S. Department of Health and Human Services
Hubert H. Humphrey Building
200 Independence Avenue, SW
Washington, DC 20201

Re: CMS-9915-P; Transparency in Coverage Proposed Rule

Dear Administrator Verma:

On behalf of our physician and medical student members, the American Medical Association (AMA) appreciates the opportunity to provide comments on the Administration's proposals to increase health care price transparency. At a time when patients are facing increasing financial burdens due to health care costs and the health care system is facing ever-growing financial strain, it is critical that patients are provided with transparent, easily accessible information about expected health care costs. However, information about health care procedures and their pricing can be complex and confusing to even the most health care-literate patients, making it critical that any transparency mandates focus only on the information most relevant to consumers and ensure this information is easy to understand.

AMA on Price Transparency

The lack of timely, standardized information about the cost of health care services has long prevented health care markets from operating efficiently and has impacted the ability of patients to manage their health care costs. Changes in available health care plans have resulted in patients increasingly assuming greater financial responsibility for care choices, and thereby increasing the demand for better information about anticipating out-of-pocket health care costs. Further, as health care markets evolve, patients are increasingly becoming active consumers of health care services. Achieving meaningful price transparency can help lower health care costs and empower patients to make more informed care decisions. The AMA supports price transparency efforts and recognizes that promoting disclosure of meaningful health care cost information can help to ensure patients choose lower-cost, higher-quality care.

As part of our support of price transparency initiatives, the AMA has recommended the following measures to expand the availability of health care pricing information that allow patients and their physicians to make value-based decisions when patients have a choice of provider or facility:

- Patient confusion and health literacy should be addressed by developing resources that help patients understand the complexities of health care pricing and encourage them to seek information regarding the cost of health care services they receive or anticipate receiving;
- All health care professionals and entities should be required to make information about prices for common procedures or services readily available to consumers;

- Physicians should readily communicate information about the cost of their professional services to individual patients, taking into account the insurance status of the patient (e.g., self-pay, in-network insured, out-of-network insured) when possible;
- Health plans should provide plan enrollees or their designees with complete information regarding plan benefits and real-time, cost-sharing information associated with both in-network and out-of-network provider services or other plan designs that may affect patient out-of-pocket costs;
- Health plans, public and private entities, and other stakeholder groups should work together to facilitate price and quality transparency for patients and physicians;
- Entities promoting price transparency tools should have processes in place to ensure the accuracy and relevance of the information they provide;
- All-payer claims databases should be supported and strengthened; and
- Electronic health records (EHR) vendors should include features that assist in facilitating price transparency for physicians and patients.

We welcome the Administration's increased focus on price transparency and are pleased to see alignment between some of the proposals put forward and AMA's recommendations above.

Proposed Requirements for Disclosing Cost-Sharing Information to Participants, Beneficiaries, or Enrollees

The AMA is pleased to see the Administration propose requirements of health plans to provide patients tools allowing them to obtain an estimate of their potential cost-sharing liability for covered items and services. We are further pleased to see the inclusion of information about the coverage status of a particular item or service, prerequisites to coverage, coverage limitations, and prescription drug pricing included in this proposal. We believe it aligns strongly with the AMA's recommendations for expanding the availability of health care cost information.

As discussed above, the availability of meaningful health care cost information is critical to ensure that patients can make fully informed health care decisions. **Access to this information, with accompanying information regarding the quality of the provider or facility in question, is essential to ensure that patients are able to make value-based choices when presented with options of either providers or facilities.** The AMA urges physicians to comply with requests for relevant pricing information from patients whenever possible, especially in instances of self-pay patients. However, in the case of insured patients, we strongly believe that health plans are best situated to provide this information to patients and support the Administration's proposal to require plans to make this information available to their enrollees.

Health plans are uniquely situated and stand alone as the one entity in the health care market with all relevant health care pricing and benefit information for an individual patient enrolled in a particular plan. Because plans are the only entity with ready access to all relevant information regarding a patient's coverage, cost sharing, current deductible status, and coverage limitations, as well as information about a provider's network status, we believe they should be the primary source for providing patient health care cost information. **We support the proposal to require plans to make this information available to patients via a patient self-service tool, so long as that tool is easily accessible and displays actionable information to patients in a manner that is understandable and easy to read.**

While we support proposals to require health plans to make this information available to patients, we caution that information about health care services and procedures can be complex and difficult to

understand for even the most health-literate patients. Benefit information is frequently confusing, necessitating that patients understand the difference between information such as deductibles and co-pays, in-network and out-of-network providers, and covered and non-covered services. Additional information, such as Current Procedural Terminology (CPT) codes, procedure/service terminology, coverage limitations or prerequisites to coverage, while necessary to display accurate pricing information, adds further complexity to the information displayed to patients. Because of the high risk for confusion among patients when providing this information, we urge the Administration to require disclosure and display of only that information most meaningful to patients—mainly information that helps patients easily understand what their estimated out-of-pocket costs may be. Additional information, such as display of rates negotiated with in-network providers, does nothing to help patients understand their potential financial liability and seek only to add to confusion already experienced by many patients in trying to navigate our complex health care system. **We strongly urge the Administration to ensure the information mandated for inclusion in these tools is limited to only that which is essential for determining patient out-of-pocket costs.** We also suggest that the Administration and health plans consider the utility of providing plan enrollees additional educational materials or functionalities to help them better understand health care cost information and serve as a guide to the information displayed in the proposed self-service tools.

Request for Information: Disclosure of Pricing Information through a Standards-Based API

The AMA supports the proposed requirement that plans and issuers provide patients with access to pricing information through an application programming interface (API). We agree with CMS that patients should have the ability to decide how their information will be used by consumer-facing apps, and we include ways CMS can incentivize app developers, plans, and issuers to keep patient health information private. Furthermore, we suggest the following:

- CMS should also require plans and issuers to provide prior authorization requirements to patients and physicians.
- While physicians must provide information to patients free-of-charge, CMS has not indicated that the same requirement applies to plans and issuers. It is unclear who will absorb the associated costs. CMS should also consider the impact to physician offices' (e.g., health information technology costs) if they elect to incorporate this information into their health record or practice management systems.
- CMS should ensure that beneficiaries and the individuals assisting them should have assurances that information provided across settings (e.g., online web portals, smartphone apps, plans' and issuers' policy booklets, etc.) contain consistent information.

Privacy

We wholeheartedly appreciate CMS' acknowledgement that consumers "should be able to share [...] data with third-party applications of their choosing, but that they should understand that they are accepting the potential privacy and security risks that come from using a third-party application that is not required to comply with the HIPAA Rules." Stories and studies abound about how smartphone apps share sensitive health information with third parties, often without the knowledge of an individual.¹ If beneficiaries

¹ Citations to these examples are provided in the body of the letter: (1) The Wall Street Journal reported that Facebook collected sensitive health and demographic data from a user's cellphone apps, regardless of whether the individual had the Facebook app

access their personal data or their family's health data—some of which are likely sensitive—through a smartphone, a patient should have a clear understanding of the potential uses of that data by app developers. Otherwise, most patients will not be aware of who has access to their medical information, how and why they received it, and how it is being used (for example, an app may collect or use information for its own purposes, such as an insurer using health information to limit/exclude coverage for certain services, or may sell information to clients such as to an employer or a landlord). The downstream consequences of data being used in this way may ultimately erode a patient's privacy and willingness to disclose information to his or her physician.

To assist in preventing this scenario, the AMA has identified an opportunity for CMS to empower patients with meaningful knowledge and control over how apps use their health data.

CMS should require that plans' and issuers' APIs check an app's attestation to:

- **Industry-recognized development guidance** (e.g., Xcertia's Privacy Guidelines²);
- **Transparency statements and best practices** (e.g., Mobile Health App Developers: FTC Best Practices³ and CARIN Alliance Code of Conduct⁴); and
- **A model notice to patients** (e.g., U.S. Office of the National Coordinator for Health Information Technology's [ONC's] Model Privacy Notice⁵).

The app could be acknowledged or listed by the API developer in some special manner (e.g., in an "app store," "verified app" list). We would urge CMS to encourage plans and issuers to limit app listings to those apps that have replied "yes" to all three attestations or, at the very least, provide those apps with a special designation on the plan's and issuer's website.

We recognize that a "yes" attestation would not ensure apps implement or conform to their attestations. However, app developer attestations would be a powerful resource for the Federal Trade Commission (FTC) in its enforcement of unfair and deceptive practices. In other words, an app developer would be strongly motivated to attest "yes" and to act in line with their attestations. We do not believe that requiring an API check for an app developer attestation would be a significant burden on health IT developers. We also specifically note that this proposal does not ask CMS to regulate apps or app developers; rather, it regulates the type of API technology that plans and issuers must adopt.

CMS can implement this requirement even if ONC does not since CMS' proposal does not require plans and issuers to use Health IT Modules certified by ONC. We firmly believe these sorts of "checks" on an app will provide a needed level of assurance to patients and would be greatly welcomed by users.

on his or her phone, and even if the individual had never signed-up for Facebook. (2) Studies reported in the British Medical Journal and Journal of the American Medical Association have demonstrated that most apps do not share privacy policies with patients, and when they do, sometimes do not adhere to them. (3) The Washington Post reported that a workplace wellness pregnancy-tracking app reports data to a woman's employer, including the woman's average age, number of children and current trimester; the average time it took her to get pregnant; whether the pregnancy is high-risk, conceived after a stretch of infertility, a C-section or premature birth; and her return-to-work timing. The app's privacy notice is 6,000 words.

² <https://xcertia.org/app-privacy-survey/>

³ <https://www.ftc.gov/tips-advice/business-center/guidance/mobile-health-app-developers-ftc-best-practices>

⁴ <https://www.carinalliance.com/our-work/trust-framework-and-code-of-conduct/>

⁵ <https://www.healthit.gov/sites/default/files/2018modelprivacynotice.pdf>

Privacy Controls in FHIR

CMS is proposing to require plans and issuers to implement open APIs consistent with the API technical standards proposed by ONC in 84 Fed. Reg. 7424 (March 4, 2019), using content and vocabulary standards that include those proposed by ONC at 45 CFR 170.213. These standards include the use of HL7 Fast Healthcare Interoperability Resources (FHIR) referenced in 45 CFR §170.299. FHIR supports data controls like segmentation; however, we are concerned those controls are an afterthought in FHIR-based API design and will become “bolt-on” functions—drastically increasing their costs and limiting their usefulness. The AMA has been told that FHIR developer efforts are first focused on “just making the technology work” and that “patient data protections and privacy controls are outside their scope.” The downstream consequences of this approach will negatively impact physicians and patients. Developers need to address privacy concerns and incorporate privacy considerations as a part of the development process of any new technology. Mechanisms to monitor and control data access, patient consent and privacy, and ensure data provenance, governance, and enforce state and federal law must be inherent in FHIR development.

In December 2019, it was reported that CMS discovered a bug in its Blue Button 2.0 API exposing the protected health information of around 10,000 Medicare beneficiaries. This resulted in access to the Blue Button API being suspended while the CMS completed a comprehensive code review. CMS determined the anomaly was due to a coding bug. That bug potentially allowed data to be shared with incorrect Blue Button 2.0 applications and the wrong beneficiaries. CMS determined 30-50 applications had been impacted by the bug.

The Blue Button platform is used by Medicare beneficiaries to authorize third-party applications, services, and research programs to access their claims data. A CMS identity management system verifies user credentials through a randomly generated unique user ID, which ensures the correct beneficiary claims data is shared with the correct third-party applications. The investigation so far has determined that the coding error began on January 11, 2018, and CMS officials determined a comprehensive review of the system was not completed, which could have detected this coding error earlier.

While CMS is not specifically recommending or requiring that plans and issuers utilize the Blue Button 2.0 API, lack of transparency and appropriate vetting of exchange and access tools will impact the trust between patients, physicians and the overall health system. This should serve as a warning to policy makers to take privacy very seriously. **Privacy safeguards must be established concurrently with any health information exchange policies—particularly when consumer-facing technology is implicated—or patient health and safety will be jeopardized.**

Patient Access to Information Through APIs

HIPAA provides individuals with a right to access the enrollment, payment, claims adjudication, and case or medical management records maintained by or for a health plan.⁶ The AMA fully supports a patient’s right of access and agrees with CMS that a patient should be able to easily receive this information from his or her health plan. Nevertheless, we note CMS’ expectation that beneficiaries will use this information to shop for care and more effectively manage costs may not be realistic for many patients. Claims information can be complex and erroneous, and patients have varying levels of health and technology literacy. Medicare populations, in particular, often need assistance navigating the complexity of the

⁶ 45 CFR 164.501 (designated record set) and 45 CFR 164.524(a)(1).

system. As such, **beneficiaries and the individuals assisting them should have assurances that information provided across settings (e.g., online web portals, smartphone apps, payer policy booklets) contain consistent information.**

Additionally, the AMA supports the following specific measures to expand the availability of health care pricing information that allows patients and their physicians to make value-based decisions when patients have a choice of provider or facility:

- Health plans should provide plan enrollees or their designees with complete information regarding plan benefits and real-time, cost-sharing information associated with both in-network and out-of-network provider services or other plan designs that may affect patient out-of-pocket costs;
- Entities promoting price transparency tools should have processes in place to ensure the accuracy and relevance of the information they provide; and
- Patient confusion and health literacy should be addressed by developing resources that help patients understand the complexities of health care pricing and encourage them to seek information regarding the cost of health care services they receive or anticipate receiving.

We include below other ways we believe CMS can build upon its proposals to provide patients with information to improve their care.

Prior Authorization Information

In addition to providing beneficiaries with adjudicated claims data and pricing information through an API, plans and issuers should provide beneficiaries (and prospective beneficiaries) and providers with information about restrictions on covered medications and services including when prior authorization (PA) or step-therapy is required. An AMA-convened workgroup of 17 state and specialty medical societies, national provider associations, and patient representatives developed best practices for prior authorization and other utilization management requirements by identifying the 21 most common provider and patient concerns (the Principles). Over 100 additional stakeholders have signed-on in support of the Principles, one of which states: “Utilization review entities should publicly disclose, in a searchable electronic format, patient-specific utilization management requirements, including prior authorization, step therapy, and formulary restrictions with patient cost-sharing information, applied to individual drugs and medical services. Such information should be accurate and current and include an effective date in order to be relied upon by providers and patients, including prospective patients engaged in the enrollment process. Additionally, utilization review entities should clearly communicate to prescribing/ordering providers what supporting documentation is needed to complete every prior authorization and step therapy override request.”⁷

To the extent that CMS has committed to utilizing the Da Vinci project both to communicate PA requirements and (in another use case) exchange data to complete PA decisions, all plans and issuers should be required to align with the project. As always, having various plans and issuers use different processes for PA exponentially increases PA burdens, so if CMS is fully committed to Da Vinci for their PA process, other plans and issuers should use this technology. **CMS should require all plans and issuers to select a single API for Da Vinci/coverage discovery and should urge other plans and**

⁷ Prior Authorization and Utilization Management Reform Principles, available at <https://www.ama-assn.org/sites/ama-assn.org/files/corp/media-browser/principles-with-signatory-page-for-slsc.pdf>

issuers to follow suit so that each plan and issuer does not have its own PA API. Multiple PA APIs would be costly and burdensome for physicians to support.

Additional Uses of APIs

APIs have enormous potential to provide valuable data to both patients and physicians. The AMA supports plans and issuers using APIs to provide complete information related to specific individuals or populations. However, there is potential for plans and issuers to include extraneous information originating from other entities which could be misinterpreted by patients or physicians as plan- or patient-specific. We are aware of instances where EHR vendors have utilized direct-to-physician marketing within the physician's workflow. These "ads" may provide benefit to patients but can also inappropriately influence physician prescribing. It is unclear the extent plans and issuers could pull data from other entities (e.g., pharmaceutical companies or pharmacy benefit managers) and seamlessly combine disparate sources into one information feed. Given the flexibility APIs provide to create a conduit of information into an EHR, physicians and patients could eventually be exposed to plan and issuer information comingled with marketing information—obscuring the information's author or impacting the information's trustworthiness. The AMA encourages CMS to study the effects of direct-to-physician advertising at the point-of-care, including advertising in EHRs, on physician prescribing, patient safety, health care costs, and EHR access.

Proposal to Require Public Disclosure of Negotiated Rate Information

As discussed above, the AMA strongly supports efforts to increase health care price transparency and urges the Administration to move forward with efforts that would help patients better understand their expected out-of-pocket costs. However, proposals to require public disclosure of information that is not relevant to determinations of patient out-of-pocket costs raises several concerns for physicians. **We strongly urge the Administration to consider the likelihood that proposals requiring disclosure of rates negotiated between in-network physicians and health plans will not have the expected impact on prices and could carry serious unintended consequences, such as negative impacts on price, competition, and patient access to care.**

Requiring the disclosure of negotiated in-network rates would have significant anti-competitive effects harming patient access to, and the quality of, health care without any countervailing consumer benefit. As the FTC has noted, while transparency can often benefit consumers and foster competition, too much price transparency can "harm competition in any market, including health care markets."⁸ The FTC has identified specific areas where transparency can benefit consumers: treatment options; available providers in the consumer's area; providers' quality performance; provider charges; out-of-pocket expenses; and explanations of benefits clearly stating what the patient owes after the insurer has paid. But as the FTC also notes, transparency with regard to some kinds of information is not particularly helpful to consumers but is "of great interest to competitors," and the FTC is "especially concerned when information disclosures allow competitors to figure out what their rivals are charging, which dampens each competitor's incentive to offer a low price, or increases the likelihood that they can coordinate on higher prices."⁹

⁸ Price Transparency or TMI, Tara Koslov and Elizabeth Jex, Office of Public Planning, U.S. Federal Trade Commission, July 2, 2015, accessible at <https://www.ftc.gov/news-events/blogs/competition-matters/2015/07/price-transparency-or-tmi>.

⁹ Id.

While the FTC’s concerns in this particular instance were focused on providers’ access to competitors’ charges, i.e., the so-called “sell side” of the health care market, the same concerns apply to transparency initiatives that would enable health insurers to figure out what their competitors are paying providers on the “buy side” of the market. Just as providers can coordinate in ways that lead to higher prices, so too can health insurers coordinate in ways that depress provider payment rates below competitive levels. This is particularly the case in health insurance markets, many of which are highly concentrated, dominated by only a few insurers that already exercise monopsony or “buyer power” in those markets.¹⁰

Insurers use of buyer power to force physician practice payment rates below competitive levels can hurt patient access to, and the quality of, care by hindering physicians’ ability to invest in new equipment, technology, training staff and other practice infrastructure. Inability to adequately fund practice infrastructure can also compromise the U.S. health care system as a whole since practice infrastructure investment, typically borne by physician practices, is necessary to participate successfully in value-based payment initiatives. Anti-competitive compensation rates can also force physicians to spend less time with patients to meet practice expenses.

The threat that depressed physician payment rates may have on patient access to, and quality of, patient care has been recognized by a number of enforcement agencies, who understand that consumers do best when there is a competitive market for physician services. For example, the U.S. Department of Justice (DOJ) has successfully challenged two health insurer mergers based in part on DOJ claims that the mergers would have anticompetitive effects in the purchase of physician services. These challenges occurred in the merger of Aetna and Prudential in Texas in 1999,¹¹ and the merger of UnitedHealth Group Inc. and Pacific Care in Tucson, Arizona and in Boulder, Colorado in 2005.¹² In a third merger matter occurring in 2010—Blue Cross Blue Shield of Michigan and Physicians Health Plan of Mid-Michigan—the health insurers abandoned their merger plans when the DOJ complained that the merger “...would have given Blue Cross Michigan the ability to control physician payment rates in a manner that could harm the quality of healthcare delivered to consumers.”¹³ In their complaint against the Anthem-Cigna merger, the DOJ and numerous state attorneys general included a claim that the merger would lead to below-market rate payments for physician and hospitals and a consequent reduction in patients’ access to medical care—a monopsony claim.¹⁴

While the context of the proposed rule does not involve health insurer mergers, the negative effects of in-network payment rate disclosure on competition could at least be as great as those resulting from an

¹⁰ For the past 18 years, the AMA has conducted the most in-depth annual study of commercial health insurance markets in the country. The AMA’s study, *Competition in Health Insurance: A Comprehensive Study of US Markets* (2019 update) utilizes the 2010 iteration of the U.S. Department of Justice (DOJ) and Federal Trade Commission (FTC) Horizontal Merger Guidelines to classify markets based on whether mergers announced in those markets would raise anticompetitive concerns. According to the 2019 update to the AMA’s *Competition in Health Insurance: A Comprehensive Study of U.S. Markets*: shows that there has been a near collapse of competition in commercial, combined HMO + PPO + POS markets. Seventy-five percent (285) of the 382 Metropolitan Statistical Areas (MSAs) commercial markets studied were highly concentrated. Further, in 91% (348) of MSAs, at least one insurer held a commercial market share of 30% or greater, and in 48% (182) of MSAs, one insurer’s share was at least 50%. *Competition in Health Insurance: A Comprehensive Study of US Markets* (2019 update) is accessible at <https://www.ama-assn.org/delivering-care/patient-support-advocacy/competition-health-insurance-research>.

¹¹ U.S. v. Aetna Inc., supra note 12, at ¶¶ 17-18; see also U.S. v. Aetna, Inc., No. 3-99 CV 1398-H, at 5-6 (Aug. 3, 1999) (revised competitive impact statement).

¹² United States v. United Health Group Inc. No. 1:05CV02436 (D.D.C., Dec. 20, 2005) (complaint), available at: www.usdoj.gov/atr/cases/f213800/213815.htm.

¹³ Blue Cross Blue Shield of Michigan and Physicians Health Plan of Mid-Michigan Abandon Merger Plans | OPA | Department of Justice, available at: <https://www.justice.gov/opa/pr/blue-cross-blue-shield-michigan-and-physicians-health-plan-mid-michigan-abandon-merger-plans>.

¹⁴ Accessible at <https://www.justice.gov/atr/file/903111/download>.

anticompetitive merger. Just as disclosure of provider charges could result in anti-competitive effects due to coordination among competing health care providers, so too could the disclosure of the payment rates a provider has agreed to accept from an insurer incentivize anti-competitive coordination among competing health insurers. For example, absent mandatory in-network payment disclosure, several health insurers in a given market are likely to pay physician practices at different rates, with some obviously being greater than others. However, making all payment rates publicly available would alert each insurer in the market to what its competitors were paying a physician practice, and most importantly, to the lowest in-network payment amounts that the physician practice accepts. This knowledge could motivate all other competing insurers to demand from the practice those lowest rates, resulting in a “race to the bottom,” where payment may become uniformly fixed at the lowest common denominator.

Any purported benefits from disclosing in-network payment rates fall far short of justifying the potential patient harm resulting from public disclosure, particularly since patients can be fully informed about their out-of-pocket financial obligation resulting from selecting a particular provider to furnish a specific item or service without irrelevant and superfluous publication of additional confusing information.

The AMA also agrees with a number of hospital and health plan stakeholders that compelling public disclosure of privately negotiated rates could raise significant constitutional concerns. As the Administration is aware, health care providers and health plans rely on the confidentiality of negotiated rates to allow them to negotiate at arm’s length and maintain competition in the marketplace. Requiring public disclosure of these confidential negotiations may potentially carry First Amendment and other constitutional implications. Further, privately negotiated rates are widely considered to be confidential trade secrets, which are generally protected from disclosure under federal, state, and common law.

Proposal to Require Public Disclosure of Historical Allowed Amount Data for Covered Items and Services from Out-of-Network Providers

While the AMA has serious concerns about public disclosure of rates negotiated between health plans and in-network providers, we support making available information about allowed amounts paid to out-of-network providers. The AMA has long advocated for such transparency in out-of-network coverage, recognizing that patients must have full knowledge of the facts to make informed decisions concerning the health insurance coverage they purchase and where, and from which providers, they seek health care services. Central to making an informed decision is, of course, understanding the amount that an out-of-network physician will charge for providing a medical service and how that compares to their health insurance’s out-of-network coverage. As such, we encourage physicians to volunteer fee information to patients and to discuss their out-of-network fees in advance of services. But only when health plans clearly disclose the scope and limitations of any out-of-network benefit (i.e., their out-of-network allowable) in language that is meaningful to the average consumer, will consumers (1) be able to shop intelligently for health insurance, and (2) be assured that the higher premiums they pay to make out-of-network care affordable actually reflect the value of the out-of-network benefit actually provided.

Request for Information: Provider Quality Measurement and Reporting in the Private Health Insurance Market

- 1. Whether, in addition to the price transparency requirements the Departments propose in these rules, the Departments should also impose requirements for the disclosure of quality information for providers of health care items and services.**

The AMA encourages the use of physician data to benefit both patients and physicians, and to improve the quality of patient care and the efficient use of resources in the delivery of health care services. The AMA supports the use of physician data when they are used in conjunction with program(s) designed to improve or maintain the quality of, and access to, medical care for all patients, and are used to provide accurate physician performance assessment. There are both challenges and risks associated with a quality public reporting mandate. Such a mandate is especially challenging when applied to specialty and surgical care, which encompasses many specialties, each of which has unique characteristics based on patient needs and often involves services that are not “shoppable.” Therefore, we caution the respective agencies against expeditiously moving forward with a requirement to disclose quality information for physician and provider health care items and services. Any new requirement must be developed in close consultation with physician specialty societies. Specialty societies must have input regarding physician measures chosen for public reporting, and opportunities to participate on workgroups or panels selecting measures for reports and/or disclosure. Specialty societies must also be given the opportunity to review and correct data before it is displayed.

Quality measures were not developed to be paired with every charge a provider bills and would require some sort of bundling of services for the information to be meaningful, and not misleading, to consumers. Quality is also currently defined by payers using primarily performance indicators that are most often process measures. Some of these measures are not fully supported by evidence and others focus on specific aspects of care that are not applicable to a particular charge. An additional challenge to using quality measures at the physician level is ensuring reliability of the performance score, because most individual physicians do not have a large enough case minimum to reliably publicly report quality information. Criteria on volume must be included in any inter-agency requirement to ensure that the publicly reported data are statistically significant and yield scores that accurately and meaningfully inform patient choice. Therefore, performance measures need to improve if they are to be used for ranking and public profiling. Misuse of measures could lead to misrepresentation of the quality of care. Therefore, to ensure accuracy of information and buy-in by the physician community, any process requires consultation with physicians and patients, working through their specialty societies and with quality measurement experts, who are the most qualified to define meaningful definitions of clinical excellence.

Private payers must also be prohibited from designing and utilizing proprietary measures, which make it impossible to audit or for providers to decipher the information. Methodologies for scoring/ranking must be made publicly available and should include a detailed description of any data used to estimate performance (i.e., the data source and specifications), use of statistical risk-adjustment techniques, the selection of performance measures, and how clinical and surgical performance was categorized. Information should also be transparent about the observation period for a given quality measure, including the differentiation between long-term follow-up and short-term outcomes. There is also a need to ensure proper clinical and social risk adjustment, as determined by the appropriate specialty society, to ensure ongoing access for patients who are at higher risk of complications and poor outcomes.

The AMA cautions against a move toward a requirement that is dependent on administrative claims quality measures. Inconsistency in definitions of terms and insufficient granularity in coding makes the data less reliable and meaningful for quality reporting. There are also no standardized methods for attribution of providers' roles in patients' episodes of care. Due to the poor attribution methodologies, the data are often not actionable by the entity being measured because of retrospective analysis and utilizing claims data for secondary purposes. In addition,

there is emerging evidence that administrative measures may also be leading to unintended consequences, such as increase in mortality.^{15,16,17}

Furthermore, while the disclosure of quality information is an opportunity for patients to proactively seek out high-quality care, any publicly required quality information should not be used to establish the standard of care or duty of care owed by a health care provider to a patient in any medical malpractice action or claim. The medical liability system is already overburdened by meritless litigation, and it is important that such publicly disclosed quality information is not misused to create new standards of care for medical liability lawsuits.

Of note, CMS has taken an iterative approach to expanding Physician Compare and has instituted high reliability standards when publicly reporting and displaying data through star ratings. Therefore, we recommend the agencies employ the same process and standards with any requirements related to quality information and to not create an entire new program, which would greatly increase administrative burden. In addition, if the agencies move forward there must be an opportunity for physicians to preview their data and file an appeal prior to publicly posting, such as an informal review similar to Physician Compare.

2. Whether health care provider quality reporting and disclosure should be standardized across plans and issuers or if plans and issuers should have the flexibility to include provider quality information that is based on metrics of their choosing, or state-mandated measures.

There is a need for some form of standardization and at the same time local flexibility. The key is the right balance. Otherwise, too much standardization leads to “teaching to the test” and runs the risk of interfering with the doctor-patient relationship, including engaging in shared decision making. Any standardization across plans must be evidence-based, and have physician input and endorsement.¹⁸ We would encourage any standardized measures to be based on recommendations from the AHIP-CMS Core Quality Measure Collaborative (CQMC).

We are also concerned with the potential for increased administrative burden given physicians’ experience related to Medicare Advantage (MA) Star Ratings Program. As the Star Ratings program has expanded and played a larger financial role on health plans’ bottom lines, the administrative demand on physicians has simultaneously increased and is impeding clinical care. A large percentage of the measures within the MA Star Ratings program is based completely on physician action, compliance and communication. In order for health plans to increase their Healthcare Effectiveness Data and Information Set® (HEDIS) scores and earn greater incentives from CMS, plans are requiring practices as part of their

¹⁵ Gupta, Ankar, et al. Association of the Hospital Readmissions Reduction Program Implementation With Readmission and Mortality Outcomes in Heart Failure. *JAMA Cardiol.* 2017. doi:10.1001/jamacardio.2017.4265. Published online November 12, 2017.

¹⁶ AMA Letter to CMS Administrator Verma- Potential Association between CMS Hospital Readmissions Reduction Program and Increased Mortality Outcomes. Sent Feb. 8, 2018. https://searchlf.ama-assn.org/undefined/documentDownload?uri=%2Funstructured%2Fbinary%2Fletter%2FLETTERS%2F2018-2-8-Letter-to-Verma-re-HRRP-and-Mortality-v4_ACM.pdf

¹⁷ PROPUBLICA. The Family Wanted A Do Not Resuscitate Order, The Doctors Didn’t. Dec. 31, 2018. <https://www.healthleadersmedia.com/clinical-care/family-wanted-do-not-resuscitate-order-doctors-didnt>. Accessed Jan. 17, 2019.

¹⁸ MacLean CH, Kerr EA, Qaseem A. Time Out—Charting a Path for Improving Performance Measurement. *The New England Journal of Medicine.* 2018.

clinical data submission requirements to submit data on all patient lab results and tests, and the plans state it is due to the Star Ratings and HEDIS requirements.

Many of the measures, particularly the HEDIS Effectiveness of Care measures, have more to do with physician quality than assessment of a health plan. The Effectiveness of Care measures are really targeting clinical quality, which is a physician or facility issue—and therefore physicians and facilities have the data. In addition, the patient experience ratings are heavily based on Health Plan Consumer Assessment of Healthcare Providers & Systems (CAHPS) that focus on physician communication and behavior. While communication between a physician and patient is important, asking the questions in a de-identified survey does not lead to quality improvement or address potential challenges patients experience when seeking care. Similar questions are also in the hospital and clinician-group CAHPS survey and are the more appropriate avenues for addressing provider communication in the context of patient experience. Without a better focus, the MA ratings program is just one more burden on physicians and does not provide beneficiaries with the information they need to determine the most appropriate and high quality MA or drug plan. Therefore, if the Departments move forward with any sort of requirement on health plans, we recommend the following:

- Allow for more general exclusions for patients with specific conditions, comorbidities or allergies from measures to ensure patient and clinical differences are accounted for and do not interfere with clinical decision-making;
- Denominators of quality measures should be appropriately defined to ensure patients for whom the treatment may not be appropriate are excluded from measurement;
- Consult and work with Agency for Healthcare Research and Quality (AHRQ) to update the Health Plan CAHPS survey to consider barriers to access, such as the intentional design of narrow networks. The current survey focuses heavily on physician; communication and behavior and is duplicative of Hospital and Clinician-Group CAHPS
- Prohibit the use of administrative claims measures, proprietary measures or algorithms; and
- Provide the ability to preview and appeal data before assessments are made and publicly reported.

3. What type of existing quality of health care information would be most beneficial to beneficiaries, participants, and enrollees in the individual and group markets? How can plans and issuers best enable individuals to use health care quality information in conjunction with cost-sharing information in their decision making before or at the time a service is sought?

The AMA recognizes the importance of incorporating a patient's voice and engaging in shared decision-making and supports the use of Patient Reported Outcome (PRO) measures when appropriate. However, we very much caution the Department of Health and Human Services (HHS) against moving forward with a single question or brief survey to measure the quality of patient experience and satisfaction due to the diversity among physician practice settings and specialties. Patient experience encompasses the range of interactions that patients have with the health care system, including their care from health plans, doctors, nurses, staff in hospitals, physician practices, and other health care facilities. Patient satisfaction is related to whether a patient's expectations about a health encounter were met. When the CMS Physician Quality Reporting Initiative (PQRI) program started, the AMA through the PCPI® extensively explored methods by which universally applicable patient satisfaction measures could be identified and developed. However, we found universal measures to be difficult to define in a way that clearly links to measuring an outcome. Therefore, the agencies working with CMS should invest time and money in expanding PRO types of measures rather than a one-size-fits-all measure approach. We would not support

mandatory adoption of a single measure that was then aggregated and scored across all physicians as each physician specialty and subspecialty is unique.

In addition, patient experience, while important, does not always correlate with better clinical outcomes and may even conflict with clinically indicated treatments. For example, a physician who recommends that a patient lose weight, stop smoking, or limit pain medications, is likely to receive a low “performance” score, even when these are clinically indicated. Therefore, tying a measure and score related to CAHPS to publicly reported ratings and accountability can be problematic, as CAHPS often depends more on patient perceptions than on good medicine. In addition, we believe CAHPS survey administration protocols are outdated and a need exists to allow for measures that use multiple modes of data collection. Allowing physicians to collect the information in the office through a tablet while the patient is in the waiting room, via smartphone app, for example, is needed. The broader the patient population physicians can reach, the more likely they are to receive good response rates.

HHS/CMS needs to also look outside of CAHPS for tools to measure patient experience, such as the CollaboRATE tool/measure. CollaboRATE is a patient-reported measure of shared decision making, which contains three brief questions that patients, their parents, or their representatives complete following a clinical encounter. The CollaboRATE measure provides a performance score representing the percentage of adults 18 and older who experience a high level of shared decision making.

Generally, it appears that collecting open-ended questions may provide valuable feedback to physicians for quality improvement purposes, but we believe it is premature to move to require publicly posting patient narratives in conjunction with CAHPS for the Merit-Based Incentive Payment System (MIPS) survey data or other surveys. Narratives can be helpful at times but context is needed to understand what happened during the patient encounter. We also believe additional research is needed to explain the potential reasons for variations across patients, especially more complex and sick patients.¹⁹ Therefore, due to the subjective nature of narratives and lack of testing, we do not see how the agencies or private plans could score the information under MIPS, Physician Compare and/or public facing website(s).

Furthermore, Clinician-group CAHPS or CAHPS for MIPS was designed to assess the group as most of the questions ask about the group as a whole so it is not appropriate to evaluate individual clinicians based on CAHPS for MIPS or Hospital-CAHPS. The survey is also geared towards assessing primary care groups and does not address patient experience related to surgical care. Therefore, goals of data collection related to patient experience must include ease of requesting feedback from patients and families, minimization of collection burden, reduction of cost to practices and increased response rates. Until HHS and CMS can achieve these goals and ensure capturing patient experience does not increase burden and results are reliable and valid, we urge HHS to not move forward with assessing individual clinicians and mandating private plans to assess and publicly post physician patient experience surveys.

4. Would it be feasible to use health care quality information from existing CMS quality reporting programs, such as QPP or the Quality Measure Inventory (QMI) for in-network providers in the individual and group markets?

¹⁹ We are aware of AHRQ conducting analysis on patient narratives but it was performed on an extremely small sample, so it is premature to make a generalizable statement and for HHS or CMS to move to implement the patient narratives in a national program. It also remains unclear how the data will be used because posted protocol on the CAHPS database on AHRQ’s website states that AHRQ has no plans at this time to accept submissions on patient narratives, nor is there an explanation offered by AHRQ on how narratives will be assessed/scored and by whom.

As part of the MIPS quality reporting, physicians are required to submit on all-payer data (except if reporting through claims), but not all physicians who accept Medicare or are in-network with a private plan participate in MIPS. In addition, some measures, such as the population health and cost measures are strictly derived from Medicare Part B claims and not representative of their private payer patients. However, we frequently hear from physicians that the all-payer data requirement is extremely time-consuming due to the amount of data entry required. It also ignores the fact that physicians are still contractually obligated to meet various other private payer quality initiatives and are assessed using different data and quality measures or program requirements. If their MIPS quality data could be potentially used to satisfy their private payer program requirements and obligations, then physicians might see the value in reporting on all patients, regardless of payer and feel comfortable with their CMS QPP data being used to satisfy private payer requirements. However, any usage of QPP data for private payer markets should not be mandatory and physicians and providers must have the ability to opt-in.

In addition, as physicians are measured on more outcome and intermediate outcome measures, their payer mix and associated patient population will affect scoring. The AMA performed a query of 2016 National Ambulatory Medical Care Survey (NAMCS) data, looking at blood pressure control rates and at a national level the rates differ by insurance types.

Insurance	Visits
All sources of payment are blank	85.5%
Unknown	67.0%
Private insurance	66.9%
Medicare	75.8%
Medicaid, CHIP or other state-based program	70.2%
Worker's compensation	49.8%
Self-pay	65.7%
No charge/Charity	N/A
Other	80.1%
OVERALL	71.8%

While combing all-payer data would increase sample sizes, any requirement would need to take into consideration how payer mix and benefit design impact outcomes. At a minimum, CMS must improve risk adjustment methodologies and risk adjust at the measure level prior to requiring and publicly reporting data.

QPP Scoring Methodologies

We are also concerned with the disconnect between QPP scoring for payment incentive purposes and for public reporting and ranking, as well as the methodology used to create quality measure benchmarks under MIPS. As we have repeatedly highlighted to CMS, MIPS awards points to physicians based on their performance relative to decile-based categories calculated from historical data (when available), and Physician Compare utilizes an achievable benchmark methodology (ABC) based on five stars. Therefore, the 5-star ratings have 5 levels of quality and MIPS creates nine levels of quality (e.g., deciles 10 through 3 and “less than 3”). Using the MIPS methodology, topped-out measures may result in fewer than nine levels, since there may not be 10 fully-defined deciles. For example, if 50 percent of physicians score 100 percent, then the 6th through 10th deciles are the same, and the MIPS deciles will appear for the top

decile (the 10th, namely the 50 percent who scored 100 percent), and for deciles 5, 4, 3, and less than 3. Through our examination (see Appendix A), we have found that the two methodologies (MIPS and 5-star) resulted in inconsistent ratings and comparisons. In several instances, physicians deemed to be of similar quality by one methodology were classified as having different levels of quality by the other methodology. Additionally, some physicians classified in the highest (or lowest) level of quality by one methodology were not classified as such by the other methodology.

The analysis in Appendix A does not attempt to determine which methodology more accurately reflects true quality or true differences in quality. However, the fact that the two methodologies produce different results when rating and ranking the same physicians implies that at least one of the methodologies is lacking and suggests that further thought and testing is necessary. If MIPS proceeds with the current methodology and the agencies adopt it, the inconsistencies demonstrated here could result in physician frustration and dissatisfaction and could ultimately lead to a lack of confidence in the MIPS and inter-agency programs. Further, these inconsistencies could send mixed signals to patients who might make incorrect assumptions about physician quality when deciding where to seek care.

As highlighted in Appendix A, we have several concerns regarding the current MIPS benchmark methodology and offer illustrative examples using the 2015 Individual Physician Compare data downloaded from the CMS website. Our main concerns with the MIPS benchmark methodology are:

1. For topped-out or highly-skewed data, thresholds are clustered close together (meaning that similar performance may not result in similar points awarded) and even relatively high performance can place a physician in one of the lower deciles. For example, a physician could score 88 percent and be in the 4th decile while another physician scores 92 percent and is in the 8th percentile. Therefore, on the same measure two physicians can perform very similarly on the measure but be awarded very different points;
2. There is a lack of consideration of the role played by random fluctuation, especially for small denominators;
3. Strictly data-driven thresholds may conflict with clinical knowledge and evidence of ideal performance or with practical considerations of quality;
4. There may be significant changes to the population of physicians and groups between the time that the historical data represents (2 years prior) to the time period to which the resulting thresholds are applied; and
5. Under certain circumstances, physician performance score under MIPS may differ significantly from their performance under the Physician Compare methodology, even for the same measure.

Therefore, prior to considering use of Quality Payment Program data for in-network providers, we urge federal agencies to work with CMS to revise the benchmark methodologies to allow measure thresholds to incorporate clinical knowledge, consider the impact of random fluctuation, and be adjusted for practical considerations of comparison and relative performance. To address the shortcomings of the existing benchmark methodologies, we suggest that CMS implement a methodology that allows for manual manipulation of thresholds. As we highlight in Appendix A, MIPS Benchmark White Paper, this would allow for enough flexibility to address the above issues when they arise. We acknowledge that this would add additional work to an already complex process, but we believe that what is most important is ensuring the fairness and clinical relevance of the measure benchmarks and making more accurate assessments about quality. We further acknowledge that there may be modifications to the methodology other than

what we suggest, which may also address our concerns and welcome the opportunity to discuss further with the Agencies and CMS.

5. Could quality of health care information from state-mandated quality reporting initiatives or quality reporting initiatives by nationally recognized accrediting entities, such as NCQA, URA, JC and NQF be used to help participants, beneficiaries, and enrollees meaningfully assess health care provider options?

While the AMA supports increased transparency in health care, we have concerns about moving forward with publicly posting accrediting organization (AO) information currently. The current accreditation surveys provide an overwhelming amount of information for consumers to digest and interpret. Much of the information contained in these reports raises issues that could alarm prospective patients yet are easily remedied during the accreditation process or during the resolution of a complaint. In addition, survey reports contain references to current medical liability cases along with personnel health and training information, which has previously been considered confidential information. While some may believe posting AO survey reports and plans of correction (PoCs) would provide a more comprehensive picture of the health care system, the AMA respectfully disagrees and believes that without proper aggregation and context, data from AO surveys and PoCs could be easily misinterpreted. Furthermore, it is unclear whether a facility's full corrective action plan would have to be posted and there are varying degrees of severity within a corrective action plan that can easily be misinterpreted.

We also believe that the statutory intent is clear that no AO surveys should be published unless an enforcement action was taken by the Secretary, which is contrary to consideration of publicly posting accrediting organization information. Section 1865(b) of the Social Security Act states:

“The Secretary may not disclose any accreditation survey (other than a survey with respect to a home health agency) made and released to the Secretary by the American Osteopathic Association or any other national accreditation body, of an entity accredited by such body, except that the Secretary may disclose such a survey and information related to such a survey to the extent that such survey and information relate to an enforcement action taken by the Secretary.”

While CMS has in the past proposed to require AO organizations themselves to post the surveys instead of CMS, the AMA believes that this would clearly still violate the intention of the statute that this information not be reported until an enforcement action was taken. In addition, the status of whether a facility is accredited is already published and includes material defects of those facilities in pending status.

Therefore, instead of moving forward with the publication of additional information, HHS should instead focus on improving the accuracy and usability of the data that is currently available to consumers through public reporting websites such as Hospital Compare, Nursing Compare, and Physician Compare. As the AMA has noted previously, physicians continue to find inaccuracies in the data published on Physician Compare. In addition, the data review and correction process continues to operate on an unrealistic timeline and remains burdensome for physicians and their staff.

Accrediting organizations, such as The Joint Commission or URAC either accredit hospitals or health plans. Individual physicians are not accredited but their practice may earn some form of certification by participating in a certification program, such as ones administered by NCQA.

Furthermore, to clarify, the National Quality Forum (NQF) is not an accrediting organization, but a voluntary consensus standards body as defined by the Office of Management and Budget (OMB) Circular A-119 that reviews and endorses measures for use in health care quality and public reporting programs. Certification or accrediting programs may submit their measures to NQF for review and the organizations may include the endorsed measures in their programs.

6. What gaps are there in current measures and reporting as it relates to health care services and items in the individual and group markets?

There is currently a gap in outcomes measures as they relate to health care services and the AMA would encourage the Agencies to look to specialty society clinical data registries to fill the gap. There are many benefits to using clinical data for public reporting:

- Clinical data more directly reflect the delivery of care.
- Unlike administrative data, which aggregates experience for system management requirements, clinical data are patient specific and can be more precisely specified to define best practice.

Therefore, we would not support or encourage the use of administrative claims data or administrative claims-based measures for public reporting due to the following reasons:

- Hospital coding compiles specific diagnosis codes into diagnosis-related groups that define reimbursement rather than reflect an accurate sequence of clinical events for clinical care delivery.
- Attribution of causality or association with adverse events to specific providers is not possible with aggregated data.
- Inconsistency in definitions of terms and coding makes the data less reliable for quality reporting. The coder must often interpret what the physician has documented, which adds the unaccountable subjective bias of both.
- There are no standardized methods for attribution of providers' roles in patients' episodes of care.

Measure developers (and CMS) previously moved away from administrative claims-based measures in part due to concerns over attribution and the inability to move to clinically meaningful outcomes measures using these data. It is a reason for the big push to adopt electronic clinical quality measures (eCQMs) and registries, because electronic tools provide for a much richer data source. For example, it is very difficult to get to intermediate outcomes such as diabetes HbA1c levels or blood pressure levels measures without requiring additional data collection. Claims also do not provide granular enough information for staging a disease. Therefore, the Agencies and CMS will be left selecting measures that may be great to look at from the community or population perspective but not appropriate to attribute to an individual physician or group practice and are so far removed from clinical practice that they do not provide meaningful and actionable data at the point of care or assist patients in deciphering care.

7. The Departments are also interested in understanding any limitations plans and issues might have in reporting in-network provider quality in the individual and group markets.

Public reporting at the individual physician level is extremely difficult to achieve due to small sample sizes. It is even more difficult at the individual surgeon level given the small number of some surgical operations performed and may result in unintended consequences, such as causing surgeons to become

risk averse or patients making misinformed assessments about quality.^{20,21} If metrics are deemed necessary and required to be publicly reported they should focus on quality of life and patient satisfaction, not mortality. Therefore, we encourage the Departments to work with and consult with physician-led organizations such as the PCPI and physician specialty societies to better understand the limitations with reporting quality information. We also would encourage the Departments to work with the AHRQ, who has expertise in public reporting and produced numerous reports on best practices.

8. The Departments seek more information about how and if quality data is currently used within plan and issuer provider directories and cost-estimator tools. The Departments also seek information on the data sources for quality information and whether plans and issuers are using internal claims data or publicly available data.

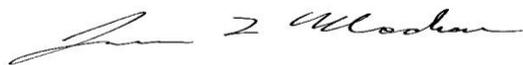
Based on the AMA's experience interacting with health plans, plans vary in their approach. Some use strictly claims data, their own pay-for-performance programs, publicly available data or a hybrid approach. There is no consistency between plans on how they design and display quality information within provider directories or cost-estimator tools.

Conclusion

As you can see, the AMA strongly supports efforts to increase the availability of health care pricing information and provide patients with appropriate tools to assist them in health care decision making and managing their health care costs. Increased availability of information estimating patient out-of-pocket expenditures, when coupled with information about the quality of the health care provider or facility in question, empowers patients to make value-based decisions about their care. We support Administration efforts to ensure health plans provide patients with tools to help estimate their out-of-pocket expenditures but urge the Administration to limit that information to only that which is truly actionable for patients. We further urge the Administration to ensure appropriate educational information is provided to patients to help them better understand and navigate the information being displayed.

While the AMA strongly supports increased price transparency for health care services, we are concerned about requirements to publicly disclose payment rates privately negotiated between health plans and providers. Public disclosure of these privately negotiated rates may hinder the market's ability to maintain robust competition, may have negative unintended consequences on health care prices, and may ultimately serve to limit patient access to care. **We strongly urge the Administration to focus transparency efforts on the information that will ultimately have the greatest impact on health care decision-making—actionable information about patient out-of-pocket costs.** The AMA looks forward to continuing to work with the Administration on the difficult issue of health care costs. Please contact Shannon Curtis, AMA Assistant Director, Federal Affairs (shannon.curtis@ama-assn.org) for any questions or to discuss further.

Sincerely,



James L. Madara, MD

²⁰ Burns, Elaine M., Pettengell, Chris., Athanasiou, Thaos., Darzi, Ara., Understanding The Strengths And Weaknesses Of Public Reporting Of Surgeon-Specific Outcome Data. *Health Affairs*, 35, no.3 (2016):415-421

²¹ Surgery & Public Reporting: Recommendations for Issuing Public Reports on Surgical Care. Surgical Quality Alliance. https://www.aans.org/pdf/Legislative/2014SQA_PublicReportingDocument.pdf. Accessed, Jan. 08, 2020.

Appendix A

Comparing Physician Compare and MIPS Methodologies and Potential Areas of Exploration and Alternatives

Executive Summary

Within the current environment of healthcare quality measurement and assessment, there are multiple programs that attempt to rank and compare the quality of care physicians provide. The Center for Medicare and Medicaid Services (CMS) Merit-based Incentive Payment System (MIPS) within the Quality Payment Program (QPP) involves awarding points to physicians based on where they fall in decile-based categories calculated from historical quality measure data (when available). Notably, this methodology differs from that of other programs, such as CMS' Physician Compare star rating public reporting program. Physician compare uses the Achievable Benchmarks of Care (ABC) methodology to place physicians into one of five categories (each with a corresponding "star rating") for purposes of helping patients compare physicians to make more informed decisions about where they seek care. In contrast, the MIPS methodology uses nine categories (and point system) to score physicians on quality measure reporting to determine whether a physician will be subject to a MIPS penalty or eligible for an incentive.

Given that physicians are evaluated under multiple programs, sometimes for the same or similar measures, the MIPS methodology raises concerns regarding the validity of the methodology under certain scenarios and the consistency of physician performance results across the multiple quality programs and their unique benchmarking methodologies. This analysis sought to explore these methodologies to identify strengths and weaknesses to better understand the implications for the physicians being measured.

Using 2015 individual Physician Compare data, we apply both the ABC (using "equal ranges") and MIPS decile methodologies to the same measures and compare the cut-offs defined by each method to see if there are differences in physician performance by methodology.

Results of our analysis suggest that when the distribution of physician performance is relatively uniform across the entire range from 0% to 100%, the two methodologies produce similar relative performance comparisons for physicians. However, when data is skewed or u-shaped, there can be significant differences in how physicians are rated by each methodology. Additionally, each methodology can be influenced by extreme values and often result in physicians needing to perform at 100% to achieve the highest performance rating.

We also propose an alternative methodology that involves manual cut-offs for the highest and lowest categories are created manually and the rest are created based on the distribution of performance. This methodology, referred to as "Manual plus Data-Driven" (M+DD), is less susceptible to extreme values and also allows administrators to incorporate clinical evidence and practical considerations into the measurement benchmarks, ensuring that categories are relevant and truly reflect high and low levels of quality.

Introduction

Currently, MIPS benchmarks utilizing historical data are created by calculating deciles of these data and awarding points to physicians based on where they fall in these deciles. Given that physicians are evaluated under multiple programs (MIPS and Physician Compare), there are concerns regarding (1) the validity of this method for given situations, and (2) the consistency of physician comparisons across different programs and benchmarking methodologies (e.g., that for the Physician Compare 5-star rating). Therefore, we sought to explore these methodologies and viable alternatives to better understand the implications for physicians.

Objective

In the analysis below, we examine the methodologies for MIPS benchmarks (deciles) and the 5-star rating (ABC with equal ranges) and compare the results they produce when applied to the same data. This allows us to compare how physicians would perform under each methodology to assess the level of consistency across measurement methodologies. For this analysis, we use the 2015 Physician Compare downloadable database, which is the most recent data available for download. To evaluate these methods, we attempt to determine how physician performance would differ under each method, and we also explore how each method performs within different shapes of distributions of physician performance. Finally, we explore alternative methods that may be more robust to the variety of performance distributions seen across the quality measures.

Methodology

Using 2015 individual Physician Compare data, we manually apply both the ABC (equal ranges) methodology (top graph of Figure 1) and the decile methodology (bottom graph of Figure 1) to the SAME measure to identify cut-offs defined by each method.

After removing zeros²², we applied the ABC and deciles methodologies as follows:

ABC:

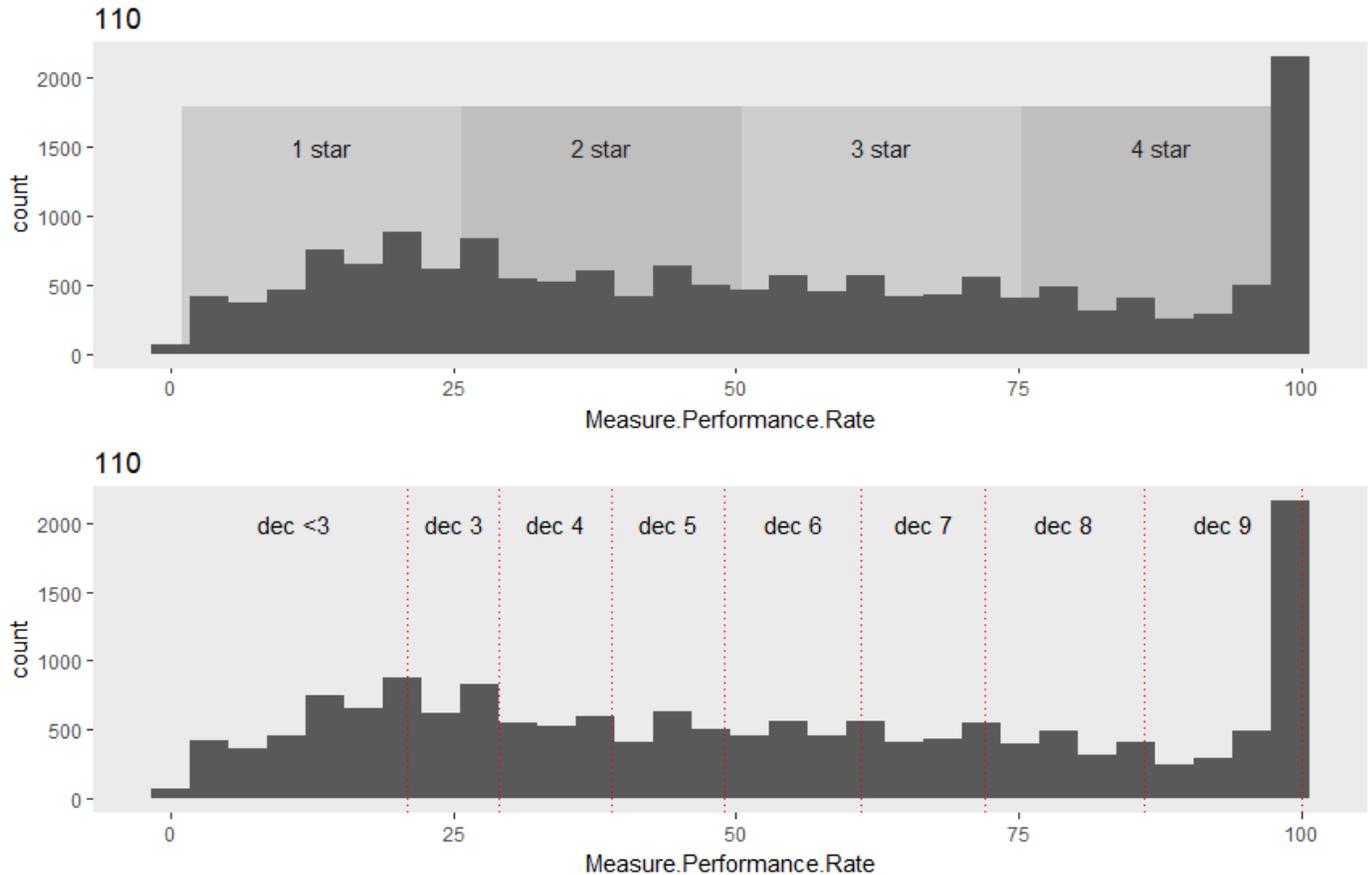
- Sort physicians in order of performance
- Identify the top physicians who account for 10% of the patient population
- Calculate an aggregate rate for those physicians and set this as the top “benchmark”
- Identify the lowest performance and set this as the bottom threshold
- Set the middle cut-offs using the “equal ranges” methodology between the top and bottom thresholds

²² Under MIPS submitting zeros is not considered satisfactory reporting and not included in the data that forms the deciles, but is included in the Physician Compare downloadable database.

Deciles:

Simply calculate the relevant deciles and set cut-offs at those deciles

Figure 1.



We recognize that these steps do not fully encompass the breadth of how CMS applies these methodologies. This is in part because the Physician Compare data includes data from all physicians who submitted data under the legacy Physician Quality Reporting System (PQRS) program, not just those who are MIPS-eligible, and creates separate ratings for each measure based on whether it was reported as an individual or group. It is also due to CMS performing additional calculations, the details of which are not publicly available and therefore we are unable to incorporate (e.g., the beta-binomial adjustment factor in the ABC method). Therefore, we understand that the cut-offs we calculate here *will likely not match those produced by CMS*. However, given that our goal is to compare characteristics of the methodologies, these data are sufficient. We also understand that the Physician Compare technical expert panel (TEP) has explored the impact of outliers and understands how they can impact ABC cut-offs. Therefore, the influence of outliers on ABC cut-offs described in our examples below may not reflect actual practice, but instead illustrate general differences between the ABC method and the decile method and help to explore the implications of these differences. In short, we recognize that our analyses are limited by our inability to replicate the full methods employed by CMS, but our goal is to understand the differences in the

methods and identify potential pitfalls they may encounter (even if these pitfalls are addressed by the full methods CMS employs).

These two methods obviously produce different numbers of categories. The ABC methodology produces 5 categories, while the decile methodology produces 9 (for MIPS, deciles 1 and 2 are grouped into a single category: “< 3”). To be able to compare these methodologies, we need to decide what would be an “equivalent” category for each methodology. That is, what decile(s) *should* a 1-star rating equate to if the methodologies were reasonable representations of one another? Below is what we have assumed for this comparison:

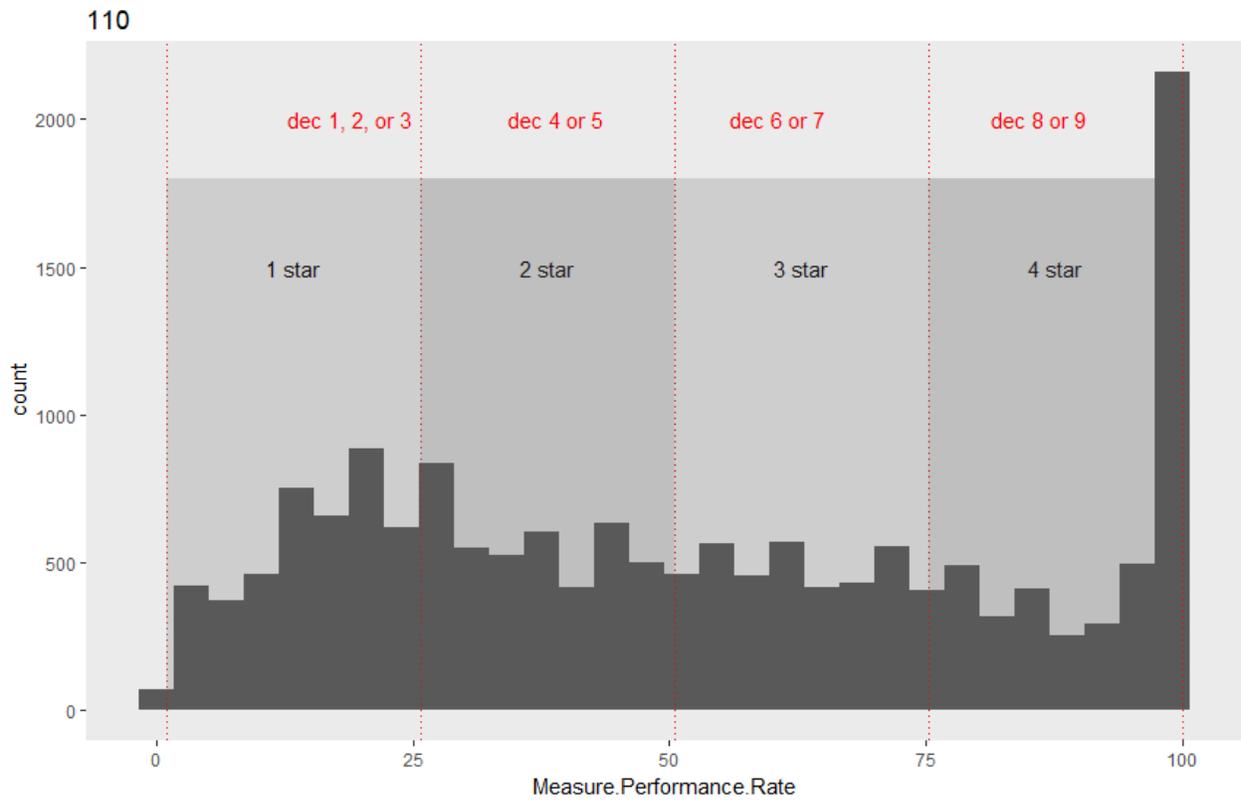
Table 1.

5-star rating	Decile(s) we consider to be comparable
5 stars	Decile 10
4 stars	Deciles 8 or 9
3 stars	Deciles 6 or 7
2 stars	Deciles 4 or 5
1 star	Deciles 3 or lower (i.e., 1, 2 or 3)

There are other groupings which may also be reasonable, but these groupings will allow us to compare physician performance between the methods and it’s likely that alternative groupings would not significantly change our findings below.

Under the assumption that the categories in Table 1 represent “equivalent” levels of quality, then the two methodologies are considered to produce identical results if the dotted red lines established by deciles line up with the edges of the shaded regions produced by the ABC methodology. That is, everyone in decile 1-3 would also be 1-star, everyone in decile 4 or 5 would be 2-star, etc, and it would look like that presented in Figure 2 below.

Figure 2.

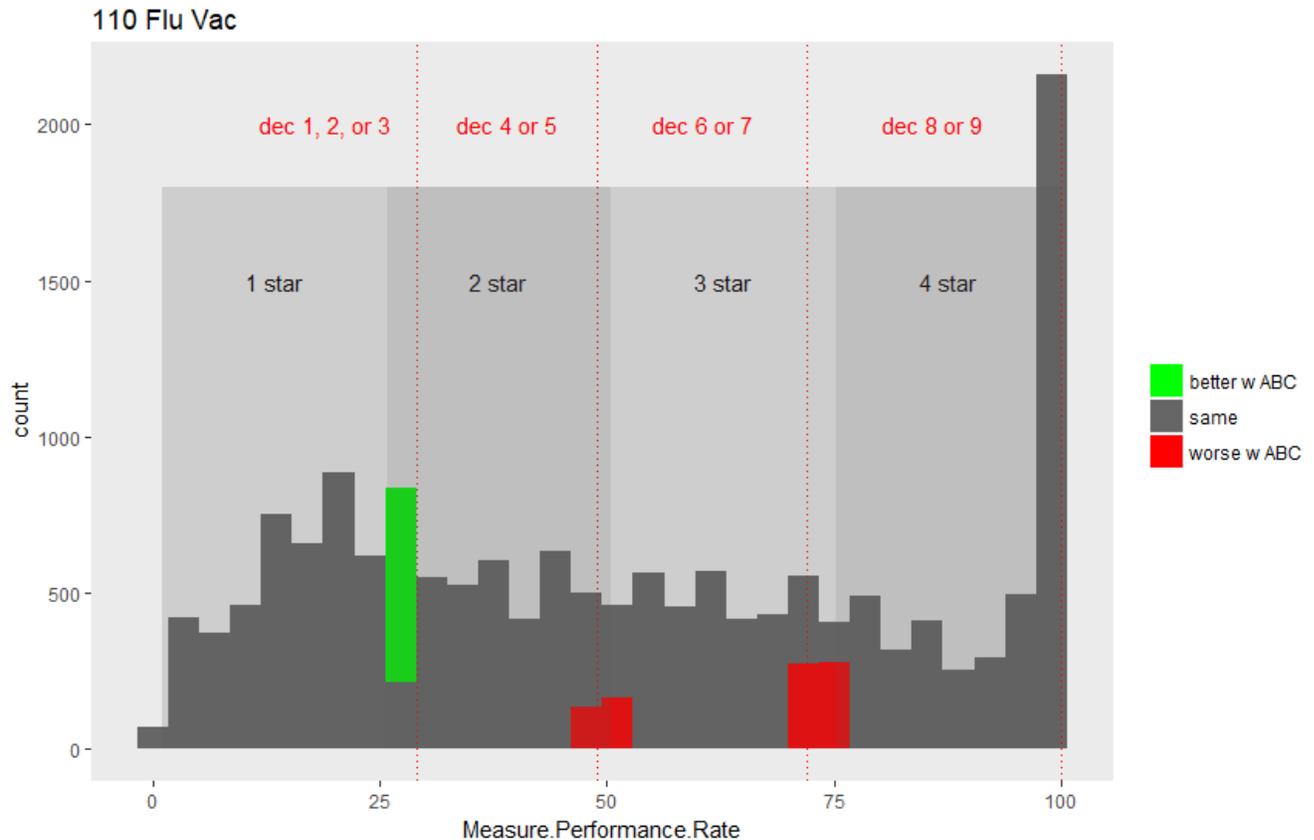


(This is an example, NOT the truth, but shows what would happen if the deciles and 5-star ratings lined up.)

Results

Example 1: A relatively uniform distribution

Figure 3.

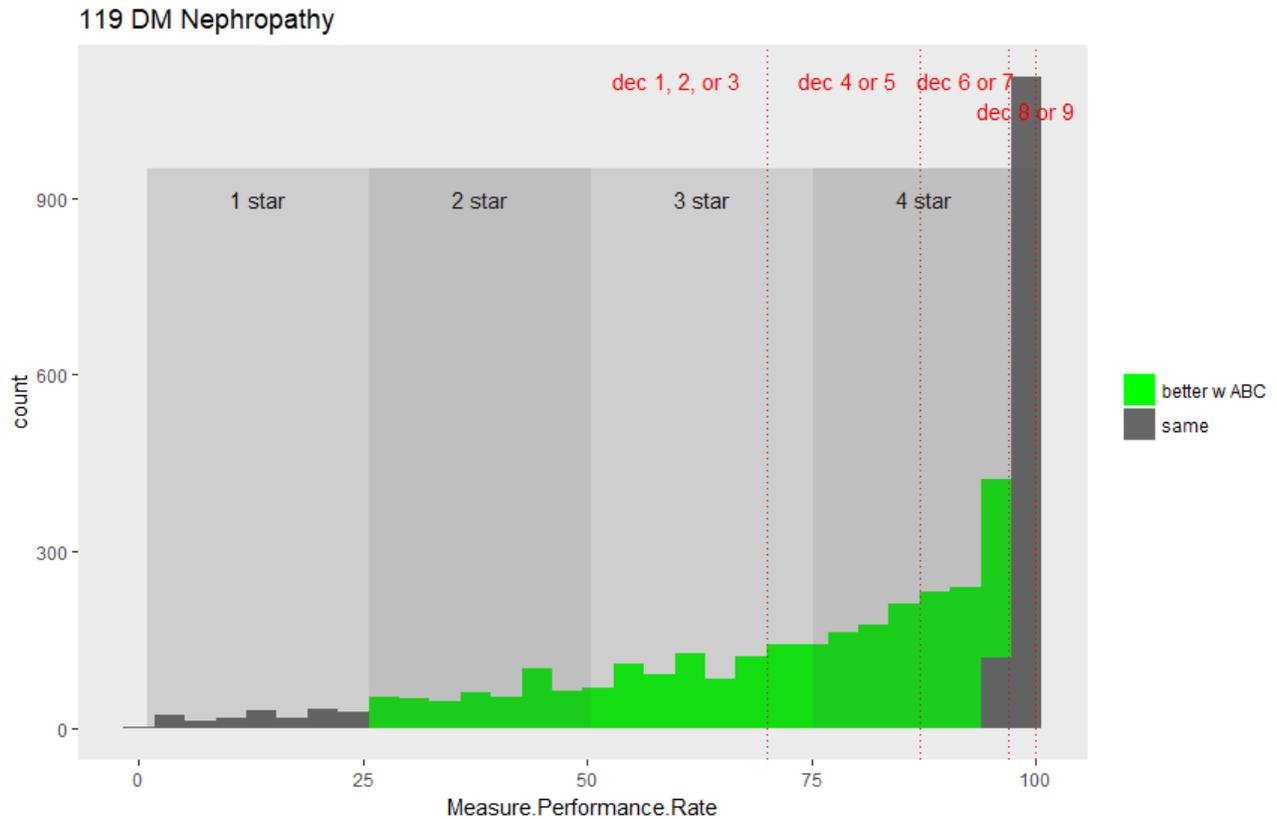


With the exception of the large number of physicians at 100%, this distribution is relatively evenly spread across the entire range from 0 to 100. And, we can see that the two methodologies produce similar relative performances for all but a handful of physicians, some of whom would fare better under the ABC methodology, and some of whom would fare better with the deciles methodology.

Note, that when utilizing the “equal ranges” method for ABC, the cut-offs are established *entirely* by the top 10% and the single lowest score. In the above graph, if the low score was closer to 25% the shaded areas would all shift to the right, so that those currently near the bottom of a particular star-rating category could drop to a lower category. What that means is that if the very worst performers (and *only* the very worst performers) improved, physicians in the middle of the distribution who maintain performance could actually have their star rating fall – again, not because they performed worse relative to those around them, but because the very worst performers performed a little less poorly.

Example 2: A skewed (or topped-out) measure

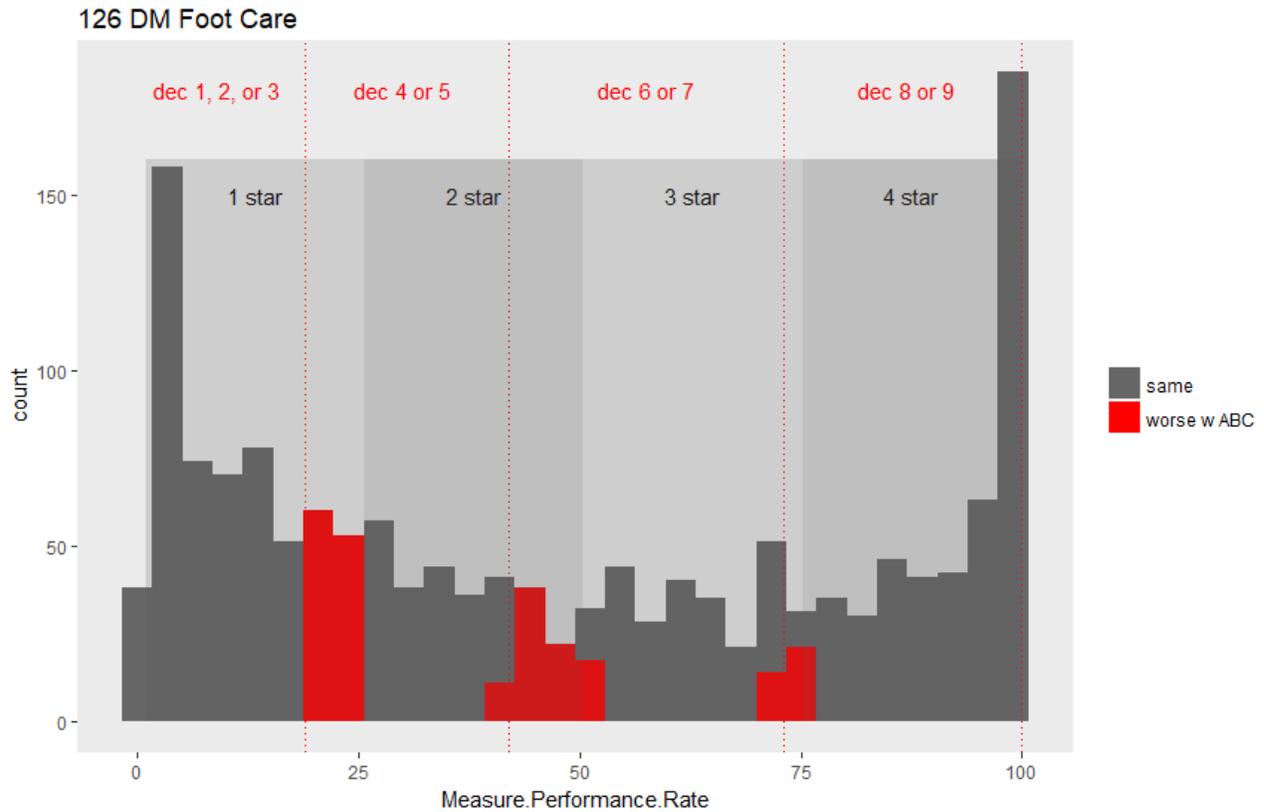
Figure 4.



When measure performance is skewed (even if it is not topped out) *but still has physicians who perform very poorly*, the two methodologies will produce very different results. In the above example, the significant number of physicians at 100% has pulled the decile cut-offs up, so that even relatively adequate performance lands a physician in a mid- to low-level decile. In contrast, since there is at least one physician who performs poorly, the ABC cut-offs are equally spaced between the top benchmark (100%) and the very bottom (1%). That is, the ABC cut-offs for this measure are identical to those in the previous example, even though the distributions look very different. That low performer pulls down all of the ABC cut-offs, and as a result, those performing at levels between 25% and about 95% fare better under the ABC methodology than the decile methodology. One could argue that the decile method sets the bar too high: physicians to perform near (but not at) 100% will fall into a mid- to low-level decile. Similarly, one might be troubled by the fact that the ABC methodology produced the exact same cut-offs as in the previous example, even though the performance distributions are very different.

Example 3. A U-shaped distribution

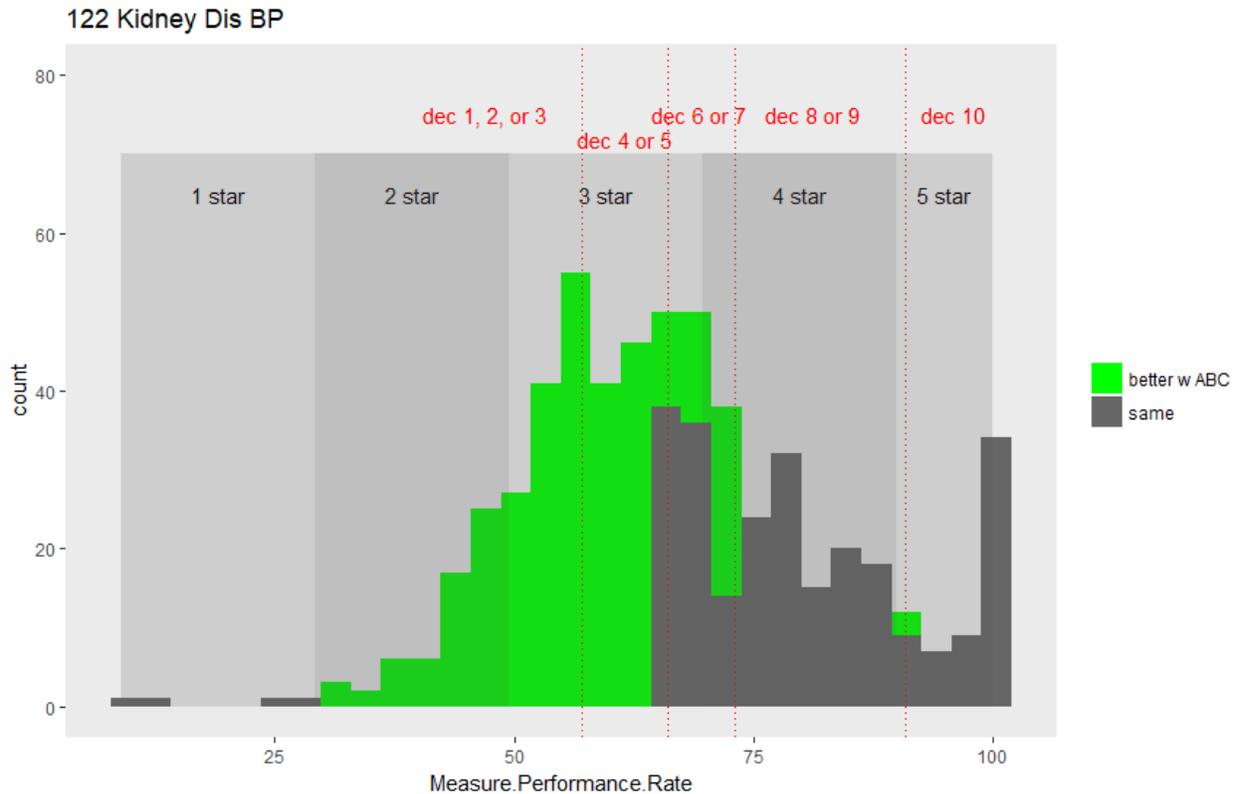
Figure 5.



In this graph, the ABC cut-offs are again the same as in the previous two graphs because there are at least 10% who perform at 100% and at least one physician who performs at 1%. Because there are a large number of physicians performing poorly, the decile-based cut-offs are pulled down, so that physicians would perform better under the decile method. This example illustrates again that the equal-range ABC method ignores the distribution of physicians between the very top and very bottom performers, and can produce identical cut-offs for very different distributions. At the same time, even with a wide range of performance results, the decile method requires that physicians perform at 100% to receive the highest rating.

Example 4. Normally distributed

Figure 6.



When roughly Normally distributed, the two methodologies produce similar cut-offs, although in the above example the significant number performing at 100% pulls up the decile cut-offs, which (in conjunction with at least one poorly performing physician to keep the ABC-based cut-offs spread across the entire range) results in a large number of physicians who would perform better under ABC. In this instance CMS should consider and explore the possibility that the very poorest performers are outliers that should be removed. If so, that would significantly shift the cut-offs for 1-star through 4-star ratings.

Thoughts and Reactions

Clearly, there are a variety of examples where these two methods would potentially produce significantly different results for physicians. This is a potential problem for physicians who may find themselves rated significantly differently for the same quality topic area by different quality reporting programs. It may also pose a problem for consumers (e.g., patients) who could face conflicting information regarding the quality of individual physicians depending on which rating they use to evaluate their options. For each methodology, there is the opportunity for extreme values to influence cut-offs. We realize that CMS can and may take steps to mitigate this issue through adjustment for small sample sizes and even removal of extreme values. However, some extreme values are not outliers: in situations where the highest cut-off is set at 100%, there could be unintended consequences. Physicians performing at 100% may be incentivized to expend energy and resources to stay at 100% rather than applying those resources in other areas where they may produce a larger marginal benefit. Said another way: if the financial losses of

dropping from 100% to 98% in Measure A are larger than the financial gains of improving from 60% to 70% in Measure B, physicians are incentivized to spend their time on Measure A, even if the improvement in Measure B would be more significant for patient care. For the ABC methodology, the examples above illustrate that it is possible that significantly different distributions can produce identical cut-offs, simply because the “equal ranges” methodology relies solely on the top 10% and lowest performers. A case could be made that data-driven cut-offs should utilize more of the data than just those from the highest and lowest performers. Additionally, as stated previously, under the ABC equal-ranges methodology, middle-of-the-pack physicians who maintain quality can experience a reduction in their star rating if the very worst physicians improve (since it’s these physicians who largely determine where cut-offs fall).

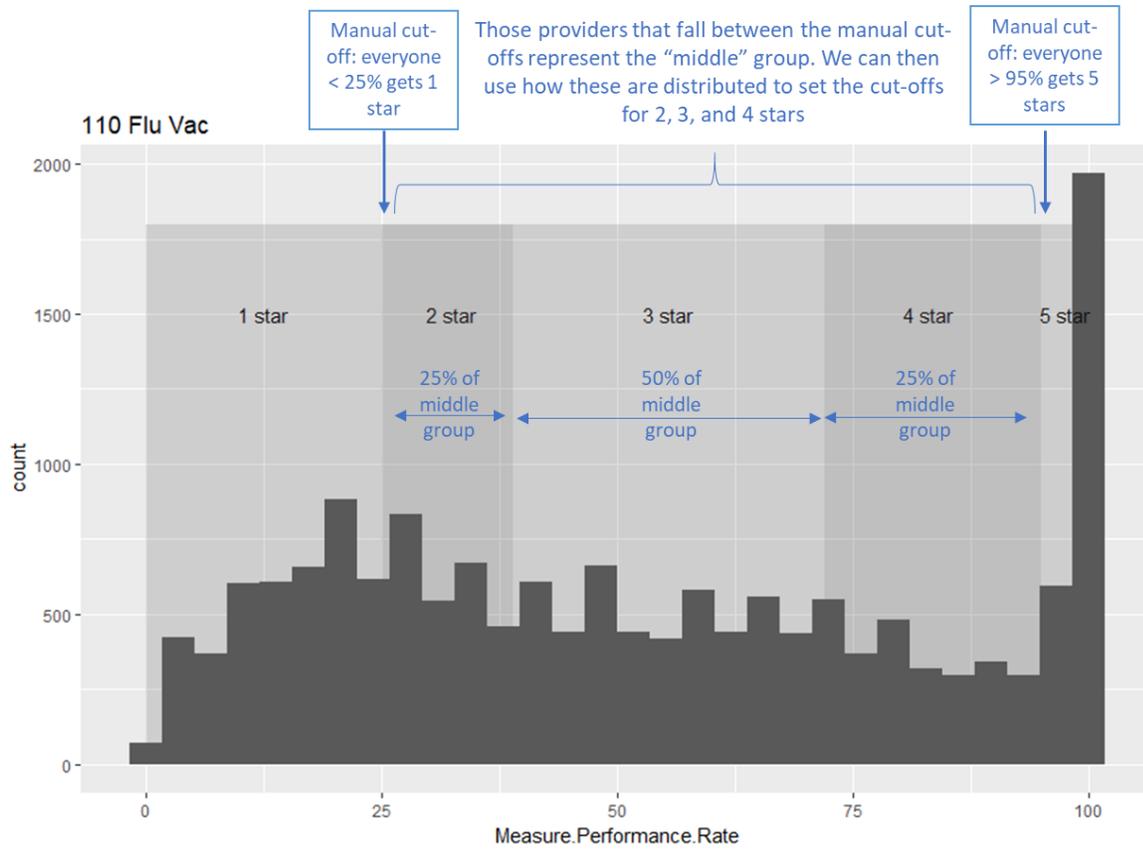
Alternative Methodologies

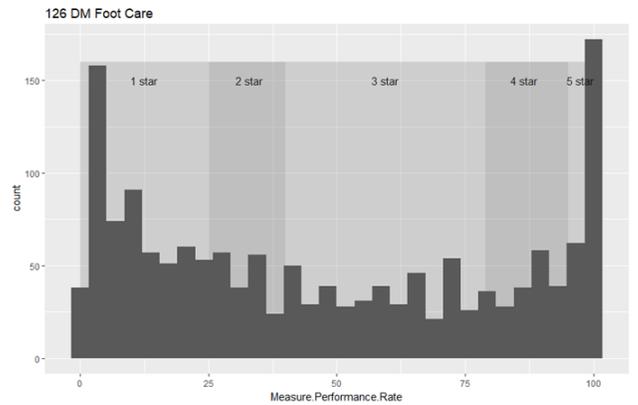
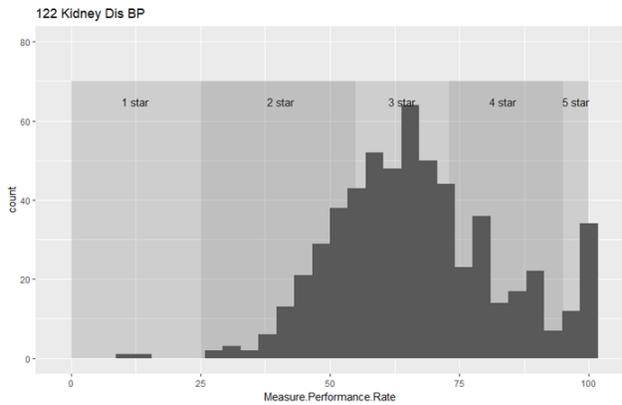
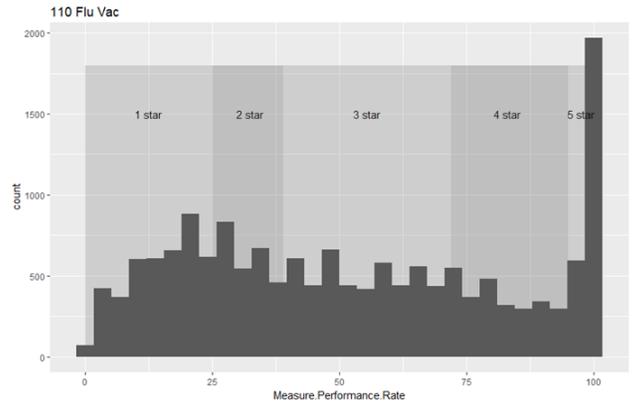
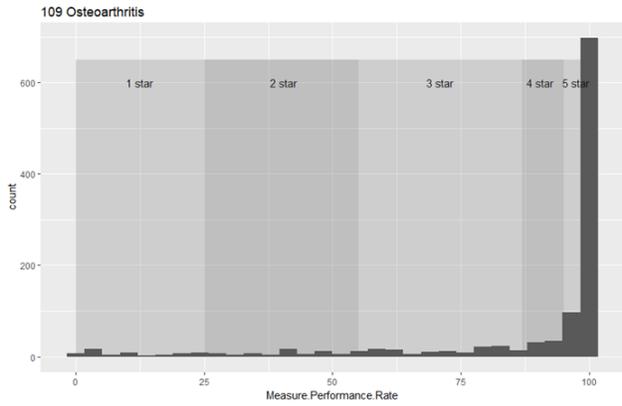
In exploring alternative methodologies, we sought to identify methods that: (1) were relatively robust to extreme values, and (2) allowed one to incorporate clinical/empirical evidence and/or practical considerations when setting cut-offs. One option would be to manually set cut-offs for the highest and lowest categories, and then allow the distribution of performance between those cut-offs to determine the other cut-offs. One example is as follows:

- Everyone scoring 95% or higher gets the highest rating (e.g., 5 stars)
- Everyone scoring 25% or lower gets the lowest rating (e.g., 1 star)
- Those scoring 26% to 94% are used to determine cut-offs for the ratings between highest and lowest (e.g., find the first and third quartile of THESE physicians and set cut-offs there)

The figure below illustrates how this would be done, and the series of graphs that follow demonstrate that while the middle cut-offs are impacted by the different distributions, the manual cut-offs ensure that high performance is always rewarded and poor performance is always penalized, regardless of the distribution.

Figure 7.

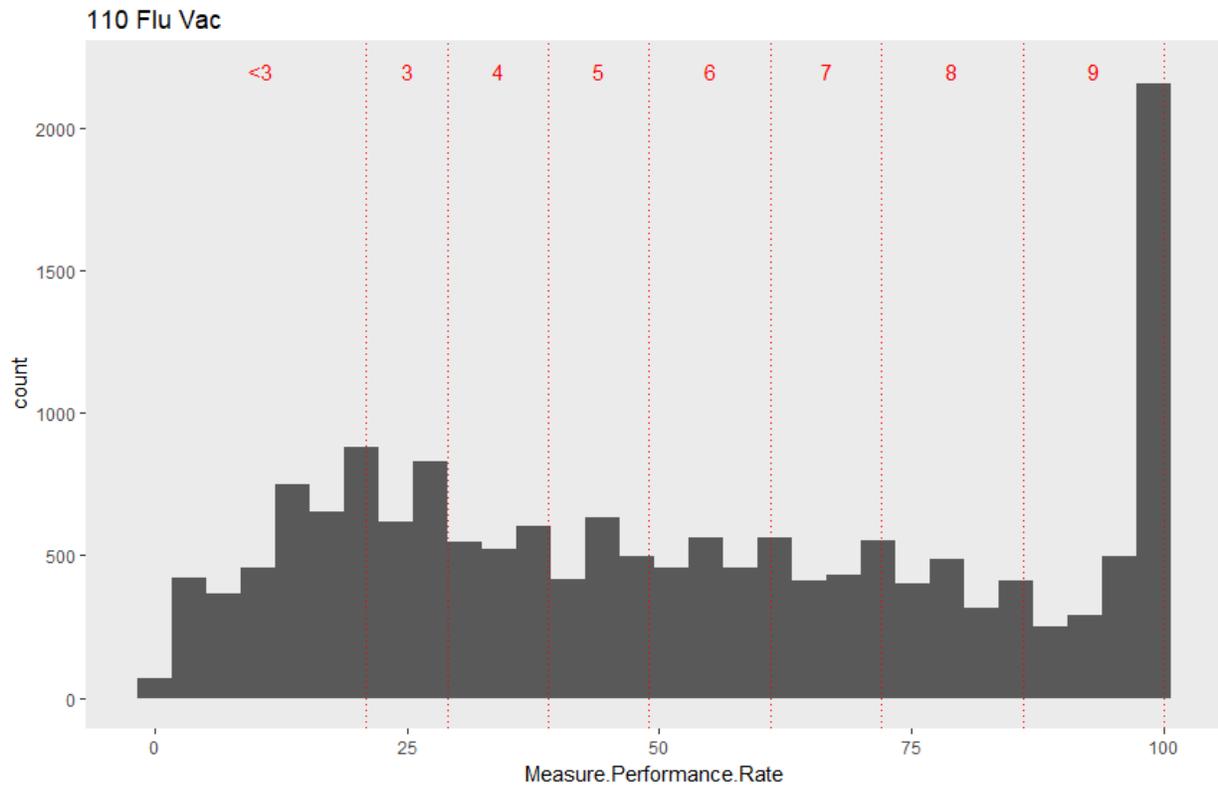




While the above example uses 5 total categories, the same thing could be done using any number of categories. In fact, if used to create 10 categories this methodology could be seen as simply a slightly-modified version of the decile method, where the top and bottom categories are manually set while the other cut-offs are data-driven in the same way the current decile method cut-offs are created.

When applying the decile method to the Flu Vac measure, we get the following cut-offs (decile 10 is 100%):

Figure 8a. Decile-based cut-offs for the Flu Vac Measure



If we apply the proposed Manual plus Data-Driven (M+DD) method, again using cut-offs of 25% and 95% for the lowest and highest cut-offs, respectively, we get Figure 8b. Because overlaying these graphs would be visually challenging (given the large number of categories), Figure 8c compares simply the cut-offs to demonstrate how the M+DD modifies the decile-based cut-offs.

Figure 8b. M+DD-based cut-offs for the Flu Vac Measure

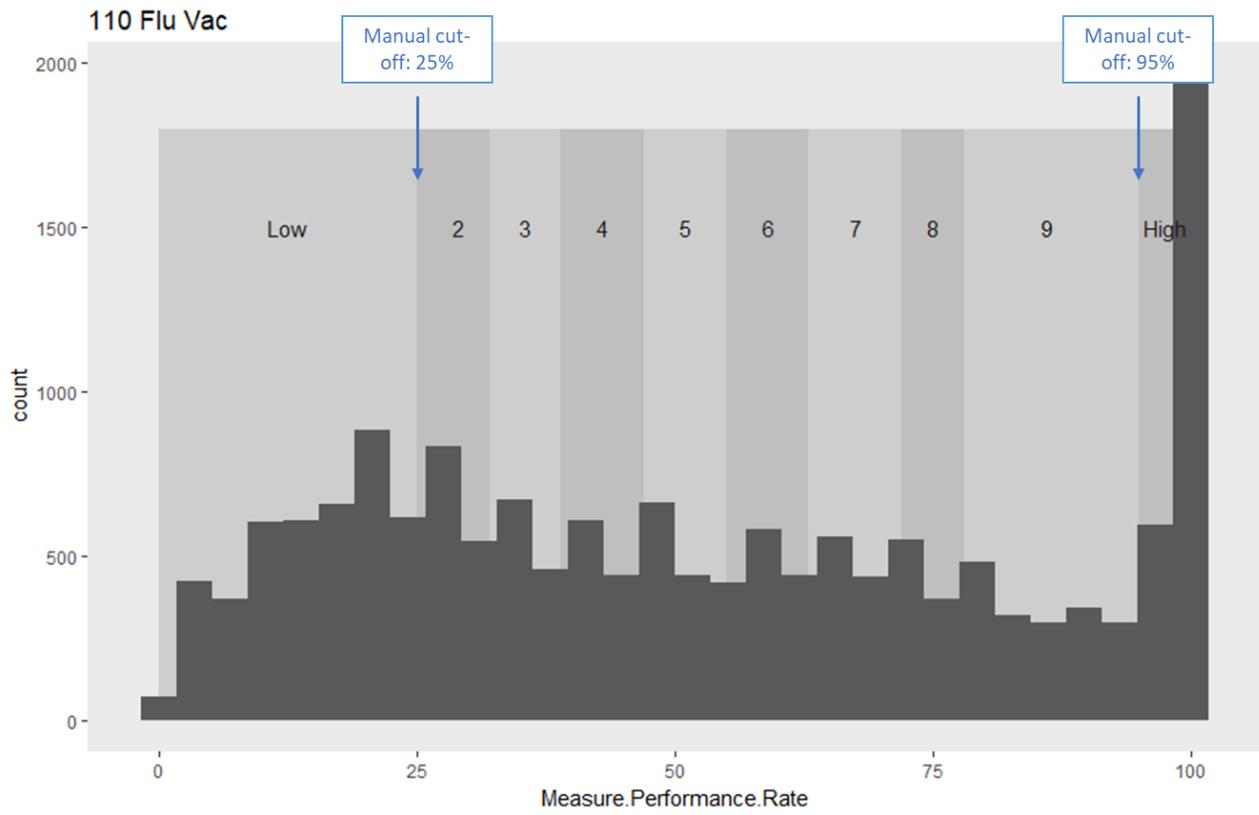
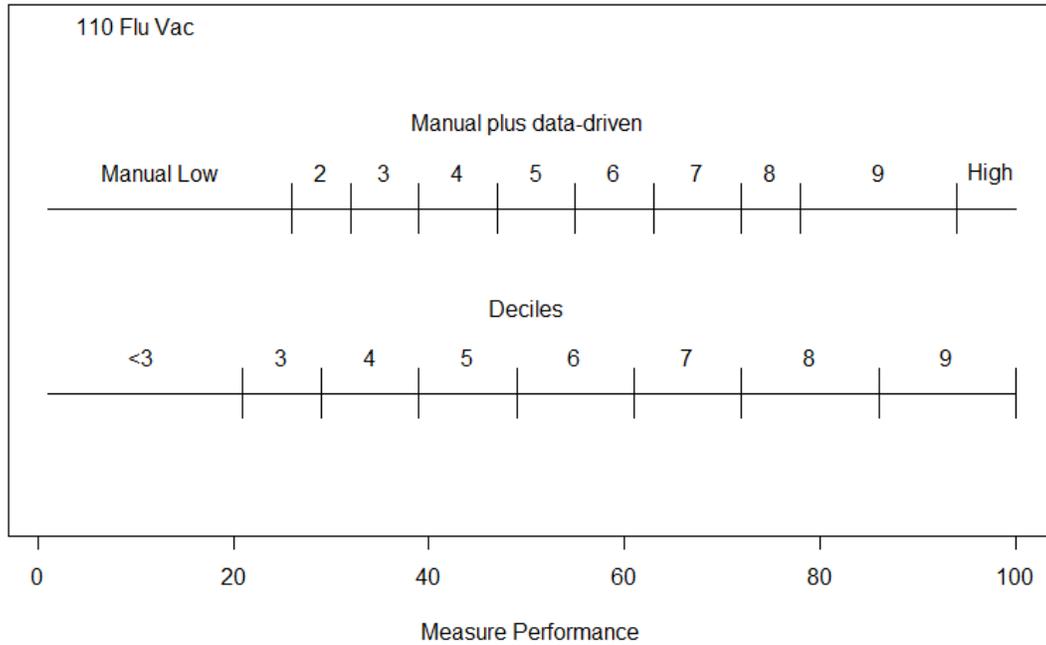


Figure 8c. Comparison of cut-offs for the Flu Vac measure



In general, the cut-offs are relatively similar (Fig 8c) because of the relatively well-behaved distribution. The main differences revolve around the extreme values: the manual cut-offs of the M+DD method result in a lower threshold for the top category and a higher threshold for the lowest category.

The example illustrated in Figures 9a-c demonstrates how the M+DD can mitigate the skewness in cut-offs produced by topped out measures, while still producing cut-offs that require high performance to achieve the highest categories.

Figure 9a. Decile-based cut-offs for the Osteoarthritis measure

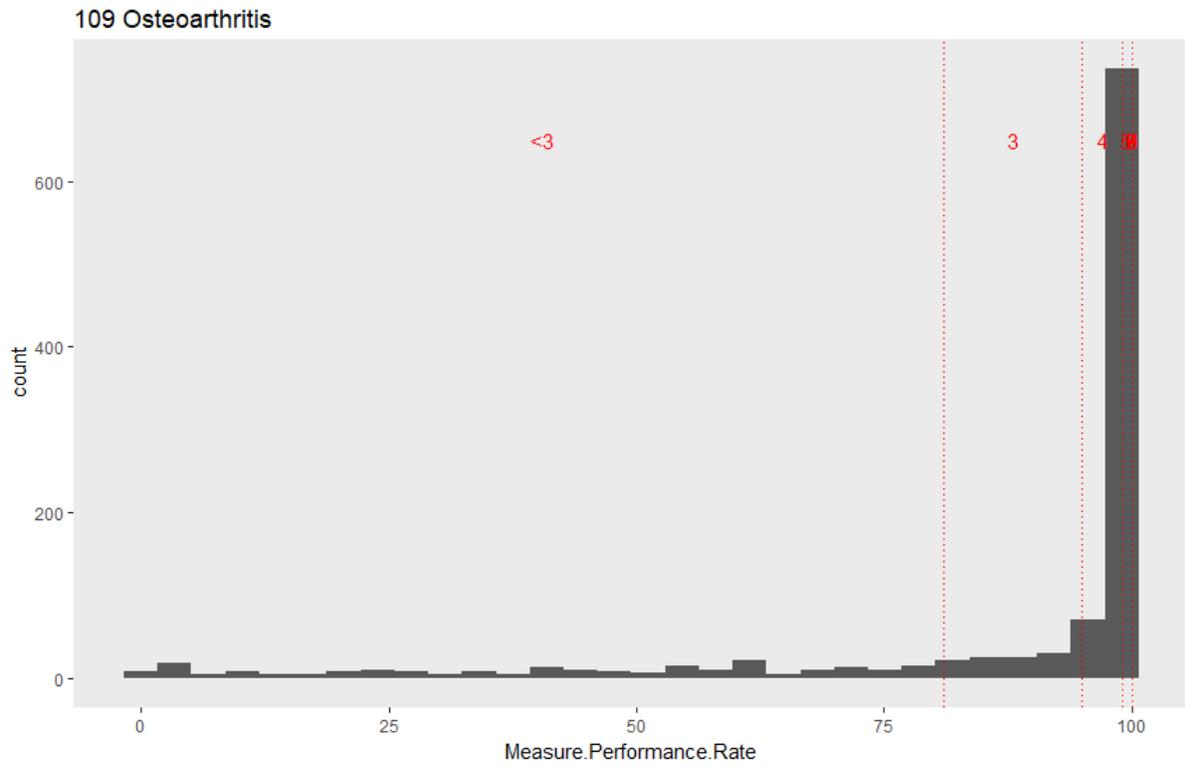


Figure 9b. M+DD-based cut-offs for the Osteoarthritis measure

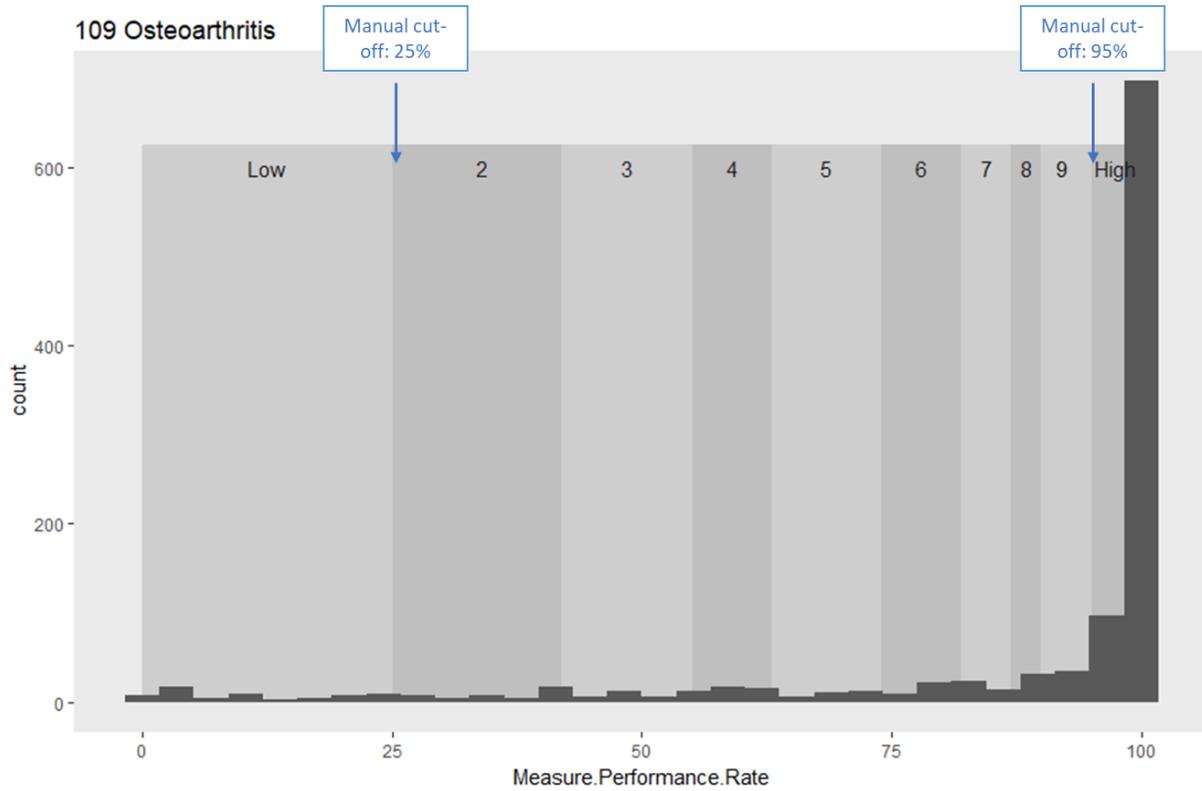
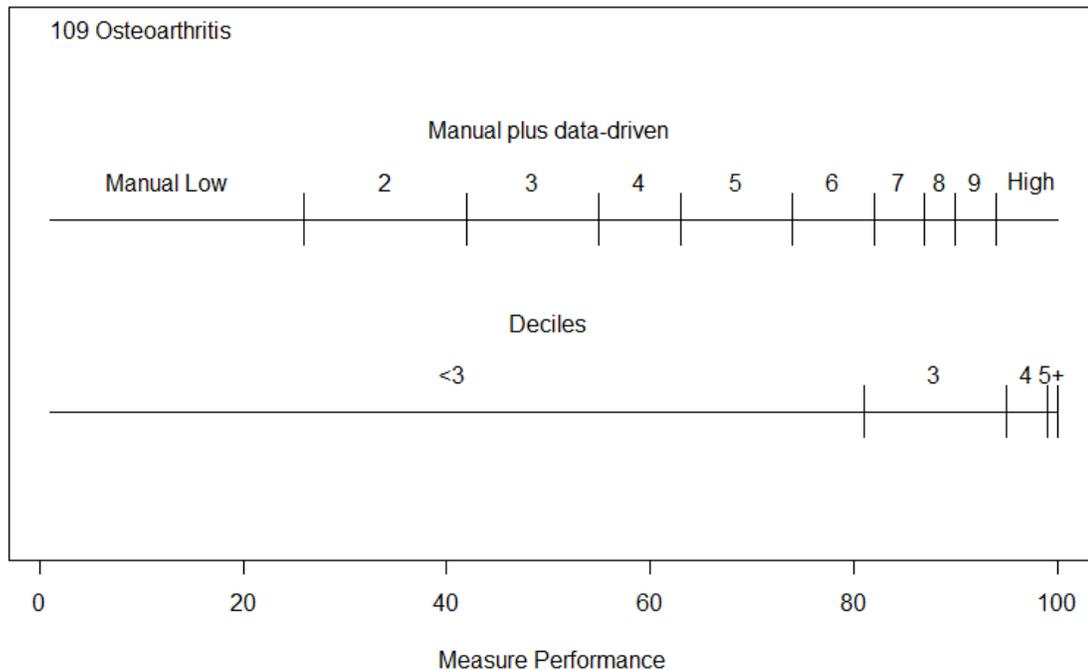


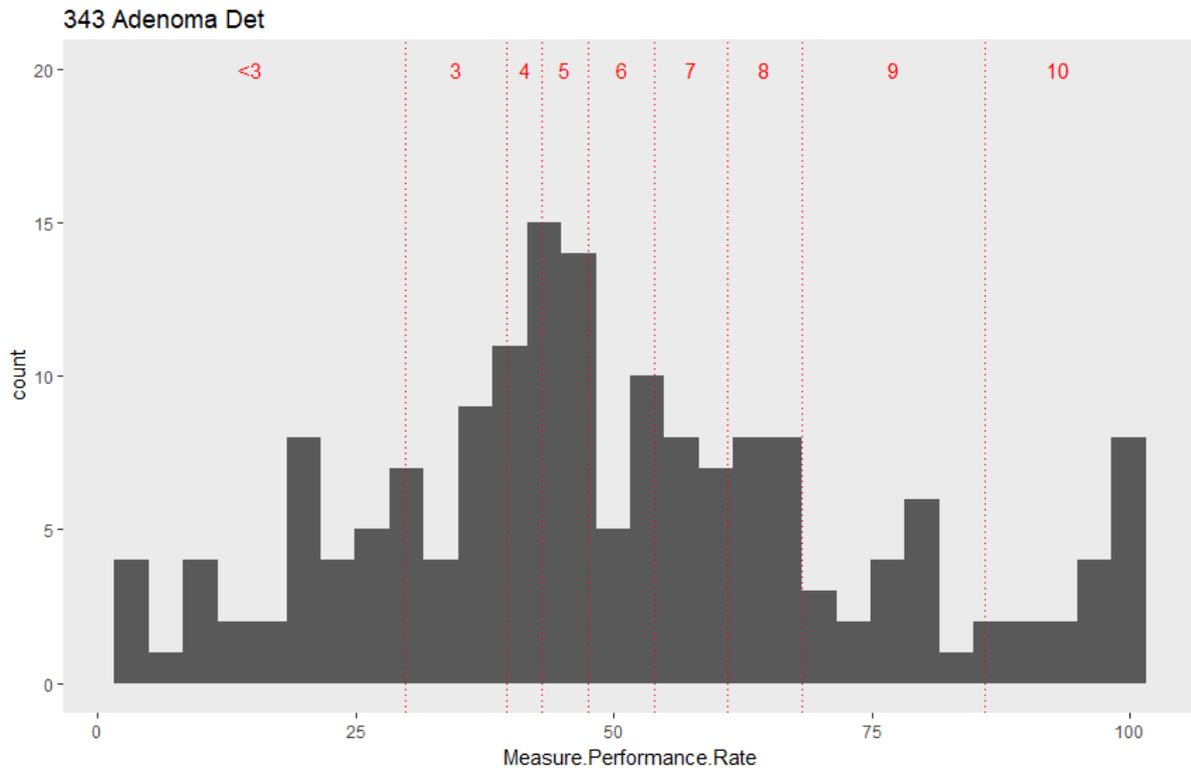
Figure 9c. Comparison of cut-offs for the Osteoarthritis measure



Here you can see that under the M+DD cut-offs, high performance is still required to achieve a high rating, but performance slightly less than 100% is not punished as severely as for the decile method. The cut-offs of 25% and 95% are somewhat arbitrary but could differ by measure and perhaps even by year as performance improves. Additionally, using the M+DD methodology one is able to consider the unique nature of each individual measure when determining cut-offs for the highest and lowest categories. This not only promotes flexibility and the use of common knowledge of the measure, but also allows one to consider clinical and/or practical aspects of the measure when setting the appropriate cut-offs. For example, the Adenoma Detection rate is a measure where determining the manual cut-offs could incorporate this type of information. The nature of the measure may suggest that the goal for physicians should be set closer to 75% to reduce the potential for false-positives. As for the lower threshold, guidelines and supporting literature suggests that the detection rate for a mixed gender population should be at least 25%, suggesting that levels below these levels represent lower levels of quality.²³ Figures 10a-c demonstrate how this information could be used to manually set more clinically relevant cut-offs than are produced by the decile method.

²³ Rex DK, Schoenfeld PS, Cohen J, et al. Quality indicators for colonoscopy. *Gastrointestinal endoscopy*. 2015;81:31-53.

Figure 10a.



(Decile 10 is at 86% while the cut-off for the lowest category is 29%)

Figure 10b. M+DD-based cut-offs for the Adenoma Detection Rate measure

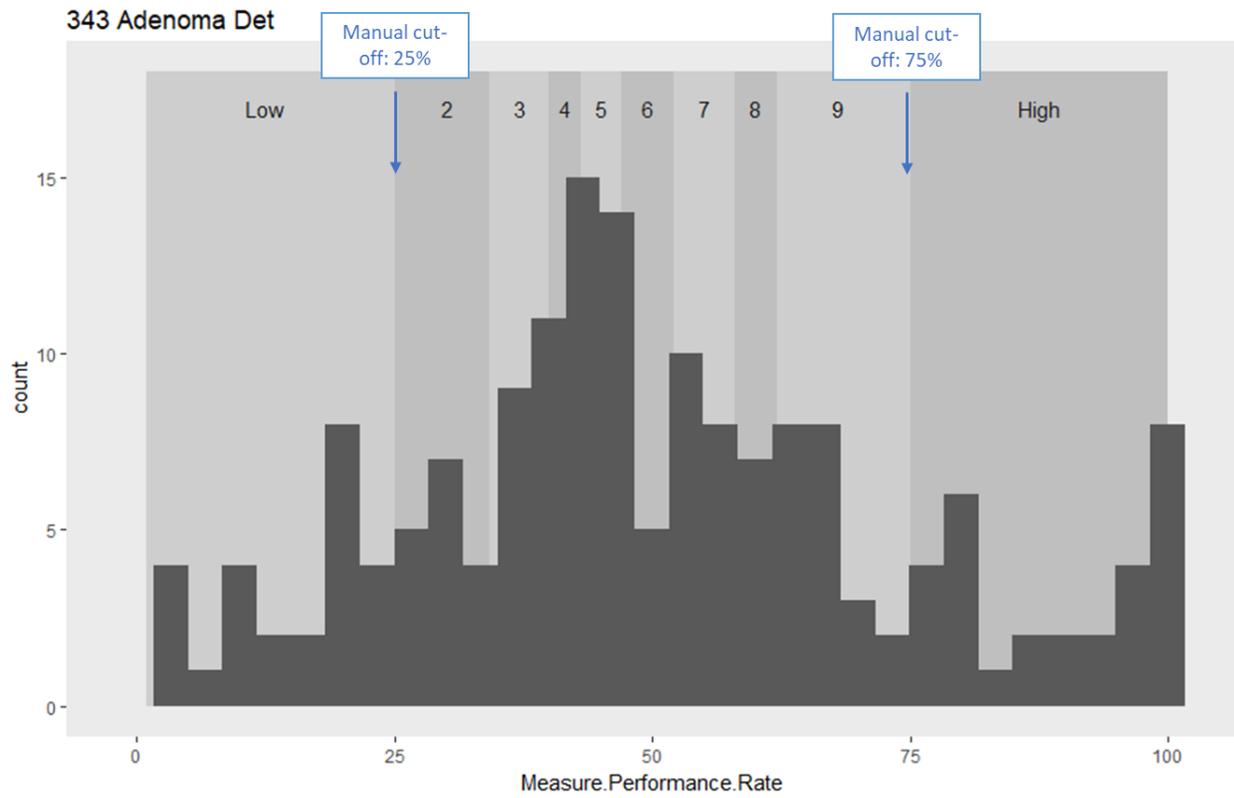
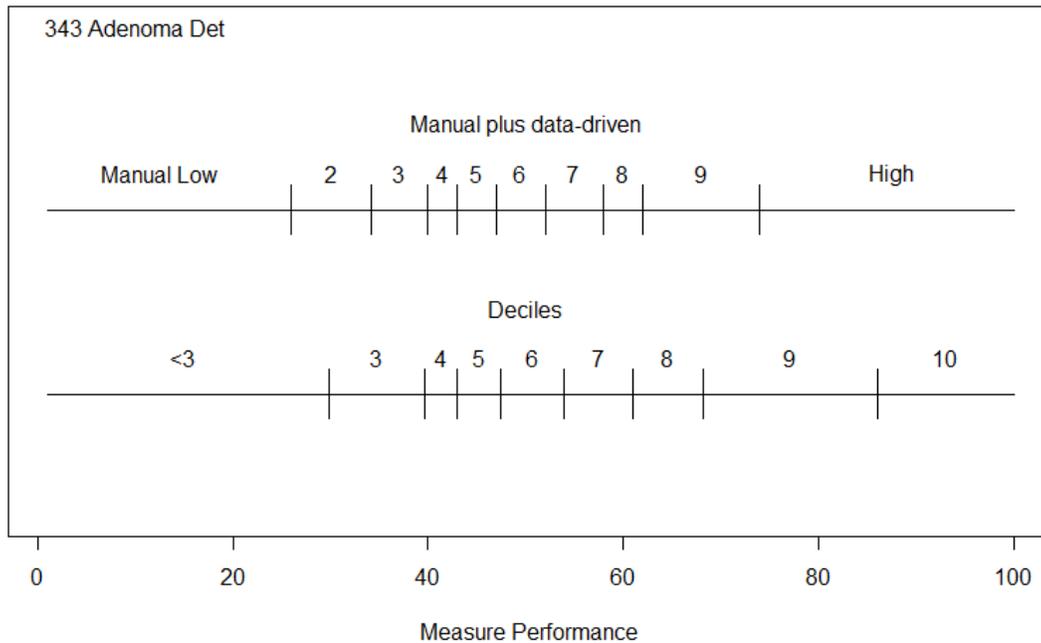


Figure 10c. Comparison of cut-offs for the Adenoma Detection Rate measure



For this measure, the M+DD cut-offs reflect current clinical opinion, practical considerations, *and* incorporate a significant portion of the current performance distribution. By allowing the top and bottom categories to be set manually, it ensures that they truly reflect levels determined to be of the highest and lowest quality. Therefore, we encourage CMS to explore an alternative methodology, such as M+DD for setting the MIPS benchmarks and Physician Compare Star Ratings.

Conclusion

Moving to a uniform and consistent methodology between the two programs will reduce administrative complexity to allow physicians to spend less time on reporting and more time with patients and on improving care, and avoid sending mixed signals to physicians and patients, as well as create a more sustainable MIPS program. It will also produce a more cohesive program and sharpen the focus on outcomes and clinical evidence as opposed to just reporting.