AMA
AMERICAN MEDICAL ASSOCIATION

JAMES L. MADARA, MD
EXECUTIVE VICE PRESIDENT, CEO

ama-assn.org
t (312) 464-5000

March 3, 2015

Andrew Slavitt
Acting Administrator
Centers for Medicare & Medicaid Services
Department of Health and Human Services
Room 445-G, Hubert H. Humphrey Building
200 Independence Avenue, SW
Washington, DC 20201

Re: Physician Compare: Request for Comments Regarding Benchmarking

Dear Acting Administrator Slavitt:

On behalf of the physician and medical student members of the American Medical Association (AMA), I appreciate the opportunity to offer comments to the Centers for Medicare & Medicaid Services (CMS) regarding the selection of benchmarks for the CMS Physician Compare website. The AMA appreciates the Agency's recent series of Physician Compare Benchmark Discussion Webinars, and its request for stakeholder input on the development of a benchmark methodology for the assessment and public posting of physician performance data on Physician Compare. As finalized in the 2014 Physician Fee Schedule Final Rule, a subset of individual eligible professional (EP) quality measures from 2014 Physician Quality Reporting System (PQRS) reporting will be reported on Physician Compare starting in late 2015. CMS currently reports only on Physician Compare group practice reporting option (GPRO) measures. Per the 2015 Physician Fee Schedule final rule, CMS has plans to further expand how data are posted and is currently evaluating and proposing potential benchmarking methodologies for posting 2016 data starting in 2017.

General Concerns with Ratings on Physician Compare

The AMA believes it is premature to move forward with a benchmarking methodology as the PQRS program was designed as a pay-for-reporting program and not a public reporting program. Therefore, PQRS data/information does not lend itself well to public reporting. For instance, medication reconciliation is a very important patient safety concern for patients contemplating knee replacement surgery. However, the appropriate question is "At what points in this episode of care should the medications be reconciled and who is responsible for this action?" Simply saying that every physician should do this is not specific enough to support efficient, patient centered care. As CMS' contractor, Westat, pointed out in the webinar, there are various ways individual EPs and group practices may participate in PQRS, and this is an important factor for public reporting. Unlike other CMS Compare websites and quality programs, such as Hospital Compare or Nursing Home Compare, EPs and group practices do not report in PQRS on a uniform set of measures through one submission mechanism, due to

specialization within medicine, and the differing practice arrangements as well as different types of practice sites.

To complicate the issue further, CMS has changed the PQRS reporting requirements three times in the last three program years, making it extremely difficult for practices to comply. At the same time, CMS has reduced the number of available measures to report on by 50 measures between 2015 and 2014 due to CMS claiming the measures are topped out, even though 75 percent of measures had a successful reporting rate below ten percent. The dearth in the measure portfolio is leaving many specialties, particularly subspecialists, without a sufficient set of measures to report on and/or being forced to report on quality measures that are not meaningful for a patient to assess care. The CMS Physician Compare Technical Expert Panel (TEP) even points out the value in maintaining a more consistent measure set for public reporting over time, as the inconsistency may be confusing to health consumers.

As this and other quality programs continue to evolve, we urge CMS officials to keep in mind that many physicians are already near—or have surpassed—the breaking point in meeting the various quality metrics and submission methods required for PQRS, the Electronic Health Records Incentive Program/Meaningful Use (MU), and the Value-Based Payment Modifier (VBM). Practices have invested tens of thousands of dollars in EHR systems, and the AMA, along with many specialties, have invested millions in developing quality measures and registries. Converting to ICD-10 will pose another hurdle for many practices. Consequently, we also urge CMS to consider adopting benchmarks that employ or build upon the existing assessment criteria in these programs, and avoid creating yet another, independent, non-streamlined quality performance assessment and feedback report for Physician Compare.

The AMA continues to have serious concerns about the accuracy and validity of reporting information about individual physicians and other EPs, versus group practices, on Physician Compare. A likely foreseeable problem CMS will run into when moving from group level reporting to individual is having an adequate sample size to make a statistically valid comparison between physicians and practices. CMS is already grappling with this issue under the VBM program, where CMS only has sufficient data for 30 percent of large group practices and CMS has enough data for groups of ten only three percent of the time.

However, if CMS moves forward, the AMA strongly urges the Agency to continue to proceed cautiously and afford ample opportunity for public comment in setting any forthcoming benchmark methodology and associated targets for Physician Compare. We appreciate CMS acknowledging the challenges with expanding the Physician Compare website and informally stating on the February 19, 2015 CMS Physician Compare webinar (hosted by Westat) that it would not move forward with reporting more robust information until 2017, based upon 2016 reporting. However, prior to CMS formally adopting a particular methodology, it should be thoroughly tested and evaluated by CMS in a transparent manner. This is crucial to ensure that the benchmark methodology is statistically valid, reliable, and accurately rates physicians and other EPs and that the quality metrics are meaningful to patients and relevant to physician specialties and practices, assess performance and/or outcomes which are truly within the control of that physician, and are risk adjusted to avoid penalizing physicians who serve a high number of patients with multiple and complex health problems and/or low socio-demographic status.

Unfortunately, health providers and quality experts have raised questions about whether current quality programs actually ensure or improve the quality of patient care.[1] Therefore, a website such as Physician Compare may actually be misleading to consumers. If CMS decides to move forward, once a methodology has been tested and vetted, it should be proposed in public rule-making (such as the Physician Fee Schedule proposed rule) to allow physicians and other stakeholders to fully review, vet, and comment on the proposed methodology. **EPs should also be afforded at least a two-year advance notice before any methodology takes effect, especially if CMS moves forward with newly proposed performance goals.**

Operational Principles for Physician Compare

The AMA appreciates the modifications that CMS has made to improve the data and operability of Physician Compare, and we urge the Agency to consider the following operational principles in guiding further improvements and modifications.

Updating Demographic Data: CMS should strive to make timely updates to the demographic data for individual EPs and group practices on an ongoing basis, incorporating recent edits made in the Provider Enrollment, Chain, and Ownership System as well as relevant claims data, and quickly correct errors pinpointed by physicians themselves. The current disclaimer that it may take up to six months to correct information on Physician Compare is completely unacceptable, especially if CMS moves forward with expanding the website. We urge CMS to commit to updating demographic data within four weeks of being notified.

Appeals Process and Preview Period: The AMA continues to urge CMS to implement a more robust appeals process for contesting Physician Compare information, and to expand the preview period for an EP to review their information beyond the current 30 days to 90 days. And if an EP or group practice files an appeal and flags their demographic data or quality information as problematic, CMS should postpone posting their information until the issues are resolved. It often takes medical practices several weeks and sometimes months to register and obtain their PQRS reports and Quality and Resource Use Reports (QRURs). It is also unclear how CMS plans to widely notify EPs of the preview period for reviewing their public ratings. We anticipate potential problems and backlogs with obtaining reports, as CMS greatly expands all of its quality programs and moves to profile all EPs.

Ensuring Statistical Validity under ACA: We also believe CMS is in violation of section 10331 (b)(6) of the Affordable Care Act (42 U.S.C. § 1395w-5), which requires CMS to have processes to ensure that timely statistical performance feedback is provided to physicians concerning the data reported under any program that is subject to public reporting on Physician Compare. Currently, CMS only provides 30 days for an EP or group practice to review their Physician Compare performance information before it goes live on Physician Compare. CMS also does not provide a feedback report to an EP or group practice until approximately six to nine months after the close of the reporting period, which provides **no** opportunity for a practice to improve their performance until they are well into the next reporting period. Therefore, it

[1] See, e.g.: McKethan, A. Jha, A. Designing Smarter Pay-for-Performance Programs. *JAMA 312; 24.* December 24/31 2014; Hershberger, P, Bricker, D. Who Determines Physician Effectiveness? *JAMA 312:24.* December 24/31, 2014; Panzer, R.J., Gitomer, R.S., et al. Increasing Demands for Quality Measurement. *JAMA 310: 18.* Nov. 13, 2013.

takes two years before a practice can actually utilize a feedback report to improve their PQRS performance, which forms the basis of what is reported on Physician Compare.

Notice of Chosen Methodology: EPs and group practices should be afforded at least two years' advance notice before any benchmark methodology takes effect, to afford practices the opportunity to prepare and change workflows and implement a quality improvement strategy.

Individual vs. Group Reporting: As noted above, section 10331 of the ACA requires any public reporting of performance information to be statistically valid and reliable. The AMA is concerned with CMS' proposal to utilize a sample size of 20 patients as testing of performance measures. Such a small sample size often demonstrates low reliability. Therefore, CMS should only continue to report at the group practice level and not at the individual EP level.

Star Ratings: The AMA supports efforts to make medical standards more comprehensible to patients. However, we continue to have serious concerns about applying star ratings or similar systems, and the dangers of misleading patients by displaying the ratings in an over-simplified graphic. Star ratings can artificially inflate and magnify minor variances in performance scores which are not statistically different. The overwhelming majority of physicians likely fall within a small range of average quality. CMS should bear this in mind and strive to identify and single out only the true outliers.

Minimum Patient Sample: We continue to believe that a 20 patient minimum sample of patients is insufficient. Acumen, on behalf of CMS, tested measures at the group practice rate using at least 25 measure-eligible cases for a select set of GPRO web-interface measures. Therefore, CMS should test measures and composites with a 20-patient attribution and provide an opportunity for public review and comment on the results to ensure reliability and validity.

Reporting Mechanisms: Comparisons between EPs and/or group practices should only be based on the same mechanism which an EP or group practice used for reporting. For instance, if an EP participates in PQRS via claims, they should only be compared against EPs who reported via claims. When evaluating PQRS results, CMS has pointed out in the past there may be different results regarding a measure's reliability depending upon which reporting mechanism was used to report the measure. A data element for a measure in claims may be captured differently and more or less consistently than its counterpart in an EHR. The National Quality Forum's evaluation criteria also require that a measure submitted for endorsement can only be considered for the data source in which it is specified and tested.[2]

Risk Adjustment: CMS should expand its risk adjustment methodology to incorporate race, income, and region type, to avoid inaccurate conclusions about quality and performance measurement that could unfairly penalize physicians who treat a number of socio-disadvantaged patients. While case mix may not play a role in certain structure and process measures, risk adjustment must occur for measures that are not fully within the measured providers' control. Ignoring factors such as patients' socioeconomic and sociodemographic situations could lead to the conclusion that physicians and practices that serve low-income patients provide lower quality care than those serving high-income patients, when the difference in scores may actually be due to differences in patient mix rather than differences in quality of care

[2] National Quality Forum Measure Evaluation Criteria.
http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx

provided. To hold physicians accountable for different outcomes, without recognizing the patient factors that contribute to those differences, would unfairly penalize them. CMS' own contractor, Acumen, demonstrated this potential problem when it tested the Diabetes Mellitus measure (DM) composite that is part of the GPRO web-interface. Acumen tested the DM composite with expanded risk adjustment to include demographic and regional characteristics (i.e., race, region, region type, household income, and home value), and the results differed from the original performance assessment.

Stakeholder Input and Communications: It would be beneficial and enhance the transparency of Physician Compare for CMS to allow public comment on the deliberations of the Physician Compare TEP. Currently, the public has no opportunity to participate and comment on the TEP's recommendations. CMS should also engage regularly with physicians and other stakeholders. With Hospital Compare, CMS conducts monthly to quarterly calls with affected stakeholders, discussing plans for expansion and notifying them in advance of the release of information. The AMA would be happy to convene something similar with the specialty societies and CMS.

Comments on Proposed Benchmark Methodologies

Composite-Level Benchmark: While we support the use of composite measures, they must be reviewed and tested. It is incorrect to presume that when individual measures are assembled into a composite, the composite will be just as valid and measure practices just as accurately as the individual measures. The overwhelming majority of measures in PQRS were developed as individual measures and most EPs report individual measures, not composites or measures groups. Even the composites that currently exist in PQRS were arbitrarily created by CMS and were not tested for use as a composite.

"Peer Group Comparison": The AMA seeks clarification in terms of how CMS would define region type or population served. We support the concept of stratifying benchmarks by "peer groups," particularly by specialty and region, but need more information to accurately evaluate the proposal. CMS would also have to take into consideration subspecialization within a specialty to allow a consumer to make an accurate comparison.

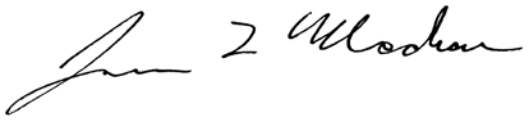
Weighted Average: With a weighted average methodology, CMS would need to have a better way of delineating between specialists, subspecialists, and high-cost specialties. Within the VBM, CMS lumps all physicians within a particular specialty together, including those in very different subspecialties. A more sensible approach to making comparisons would be to weight by related conditions. The flaw with the alternative weighted average methodology that CMS has utilized in the past is that high-cost specialties are automatically penalized. For instance, within internal medicine, an oncologist would automatically have higher costs than an internist or other primary care physician, given the greater expenses required for treating patients with cancer. Thus, a weighted average cannot lead to accurate results, without stratifying by specialty and subspecialty.

Performance Goal: The AMA seeks further clarification by CMS on utilizing the Government Performance and Results Act as a basis for constructing the performance goal methodology. While the performance goal methodology sounds like a suitable methodology for quality improvement, operationalizing it will be very difficult given the poor feedback loops back to EPs and group practices by CMS. We also would not support any performance goal of 100 percent, as there are instances when meeting a measure would be contrary to appropriate care for certain patients and there would need to be

room for exceptions. Also, to be considered a successful reporter under PQRS, an EP or group only has to report the measure successfully 50 percent of the time. Furthermore, as more practices become reliant on third parties such as EHR or registry vendors to participate in PQRS, they are dependent on an outside entity being ready at the start of the performance period to ensure they can comply with PQRS and a performance goal set by CMS.

We appreciate this opportunity to share the views of the AMA regarding the Physician Compare website. If you should have any questions regarding this letter, please feel free to contact Koryn Rubin, Assistant Director, Division of Federal Affairs, at koryn.rubin@ama-assn.org, or 202-789-7408.

Sincerely,



James L. Madara, MD

cc: PhysicianCompare@Westat.com