AMA
AMERICAN MEDICAL ASSOCIATION

JAMES L. MADARA, MD
EXECUTIVE VICE PRESIDENT, CEO

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

January 13, 2015

George Isham, MD
Elizabeth McGlynn, PhD
Co-Chairs
Measure Applications Partnership
National Quality Forum
1030 15th Street, NW, Suite 800
Washington, DC 20005

Re: MAP 2015 Considerations for Implementing Measures in Federal Programs: Draft for Public Comment, December 2014

Dear Dr. Isham and Dr. McGlynn:

The American Medical Association (AMA) is pleased to have the opportunity to comment on the Measure Applications Partnership (MAP) December 2014 draft report, *MAP 2015 Considerations for Implementing Measures in Federal Programs*. We commend the talent and dedication of the MAP staff in synthesizing the key themes of the various workgroups. This is a major achievement under a very short timeline. Continuing our longstanding commitment to quality improvement initiatives that enhance the quality of care provided to patients, the AMA offers the following comments.

Process and Approach for MAP Pre-Rulemaking Deliberations 2015

We agree that providing a preliminary analysis by staff was helpful in improving efficiencies in the review process and allowed the workgroups to better focus on areas where consensus needed to be built. That said, given the importance of this analysis, we would like to point out an area of concern around the third question that asks whether the measure under consideration is tested for the appropriate setting and/or level of analysis and if not, then could it be adjusted to use in the program (page 5, Process and Approach). Additional detail on how the National Quality Forum (NQF) staff was able to make the determination on that second question, when the measure did not currently meet the program setting and/or level of analysis, should be described in this report. Adjusting a measure to apply to a different setting or to attribute to a different provider can be very complex. Asking staff to make an evaluation on whether this is reasonable in the absence of data to demonstrate adequate reliability and validity, and to ensure appropriate attribution and assurances of sufficient sample sizes, does not seem prudent. We support the use of "conditional support" when the measure construction differs from the program needs. However, decisions recommending that it would be appropriate to modify the measure must be based on solid evidence and data and not be subjective.

In addition, we believe that there is an error in box 2 of Figure 2, MAP Preliminary Analysis Algorithm for Earlier Stage Measures on page 6, Process and Approach. It would seem more likely that the logic

would be "No" in between boxes 2 and 3 since the purpose of this analysis is to look at measures that are not yet fully developed and/or endorsed.

Clinician-Overarching Themes

"High Value Measures" or "Measures that Matter"

Moving to more "high value" measures or "measures that matter" are important goals of the AMA and the house of medicine. However, as the MAP report points out, there is not uniform agreement on which measures have the greatest, or incrementally more, value in driving results. What might be considered high value may differ by specialty or patient, as well as vary depending upon its intended purposes. The AMA acknowledges the need to incorporate more outcome measures. However, in many areas of patient care we do not yet have high quality outcome measures with enough specificity to drive improvement. In these situations, structure and process measures that align closely to an outcome based on research play a very important role in improving care, as they provide insight into the changes that need to occur to support improvement in outcomes. In addition, it is important to note that outcome measure development is inherently more complex and difficult than process measure development. There are a number of methodological considerations that warrant careful review and analysis, including perhaps most notably risk adjustment.

Outcome measures at the physician level can be particularly challenging to construct for two primary reasons—small sample sizes and the difficulty of identifying outcomes for which the physician can be held accountable (i.e., those outcomes that are not largely dependent on factors outside of the quality of care received). For example, outcome measures are more reasonable for surgical procedures as opposed to chronic conditions. Infrastructure challenges may also prevent measure developers from developing such outcome measures. These can involve problems with capturing patient reported or experience of care measures in the electronic health record (EHR), as well as measures reliant upon EHR interoperability. Furthermore, as we begin to address measuring overuse and appropriateness in various clinical domains, it must be acknowledged that there are patient scenarios which are not addressed by the current available evidence and not all patient scenarios may easily led themselves to capture by appropriateness and overuse criteria. Areas that fall in this grey zone are not good subjects and focus areas for performance measurement.

In reality, performance measurement is a relatively new science in health care and therefore, is changing at a rapid pace. Given the current measure development cycle, we are probably three to five years away from having a set of measures that are meaningful and appropriate for widespread use by multiple stakeholders. Therefore**, we disagree with the suggestion that linking "greater incentives for those who report on high-value measures might prompt faster development of high-value measures in other condition/topic areas."** This statement assumes that individual physicians can wield sufficient influence on which measures are developed, when development and testing takes significant time and resources and individual physicians are not the sole developers of measures for CMS programs. **Holding physicians accountable for something that is not necessarily within their direct control seems imprudent, and we recommend that this sentence be removed.**

We are glad to see the MAP acknowledge that measurement gaps could arise when measures are removed from programs, and the challenge this is posing. For the 2015 Physician Quality Reporting System (PQRS), the Centers for Medicare & Medicaid Services (CMS) removed 50 measures from the program, increased the requirements threefold for avoiding a payment cut, and required physicians to report on one

"cross cutting" measure, based upon recommendations made in the 2014 MAP report. The changes in requirements coupled with the elimination of measures are leaving many physicians, especially subspecialists, unable to meaningfully participate in PQRS. These new requirements are not positively contributing to valuable performance improvement at the point of care; they are turning the program into a check-box exercise for many physicians. A flexible approach is critical to ensuring that relevant measures are available to as many physicians as possible. Many new physicians will begin participating in the PQRS, EHR Incentive/Meaningful Use (MU) and Value-Based Payment Modifier (VM) programs as they are now strictly in penalty phases, and existing physicians must be supported to ensure the continued ability to participate in the programs. The MAP must also keep in mind that PQRS participation serves as the gateway for success or failure under the VM.

Aligning Measurement Requirements

We applaud CMS for taking steps to further align requirements with the various physician quality programs and are glad to see the MAP highlight measure alignment as a high priority. Measure alignment is one step in reducing the administrative burden with the various quality programs. We continue to believe more work must be done to fully align the programs and allow physicians to report once to avoid payment adjustments. In order for PQRS reporting to count towards MU Quality Reporting, a physician must take into consideration the following detailed rules and requirements:

- PQRS quality measures must be reported for a full year, as opposed to 90 days, so first-year MU participants must report twice.
- Some of the MU electronic clinical quality measures (eCQMs) include "look back" or "look forward" periods requiring data outside of the PQRS and VM reporting periods. If CMS cannot calculate a performance rate for that eCQM, a physician would be subject to both PQRS and VM penalties.
- For MU, it is acceptable to report zeroes on measures (including not having any denominator eligible patients given the limited number of eCQMs). This is not permissible for the PQRS EHR reporting option or any other option under PQRS. If a physician does not have any data on Medicare patients (i.e., none of their Medicare patients fall into the denominator of any of the quality measures which their EHR incorporates), then the physician needs to report separately for PQRS.
- Physicians are only able to report MU measures through the EHR reporting option, even though there are a limited number of eCQMs that do not meet the needs of all physician specialties. If a physician reports under PQRS through claims, registry, or qualified clinical data registry, it will not count towards MU.

An additional area of much needed alignment involves the various quality reporting requirements and measures for physicians under Medicare Advantage (MA) plans. Each plan has its own unique set of various measures and requirements that are often unaligned with the PQRS program, the MU quality requirements, and/or VM. A portion of the requirements are often based on the CMS requirements set for the MA 5-Star Rating program. However, there are instances when the measure(s) are not applicable or even conflict with appropriate physician care, and the plan does not allow exclusions in such circumstances because the plan would score poorly. CMS has set benchmarks of 90 percent or higher for plans to achieve 5 stars on administrative claims measures, which the AMA does not support. Thus, in certain instances, the quality reporting requirements actually penalize physicians for delivering appropriate and effective care.

MAP 2015 Considerations for Implementing Measures in Federal Programs: Clinicians

Background

We recommend that the wording in the last paragraph on page 3, Clinicians, be reviewed to clarify its intent and potential redundancy with respect to the beginning of the report.

Parsimony and Alignment Across Programs

We caution the MAP to assume that moving toward a "greater focus on selecting composite measures, appropriate use measures, and outcomes can promote greater parsimony and reduce the burden of measurement for professionals" (first paragraph, page 7, Clinicians). Moving toward these high-value measures will require additional diligence and improvement efforts on the part of providers and it will not necessarily promote reduced measurement burden. We recommend that this sentence be revised to reflect the increased efforts these types of measures will require of providers and not downgrade the extent of this work.

Incentives for More Meaningful Measurement

We strongly support the recommendations to encourage steady financial support for measure development and potential avenues to encourage increased participation while leveraging current measure development, testing, and maintenance needs. However, as the field of measure developers expands, there is an increased risk of un-harmonized measures and duplicative efforts. It is also prudent to consider the need to provide incentives to coordinate efforts and co-produce CQMs. It is imperative that measure developers have the necessary expertise with eCQM standards currently in use such as the Quality Data Model, Health Level-7 (HL7), and HQMF (Health Quality Measures Format) eMeasure, and are involved in national efforts focused on the future direction of health care standards.

As part of this financing effort, we urge the MAP to make the recommendation for CMS to undertake a study, working with the AMA-convened Physician Consortium for Performance Improvement® (PCPI®) and other affected stakeholders, on evaluating the current PQRS measure portfolio given the significant reduction in the number of available measures, as well as the emphasis on "high value measures." This analysis should identify and evaluate: which measures are clinically relevant to which specialty, which modality(ies) support PQRS measure reporting, the amount of effort, resources, and time needed to prepare certain measures for EHR and/or registry capture, and how many measures are available for reporting for each of the six NQS domains by specialty and sub-specialty. This measure portfolio evaluation will help CMS, medical specialties, and measure developers understand the gaps as it relates to individual medical specialties and identify priority areas to better focus resources for quality measure development, testing, and implementation.

The MAP also points out that eligible professionals (EPs) must invest in infrastructure to support reporting data, such as EHRs and registries, and there is an urgent need for available sites to test measures. To resolve this dilemma, the MAP suggests that CMS consider innovative incentives, such as waiving non-participation penalties in the various quality programs. The AMA supports this concept, and we would strongly support statutory changes that would reinstate incentives within the physician quality programs. The AMA has long believed that physicians should be supported for improving care. We also would like to point out that increasing incentives under a budget-neutral program such as the VM also leads to penalties for other participants, as budget neutrality creates winners and losers. The AMA does

not support this approach, as we believe as a fundamental principle that quality should be rewarded and its value should not be diminished by creating artificial competitions and comparisons between physicians.

Considerations for Specific Programs

Physician Quality Reporting System, Physician Compare, Physician Value-Based Payment Modifier

We suggest that the MAP review the program objectives outlined for PQRS, as several appear to be statements rather than true program objectives (such as the second and third bullets). We have concerns and do not recommend that all available measures in PQRS be made available under Physician Compare, especially new measures. Physicians should have at least two years of experience reporting on a measure before the results are publicly reported on Physician Compare. This phase-in is needed to gain experience with the measure and a better understanding of the reporting requirements for the measure, especially if there is a movement towards more outcome measures within PQRS. We are also concerned that many measures developed for PQRS were not tested for use on Physician Compare. In addition, until a measure has been fully specified it should not be made available under MAP review for use on Physician Compare. It is impossible to fully evaluate a measure for intended use until construction is complete. Furthermore, the AMA does not support public reporting of outcomes measures that are attributed to individual physicians, as these should only be attributed to a group. At the aggregate level, a physician's patient mix is very diverse and it is often very difficult to accurately attribute key outcomes to the individual physician.

We also request for future MAP deliberations to include exact categorization of measures. The Quality and Resource Use Reports (QRURs) and VM programs have multiple facets and it would be helpful if CMS actually categorized the measures based upon how they are classified in the programs. Without providing the specific bucket(s) the measures fall under, it is difficult to provide feedback on the appropriateness of the measures. For instance, for VM, CMS calculates a cost and quality score. However, the quality score has two buckets. One quality bucket deals strictly with PQRS measures and the other quality bucket consists of claims-based measures that CMS automatically calculates, regardless of whether it is appropriate to apply that measure to a particular physician or if it was tested at the individual physician level.

The AMA urges the MAP to be mindful and fully consider the measure development lifecycle when making its recommendations to CMS. Last year the MAP classified many measures as "topped out" and recommended the need for a core set of measures, but didn't take into consideration the implementation logistics. Therefore, its recommendation added to the complexity of the program. We continue to believe it is premature and short-sighted to remove a measure as "topped out" simply because it is deemed as having a high performance rate, particularly when the EP reporting rates within the PQRS program are so low. There is now a significant gap in the measure portfolio due to the number of measures CMS removed for reporting in 2015, with minimal advanced warning to physicians or measure owners. The AMA also does not believe that performance rates alone provide a valid reason to consider a measure "topped out." Removal from PQRS of any measure as "topped out" must be based upon consideration of several factors, including reporting rate and performance rate, at a minimum. A higher reporting rate or threshold may be indicated before decisions are made on measures based solely upon performance rates. Additionally, high performance rates (close to 100 percent) on some measures among reporting EPs may be partly attributable to intensified improvement efforts motivated by the reporting opportunities.

Therefore, removal of these measures from PQRS may result in a drop in performance as well as the quality of care.

Therefore, **we urge the MAP to recommend that CMS require at a minimum a three-year phase out period for removal of any measures, to allow the submission of new measures within the current Call for Measures timeframe and prevent gaps in the measure portfolio.** The proposed timeframe is also consistent with our previous request of CMS for three years of stability in CMS' quality programs to allow EPs time to adjust to changes in programs and make improvements in their practice. Under the current process for incorporating new measures into physician quality programs, CMS requires a measure developer to submit a measure two years prior to the start of the program year. To consider a measure for the 2016 PQRS program, CMS had to receive the measure information by June 2014—a gap of almost two years. Prior to the requirement of MAP review, it took only six to twelve months for a measure to be included in a CMS program.

Medicare Shared Savings Program

The MAP Clinician Workgroup reportedly reviewed 107 PQRS measures for potential reporting by accountable care organizations (ACOs) in the MSSP. The Workgroup specifically supported additional MSSP measures involving "high-value," multiple chronic conditions, multiple or across settings, overuse, composite measures (including imaging overuse), cross-cutting measures (such as for pain management); outcome measures (especially for cancer and accuracy of diagnosis and screening methods), and HEDIS health plan measures aligning with Medicare Advantage and private health plans.

The AMA believes the MAP has failed to take into account the design and purpose of ACOs, and how these additional measures could actually impair ACO operations and ultimately harm Medicare beneficiaries. It is important to ensure that ACO savings are not achieved through withholding or limiting care, but there is no direct link between more quality measures and better protection for beneficiaries. Excessive measures would make the program overly burdensome, deter ACO participation, and deny patients the benefits of better care coordination. Adding measures increases the burden on ACOs with no compensation for the additional time, and requires ACOs to shift focus to different aspects of clinical care, change data collection and analysis systems, etc. Frequent changes are a recipe for failure of this vital Medicare program. The AMA has urged CMS to provide more stability for Medicare ACOs by setting quality standards for the entire three-year agreement period and only changing them during that time if both CMS and the majority of ACOs agree the change is needed.

It is not feasible to include all quality measures in the ACO program, or to align ACO measures with all other programs or private payers. Clear principles are needed to guide measure priority and selection. Quality measures should be primarily designed to *protect* beneficiaries from inappropriate reductions in services by ACOs. We believe ACOs have a disincentive to provide preventive care services and might be underused, thus having a measure on the use of preventive care services could help protect patients against possible underuse. Providing expensive preventive services to these patients increases the beneficiaries' total cost of care, which counts against the ACO financial benchmarks, reducing the chances that they will achieve savings to share and increasing the chances that they will incur losses. These calculations are made each year, so in reality the only investments that are in ACO's interest to make are those that have a return within that same year. If the ACO spends $50,000 to provide colonoscopies to 100 patients which prevent hundreds of thousands of dollars in advanced cancer care for a couple of those people six months later, that would be worthwhile. But tests that prevent costs years down the road do not help the ACO. Therefore, a quality measure could help address this imbalance in

financial incentives and encourage ACOs to engage in outreach to get people in for colonoscopies and other preventive care.

If the MAP wants quality measures to *improve* care for beneficiaries, then the measures should focus on areas where: (a) beneficiaries are receiving poor care today; and (b) it is feasible for an ACO to make changes in care that would improve care in those areas using the limited resources available in the shared savings program. If the goal is quality improvement—rather than preservation of current quality—then the shared savings formula needs to be restructured to ensure that adequate resources are directed to providers to achieve this.

With respect to the MAP Workgroup's specific suggestions for potentially adding new measures:

- *"High-value," cross-cutting, and "topped-out" measures:* The same rationale for why these initiatives are bad for physicians and other clinicians applies equally for ACOs. Measures should be retired when they are no longer supported by clinical evidence or their use could lead to undesirable consequences. Measures should not be retired simply because they are "topped-out."
- *Managing multiple chronic conditions:* ACOs are accountable for total spending, and thus have a natural incentive to improve coordination of care for multiple chronic conditions. There is no need for additional measures to address this issue.
- *Measures for/across multiple settings:* Adding measures specific to particular settings can be very problematic if ACOs change the mix of patients in different settings. For example, if an ACO arranges for more patients to receive home health care rather than care in a skilled nursing facility (SNF) after discharge, the acuity level of patients in both home health and SNF in that community will be higher than in other communities, which could make the ACO appear worse on quality measures in those settings. If fewer patients use SNFs, then the smaller, higher acuity group of SNF patients might have a higher readmission *rate*, even though there would be fewer readmissions in total than before.
- *Overuse/composite measures:* The design of the ACO program acts as an incentive to underuse, not overuse, particular services, because the overuse of services decreases potential ACO savings. The use of imaging services has been decreasing for the past several years. Under the Protecting Access to Medicare Act, Medicare will require the consultation of "appropriate use criteria" for ordering of advanced diagnostic imaging. Thus we fail to see the need for overuse measures.
- *Outcome measures:* We believe it is important to move away from process-based measures toward outcome-based measures. However, the measured outcomes must be within the control of the ACO, with effective risk adjustment to avoid penalizing ACOs that manage the care of beneficiaries with more needs and to avoid causing access problems for such beneficiaries.
- *HEDIS health plan measures:* Aligning ACO measures with those of other programs and private plans is ideal in theory. But the MAP must also consider the considerable burden of adding new measures on individual ACOs, which detract from the resources needed to maintain a high level of care.

Medicare and Medicaid EHR Incentive Programs

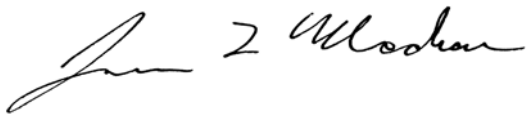
The MAP notes that the eCQMs under consideration tend to be limited by today's EHR environment reality rather than push to a system of greater interoperability and health information exchange in which measurement is readily performed without additional burden on the providers. While the AMA understands the desire to push the envelope, especially if the goal is to reduce physician burden, it is very

expensive to make changes and capture the necessary information that the MAP would like to see with eCQMs within the EHR. The majority of EHRs are episode based, so often it is not feasible to capture longitudinal information to measure outcomes, especially when a given patient is unlikely to receive all health care services through a single integrated delivery system.

Conclusion

We appreciate the opportunity to provide our comments and look forward to continuing our work with the MAP to ensure adoption of quality measures in physician federal quality programs that result in effective and broad participation and improvements in the delivery of care. Should you have any questions, please feel free to contact Koryn Rubin, Assistant Director, Federal Affairs at koryn.rubin@ama-assn.org or 202-789-7408.

Sincerely,



James L. Madara, MD