

Maternity Multi-Stakeholder Action  
Collaborative Session 3:  
Quality Measures Part 2



# Table of Contents

<b>Recommended Steps to Selection Performance Measures Selection for Incorporation into a Maternity Alternative Payment Model .....</b>	<b>3</b>
<b>New York Department of Health Maternity Care Quality Measure Summary .....</b>	<b>6</b>
<b>New York Department of Health Maternity Care Value Based Payment Quality Measure Set Measurement Year 2017 .....</b>	<b>17</b>
<b>Developing a State-based Quality Measurement Program Using an Episode-of-Care Framework: Recommendations for State Purchasers .....</b>	<b>22</b>
<b>Considerations for State Development of Performance Measure Sets .....</b>	<b>30</b>



**Recommended Steps to Performance Measures Selection for Incorporation into a Maternity  
Alternative Payment Model  
For use in the LAN Maternity Action Collaborative (MAC) Quality Measurement Meetings  
March 3 and March 20, 2017**

Objective: Assist MAC kick-off meeting participants to understand the process steps to take in order to develop a measure set to be used for a maternity APM.

The aim of performance measurement in maternity care alternate payment models is to accelerate movement to high-value maternal-newborn care, through better care, better outcomes and experience, and wiser spending. The following steps outline a process for developing a measure set. They are informed by experience across multiple states in measure set development exercises. While it is important to follow most of the steps in an order, some iterative discussion will naturally occur, especially with the later steps.

**Step 1: Determine who should be participating in the measure selection process.**

- Just those party to the contract(s), or other interested parties (e.g., consumers)?
- How large a group?
- Mix of clinical and measurement expertise?

**Step 2: Determine whose performance is to be measured.**

*Options include:*

- Maternity care providers
- Hospitals
- Birth centers
- Neonatologists
- Pediatricians
- Some combination of the above

**Step 3: Identify the intended use(s) of the measure set.**

*Options include:*

- To adjust payment in some fashion (exactly how is a separate conversation)
- To monitor performance without financial consequence
- To feedback performance information to service providers for use in quality improvement
- To test new measures for potential future use

**Step 4: Identify the criteria to be used to inform measure selection.**

*Sample criteria for individual measures*

1. Evidence-based and scientifically acceptable
2. Has a relevant benchmark
3. Not greatly influenced by patient case mix
4. Fosters accountability for outcomes, using woman-reported data whenever feasible
5. Consistent with the goals of the program
6. Feasible to collect
7. Aligned with other measure sets
8. Promotes increased value
9. Addresses an opportunity for maternity care quality improvement
10. Potential to transform maternity care quality, outcomes and value
11. Sufficient denominator size

*Criteria for the measure set as a whole*

1. Representative of the array of services provided, including prenatal, intrapartum and postpartum/newborn phases of care
2. Representative of the diversity of patients served
3. Not unreasonably burdensome to payers or providers
4. Measures multiple levels of care, including facility and clinician/group
5. Includes -- whenever feasible -- woman-reported outcome and experience of care measures

**Step 5: Identify the process by which measure selection decisions will be made.**

- Group consensus or voting?
- One or more rounds of review?
- Explicit (e.g., with scoring) or implicit use of selection criteria?

**Step 6: Identify populations and performance domains for measurement**

*Options include:*

- Populations: all women, women with substance use disorder, women with mental illness, other women with high-risk pregnancies
- Performance domains: Prenatal Care, Labor and Delivery, General Newborn, High-Risk Newborns, Maternal Complications, Emergency Care, Postpartum Care

**Step 7: Identify candidate measures.**

*Options include:*

- Measures currently in use by participating providers and payers
- Measures found in national measure sets
- Measures that address a priority opportunity for performance improvement
- Items evaluated in research studies that can fill crucial measure gaps, such as woman-reported outcome and experience of care measures

**Step 8: Identify potential data sources and operational means for obtaining data, including timeliness.**

*Options include:*

- Clinical data – from EHRs and/or HIE (if available)
- Claim data
- Survey data – provider and patient

**Step 9: Estimate desired size of the measure set.**

**Step 10: Determine whether all-payer or payer-specific data will be used in contracts.**

**Step 11: Begin measure selection process by reviewing individual measures.**



**Department  
of Health**

# Maternity Care Quality Measure Summary

Draft

May 2016

NYS Medicaid Value Based Payment

# Maternity Clinical Advisory Group (CAG) Quality Measure Recommendations

## Introduction

Over the course of three meetings, the Maternity CAG has reviewed, discussed and provided feedback on the proposed maternity bundle to be used to inform value based payment contracting for Levels 1-3.

A key element of these discussions was the review of current, existing and new quality measures used to measure relevant for the maternity bundle. This document summarizes the discussion of the CAG and their categorization of outcome measures.<sup>5</sup>

## Selecting quality measures: criteria used to consider relevance<sup>6</sup>

In reviewing potential quality measures for utilization as part of a VBP arrangement, a number of key criteria have been applied across all Medicaid member subpopulations and disease bundles. These criteria, and examples of their specific implications for the Maternity VBP arrangement, are the following:

### Clinical relevance

#### **Focused on key *outcomes* of integrated care process**

*i.e. outcome measures (postpartum depression) are preferred over process measures (screening for postpartum depression); outcomes of the total care process are preferred over outcomes of a single component of the care process (i.e. the quality of one type of professional's care).*

#### **For process measures: crucial evidence-based steps in integrated care process that may not be reflected in the patient outcomes measured**

*i.e. focus on postpartum contraceptive care is key but will not be captured in outcomes of current maternity episode*

#### **Existing variability in performance and/or possibility for improvement**

*i.e., blood pressure measurement during pregnancy is unlikely to be lower than >95% throughout the State*

### Reliability and validity

#### **Measure is well established by reputable organization**

*By focusing on established measures (owned by e.g. NYS Office of Patient Quality and Safety (OQPS), endorsed by the National Quality Forum (NQF), Healthcare Effectiveness Data and Information Set (HEDIS) measures and/or measures owned by organizations such as the Joint Commission, the validity and reliability of measures can be assumed to be acceptable*

---

<sup>5</sup> The following sources were used to establish the list of measures to evaluate: existing DSRIP/QARR measures; AHRQ PQI/IQI/PSI/PDI measures; CMS Medicaid Core set measures; other existing statewide measures; NQF endorsed measures; measures suggested by the CAG.

<sup>6</sup> After the Measurement Evaluation Criteria established by the National Quality Forum (NQF), [http://www.qualityforum.org/uploadedFiles/Quality\\_Forum/Measuring\\_Performance/Consensus\\_Development\\_Process%E2%80%999s\\_Principle/EvalCriteria2008-08-28Final.pdf](http://www.qualityforum.org/uploadedFiles/Quality_Forum/Measuring_Performance/Consensus_Development_Process%E2%80%999s_Principle/EvalCriteria2008-08-28Final.pdf)

### **Outcome measures are adequately risk-adjusted**

*I.e. measuring ‘% preterm births’ without adequate risk adjustment makes it impossible to compare outcomes between providers*

### **Feasibility**

#### **Claims-based measures are preferred over non-claims based measures (clinical data, surveys)**

*I.e. ease of data collection data is important and measure information should not add unnecessary burden for data collection*

#### **When clinical data or surveys are required, existing sources must be available**

*I.e. the vital statistics repository (based on birth certificates) is an acceptable source, especially because OQPS has already created the link between the Medicaid claims data and this clinical registry*

#### **Data sources preferably are patient-level data**

*Measures that require random samples (e.g. sampling patient records or using surveys) are less ideal because they do not allow drill-down to patient level and/or adequate risk-adjustment, and may add to the burden of data collection. An exception is made for such measures that are part of DSRIP/QARR.*

#### **Data sources must be available without significant delay**

*I.e. data sources should not have a lag longer than the claims-based measures (which have a lag of six months). This is an issue with the vital statistics repository, for example, which have a one year lag (at least for the NYC data).*

### **Meaningful and actionable to provider improvement in general**

**Measures should not only be related to the goals of care, but also something the provider can impact or use to change care.**

## **Categorizing and Prioritizing Quality Measures**

Based on the above criteria, the CAG discussed the outcome measures in the framework of three categories:

- **Category 1** – Category 1 is comprised of approved outcome measures that are felt to be clinically relevant, reliable and valid, and feasible.
- **Category 2** – Category 2 outcome measures were felt to be clinically relevant, valid and probably reliable, but where the feasibility could be problematic. These outcome measures should be investigated during the 2016 or 2017 pilot but would likely not be implementable in the immediate future.
- **Category 3** – Category 3 measures were decided to be insufficiently relevant, valid, reliable and/or feasible.

Ultimately the use of these measures, particularly in Category 1 and 2 will be developed and further refined during the 2016 (and possibly 2017 pilots). The CAG will be re-assembled on a yearly basis during at least 2016 and 2017 to further refine the Category 1 and 2 measures.

The HCl<sup>3</sup> grouper creates condition-specific scores for Potentially Avoidable Complications (PACs) for each condition. The ‘percentage of total episode costs that are PACs’ is a useful measure to look for potential improvements; it cannot be interpreted as a quality measure. PAC counts however, can be considered clinically relevant and feasible outcome measures. For Maternity Care, however, the PAC counts are low, and the events that the grouper considers to be PACs are not all considered validated outcome measures by the CAG. (Individual PACs may be ‘mined’ to be considered to be future quality measures, such as post-partum depression etc.)



## Maternity CAG Recommended Quality measures – Category 1 and 2

	#	Measure	Measure Steward/Source
<b>Category 1</b>	1	Frequency of Ongoing Prenatal Care	National Committee for Quality Assurance
	2	Prenatal and Postpartum Care (PPC)	National Committee for Quality Assurance
	3	% of Vaginal Deliveries With Episiotomy*	Christiana Care Health System
	4	Vaginal Birth After Cesarean (VBAC) Delivery Rate	Office of Quality and Patient Safety (eQARR)
	5	C-Section for Nulliparous Singleton Term Vertex (NSTV) (risk adjusted)*	The Joint Commission
	6	% of Early Elective Deliveries*	The Joint Commission
<b>Category 2</b>	7	Antenatal Steroids*	The Joint Commission
	8	Antenatal hydroxyl progesterone	Texas Maternity Bundle
	9	Experience of Mother With Pregnancy Care	New
	10	Appropriate DVT Prophylaxis in Women Undergoing Cesarean Delivery*	Hospital Corporation of America
	11	Intrapartum Antibiotic Prophylaxis for Group B Streptococcus (GBS)*	Massachusetts General Hospital
	12	Birth Trauma Rate – Injury to Neonate	Agency for Healthcare Research & Quality- Quality Indicators
	13	Live Births Weighing Less than 2,500 Grams (risk adjusted)	Bureau of Vital Statistics
	14	% Preterm births	Bureau of Vital Statistics
	15	Under 1500g Infant Not Delivered at Appropriate Level of Care*	California Maternal Quality Care Collaborative
	16	Postpartum Blood Pressure Monitoring	Texas Maternity Bundle
	17	LARC uptake	CMS - set of 'Contraceptive Use Performance Measures' for Medicaid
	18	Neonatal Mortality Rate	New York State Prevention Agenda
	19	Hepatitis B Vaccine Coverage Among All Live Newborn Infants Prior to Discharge*	Centers for Disease Control and Prevention
	20	% of Babies Who Were Exclusively Fed with Breast Milk During Stay*	The Joint Commission
	21	Monitoring and reporting of NICU referral rates	New
* = NQF Endorsed			

**CAG categorization and discussion of measures**

Topic	#	Quality Measure (* = NQF Endorsed)	Type of Measure	Measure Steward/Source	DSRIP	QARR	HEDIS	Data Required		Category	Quality Measure Categorization & Comments
								Medicaid Claims Data	Vital Statistics <sup>1</sup>		
Prenatal Care	1	Frequency of Ongoing Prenatal Care	Process	National Committee for Quality Assurance/HEDIS		X	X	YES	-	1	Scores high on all criteria. HEDIS measure in QARR.
	2	Prenatal and Postpartum Care (PPC)	Process	National Committee for Quality Assurance / HEDIS		X	X	YES	-	1	Scores high on all criteria. HEDIS measure in QARR.
	3	Behavioral Health Risk Assessment	Process	American Medical Association – convened Physician Consortium for Performance Measurement <sup>®</sup> (AMA-PCPI)				NO	YES	3	Low relevance since this measure only looks at whether or not the screening was done. Vital statistics data on this topic have limited reliability. Postpartum depression is being considered as a Potentially Avoidable Complication (PAC) in the Maternity bundle.
	4	Antenatal Depression Screening	Process	Texas Maternity Bundle				NO	NO	3	As the previous measure, with the addition that this measure is not included in the vital statistics dataset.
Screening / Prevention	6	Risk-Appropriate Screening During Pre-Natal Care Visits (Gestational Diabetes)	Process	AHRQ guideline: National Collaborating Centre for Women's and Children's Health. Antenatal care: routine care for the healthy				NO	YES	3	Clinically relevant, but should be focused on broader set of risk factors. More relevant to focus on outcome measure – many of the complications of not doing this screening properly will be captured as Potentially Avoidable Complications (PACs). Risk-appropriate screening is currently an OPDS quality improvement target. Measures that may be forthcoming from this project could at a later stage be considered by the CAG.

Topic	#	Quality Measure (* = NQF Endorsed)	Type of Measure	Measure Steward/Source	DSRIP	QARR	HEDIS	Data Required		Category	Quality Measure Categorization & Comments
								Medicaid Claims Data	Vital Statistics <sup>1</sup>		
				pregnant woman.							
	7	Antenatal Steroids*	Process	The Joint Commission				NO	YES	2	Clinically very relevant because it is a key intervention to increase the incidence of fetal maturation (reduce respiratory distress, reduce intraventricular hemorrhage, and reduce neonatal death). The size of the relevant population is small. In addition, the quality of these data in the vital statistics is deemed to be questionable. Given the clinical relevance, these issues merit further attention during the 2016 Pilot. One concern that was mentioned was that 'receiving the full course' could be too high a goal.
	8	Antenatal Hydroxyl Progesterone	Process	Texas Maternity Bundle				NO	YES	2	Clinically very relevant because it is a key intervention to reduce the incidence of preterm births. The size of the relevant population is small. In addition, the quality of these data in the vital statistics is deemed to be questionable. This specific intervention is not yet an established process measure. Given the clinical relevance, these issues merit further attention during the 2016 Pilot.
	9	Antenatal Blood Pressure Monitoring	Process	Not available				NO	NO	3	Low feasibility and low clinical relevance because of expected uniformly high score.
Organization	10	Shared Decision Making	Process	Informed Medical Decisions Foundation				NO	NO	3	This measure was suggested by clinical experts. Although the clinical relevance is high, the feasibility is low and this is currently not standard practice.
Experience	11	Experience of Mother With Pregnancy Care	Outcome	New				NO	NO	2	To be further discussed during pilot. The experience (or perhaps even Patient Reported Outcomes) of maternity care is of course highly clinically relevant and a focus on this quality aspect is a core element of DSRIP and the NYS Medicaid VBP roadmap. The feasibility of this measure, however, is currently very low, because the required data for this measure is currently not even gathered.

Topic	#	Quality Measure (* = NOF Endorsed)	Type of Measure	Measure Steward/Source	DSRIP	QARR	HEDIS	Data Required		Category	Quality Measure Categorization & Comments
								Medicaid Claims Data	Vital Statistics <sup>1</sup>		
Vaginal Delivery	12	% of Vaginal Deliveries With Episiotomy*	Process	Christiana Care Health System				NO	YES	1	Episiotomies are increasingly seen as mostly unnecessary. Scores high on all criteria.
	13	Degree Perineal Laceration During Vaginal Delivery	Outcome	Beth Israel Deaconess Medical Center				NO	YES	3	The CAG considered this measure to create the wrong incentive: overuse of C-sections or episiotomies was seen as a worse side effect than the (small) chance on significant lacerations. Moreover, this is already captured as a PAC.
	14	Vaginal Birth After Cesarean (VBAC)	Process	Office of Quality and Patient Safety (eQARR)	X			NO	YES	1	Key QARR measure, calculated by OQPS.
	15	C-Section for Nulliparous Singleton Term Vertex (NSTV) (risk adjusted)*	Outcome	Office of Quality and Patient Safety (eQARR)		X		proxy	YES	1	Key QARR measure, calculated by OQPS.
C-Sections	16	Appropriate DVT Prophylaxis in Women Undergoing Cesarean Delivery*	Process	Hospital Corporation of America				NO	NO	2	Clinical relevance is high: preventing DVT in maternity care in general is one of the three major initiatives of the motherhood initiative in NYS, together with post-partum hemorrhage and high post-partum blood pressure. During the pilot, a discussion with ACOG NYS will be continued on the feasibility of linking their database to MIDW data.
	17	Appropriate Prophylactic Antibiotic Received Within One Hour Prior to	Process	Massachusetts General Hospital / Partners Health Care System				NO	NO	3	Information not available. Can't tell when the antibiotic is given. Process measure; outcomes are captured in PACs.



Topic	#	Quality Measure (* = NOF Endorsed)	Type of Measure	Measure Steward/Source	DSRIP	QARR	HEDIS	Data Required		Category	Quality Measure Categorization & Comments											
								Medicaid Claims Data	Vital Statistics <sup>1</sup>													
Prevention	18	Surgical Incision for Women Undergoing Cesarean Delivery* Intrapartum Antibiotic Prophylaxis for Group B Streptococcus (GBS)*	Process	Massachusetts General Hospital				NO	NO	2	As DVT prophylaxis.											
												20	Obstetric Trauma Rate – Vaginal Delivery With Instrument	Outcome	Agency for Healthcare Research & Quality - Quality Indicators				YES	YES	3	The CAG considered this measure to create the wrong incentive: overuse of especially C-sections to reduce this score was seen as a worse side effect than the (small) chance on significant lacerations. Moreover, this is already captured as a PAC.
Overall	22	% of Early Elective Deliveries*	Outcome	The Joint Commission	X			NO	YES	1	DSRIP measure. High score on all criteria											

Topic	#	Quality Measure (* = NOF Endorsed)	Type of Measure	Measure Steward/Source	DSRIP	QARR	HEDIS	Data Required		Category	Quality Measure Categorization & Comments
								Medicaid Claims Data	Vital Statistics <sup>1</sup>		
POST DELIVERY MOTHER CARE	23	Live Births Weighing Less than 2,500 Grams (risk adjusted)	Outcome	Bureau of Vital Statistics		X	X	proxy	YES	2	Clinical relevance is high, and measure is widely used and part of QARR. Yet CAG members question how much influence providers really have on this outcome. Ethnicity can play a significant role. Adequacy of risk adjustment needs to be further investigated during pilot (there is already a very advanced model created by OQPS).
	24	% Preterm births	Outcome	Bureau of Vital Statistics	X			NO	YES	2	Although this is a DSRIP measure, this is a Domain 4 measure, reported at State level and not risk-adjusted. Given the importance of this topic, this could be further investigated during the pilot.
	25	Under 1500g Infant Not Delivered at Appropriate Level of Care*	Process	California Maternal Quality Collaborative				NO	YES	2	Clinical relevance high. Also important measure to 'counteract' potential unwanted effect of saving costs by underutilizing adequate but more costly care. Can create difficult discussions on access of care. To be investigated.
	26	Prenatal and Postpartum Care (PPC)	Process	National Committee for Quality Assurance / HEDIS	X		X	YES	-	1	Measure discussed above (prenatal care).
	27	Postpartum Blood Pressure Monitoring	Process	Texas Maternity Bundle				NO	NO	2	Clinically relevant, but data is currently absent.
Monitoring	28	Postpartum Depression Screening	Process	American College of Obstetricians and Gynecologists				NO	NO	3	It's important to do the screening, but even more important to have the correct follow up. The follow up is not measured with this indicator.
	29	Postpartum Glucose Intolerance / Diabetes Screening	Process	Suggested by ACOG, CDC and ADA				NO	NO	3	It's important to do the screening, but even more important to have the correct follow up. The follow up is not measured with this indicator.
Screening											



Topic	#	Quality Measure (* = NOF Endorsed)	Type of Measure	Measure Steward/Source	DSRIP	QARR	HEDIS	Data Required		Category	Quality Measure Categorization & Comments		
								Medicaid Claims Data	Vital Statistics <sup>1</sup>				
Contraceptive Use	30	Use of Most or Moderately Effective Contraceptive Services, Postpartum	Process	CMS - set of 'Contraceptive Use Performance Measures' for Medicaid				YES	NO	2	Highly relevant, feasible and valid. Reliability requires additional investigation. CAG suggests broadening the measure to overall contraceptive use (not merely counseling). A caveat is that it is difficult to establish a percentage that is 'adequate', since simply striving to 'as high as possible' would create a dangerous incentive.		
								YES	YES			2	Clinical relevance is high. Small numerators may create low reliability, and risk adjustment needs to be adequate.
								YES	NO				
NEWBORN	31	Neonatal Mortality Rate	Outcome	National Committee for Quality Assurance / HEDIS				YES	YES	2	Scores high on all criteria, except possibly the room for improvement.		
								YES	NO				
								YES	NO				
Overall	32	Hepatitis B Vaccine Coverage Among All Live Newborn Infants Prior to Discharge*	Process	Centers for Disease Control and Prevention				YES	NO	2	High score on all criteria, the CAG suggests that some adaptations are made to the definition. 'Exclusive' seems inappropriately strict. Combining breastfeeding with bottle feeding in the beginning can help rather than higher ongoing breastfeeding. Options could be to modify the measure to "predominantly breastfed" rather than "exclusively breastfed". These data are available in vital statistics.		
								NO	YES				
NICU	33	% of Babies Who Were Exclusively Fed with Breast Milk During Stay*	Process	The Joint Commission				NO	YES	2	It will be critical to monitor the referral rates to Level 4 to ensure providers are not over-referring babies to Level 4 level of care.		
								YES	YES				
Referral Rates	34	Monitoring and reporting of NICU referral rates	Process	New				YES		2			

1. Source: <http://www.nyc.gov/html/doh/downloads/pdf/birth/limited-us08.pdf>  
 2. CMS has created a set of 'Contraceptive Use Performance Measures' for Medicaid. The indicator: '% of women ages 15-44 who are at risk of unintended pregnancy that adopt or continue use of long-acting reversible contraception (LARCs)' is on that list. [www.medicare.gov/medicaid-clip-program-information/dx-tb01sc/qualtrics-data/download/contraceptive-measures-fact.pdf](http://www.medicare.gov/medicaid-clip-program-information/dx-tb01sc/qualtrics-data/download/contraceptive-measures-fact.pdf)  
 3. Neonatal Mortality Rate is a key public health measure that is part of the State's Prevention Agenda ([www.health.ny.gov/prevention/prevention\\_agenda/healthy\\_mothers/](http://www.health.ny.gov/prevention/prevention_agenda/healthy_mothers/))  
 4. Claim data can identify specific conditions. If these measures are only for preterm babies, we need the vital statistics to identify the prematurity.



## Appendix A:

### Meeting Schedule

<b>Date</b>	<b>Agenda</b>
<b>CAG #1</b> 7/21/2015	Part I A. Introduction to Value Based Payment B. Clinical Advisory Group Roles and Responsibilities C. HCl <sup>3</sup> 101- Understanding the HCl <sup>3</sup> Grouper and Development of Care Bundles  Part II A. Maternity Bundle – Definition
<b>CAG #2</b> 8/11/2015	1. Bundle Criteria 2. Characteristics of the Maternity Population in the Medicaid Data 3. Risk Adjustment for Maternity Care 4. Performance Measurements
<b>CAG #3</b> 9/9/2015	1. Welcome & Recap 2. Outcome Measures for Maternity Episode 3. Conclusion and Next Steps





**Department  
of Health**

# Maternity Care

Value Based Payment Quality Measure Set  
Measurement Year 2017

Updated March 13, 2017

The 2017 Maternity Care quality measure set was created in collaboration with the Maternity Clinical Advisory Group (CAG) and the New York State (NYS) Value Based Payment (VBP) Workgroup. The measure set is closely aligned with existing measures sets used in the Delivery System Reform Incentive Payment (DSRIP) Program, the Quality Assurance Reporting Requirements (QARR) and the State's Vital Statistics maternity care measures. The measure set is intended to encourage providers to meet high standards of patient-centered clinical care and care coordination across multiple care settings throughout the maternity care episode.

The measure set includes measures classified by category based on an assessment of reliability, validity, and feasibility, and according to suggested method of use (either Pay for Reporting (P4R) or Pay for Performance (P4P)).

## MEASURE CLASSIFICATION

In June of 2016, the Maternity CAG [published recommendations](#) to the State on quality measures, data, and support required for providers to be successful. Additionally the report addressed other implementation details related to a VBP Maternity Arrangement. Upon receiving the CAG recommendations, the State conducted further feasibility review and analysis to define a final list of measures for inclusion during the 2017 VBP Measurement Year (MY). Each measure has been designated by the State as Category 1, 2, or 3 with associated recommendations for implementation and testing for future use in VBP arrangements.

### Category 1

Category 1 quality measures as identified by the Maternity CAG and accepted by the State are to be reported by VBP

Contractors. These measures are also intended to be used to determine the amount of shared savings for which VBP contractors are eligible<sup>1</sup>.

The State classified each Category 1 measure as either P4P or P4R:

- **P4P** measures are intended to be used in the determination of shared savings amounts for which VBP Contractors are eligible. Measures can be included in both the determination of the target budget and in the calculation of shared savings for VBP Contractors.
- **P4R** measures are intended to be used by the Managed Care Organizations (MCOs) to incentivize VBP Contractors for reporting data to monitor quality of care delivered to

### Categorizing and Prioritizing Quality Measures



#### CATEGORY 1

Approved quality measures that are felt to be both clinically relevant, reliable and valid, and feasible.



#### CATEGORY 2

Measures that are clinically relevant, valid, and probably reliable, but where the feasibility could be problematic. These measures should be investigated during the 2017 pilot program.



#### CATEGORY 3

Measures that are insufficiently relevant, valid, reliable and/or feasible.

<sup>1</sup> A Path toward Value Based Payment: New York State Roadmap for Medicaid Payment Reform. Annual Update. June 2016.



members under the VBP contract. Incentives for reporting will be based on timeliness, accuracy, and completeness of data. Measures can be reclassified from P4R to P4P through annual CAG and State review or as determined by the MCO and VBP Contractor.

### **Categories 2 and 3**

Category 2 measures have been accepted by the State based on agreement of measure importance, validity, and reliability, but flagged as presenting concerns regarding implementation feasibility. These measures will be further investigated in the VBP pilots. The State requires that VBP Pilots select and report a minimum of two Category 2 measures per VBP arrangement (or have a State and Plan approved alternative). VBP Pilot participants will be expected to share meaningful feedback on the feasibility of Category 2 measures when the CAGs reconvene. The State will discuss measure testing approach, data collection, and reporting requirements with VBP pilots at a future date.

Measures designated as Category 3 were identified as unfeasible at this time or as presenting additional concerns including accuracy or reliability when applied to the attributed member population for the maternity arrangement. Several measures in the original CAG report were removed for this reason and therefore no longer in the Category 1 or 2 measure list. These measures will not be tested in pilots or included in VBP at this time.

### **MEASUREMENT YEAR 2017 MEASURE SET**

The measures and State determined classifications provided on the following pages are recommendations for MY 2017. Note that measure classification is a State recommendation and implementation is to be determined between the MCO and VBP Contractor.

Measure sets and classifications are considered dynamic and will be reviewed annually. Updates will include additions, deletions, reclassification of measure category, and reclassification from P4R to P4P based on experience with measure implementation in the prior year. During 2017, the CAGs and the VBP Workgroup will re-evaluate measures and provide recommendations for MY 2018.



## Category 1

The Category 1 table displays the Category 1 Maternity Measure set, arranged alphabetically, and includes measure title, measure steward, the National Quality Forum (NQF) number and/or other measure identifier (where applicable), and State determined classification for measure use.

Measure	Measure Steward	Measure Identifier	Classification
C-Section for Nulliparous Singleton Term Vertex (NSTV)	The Joint Commission (TJC)	NQF 0471	P4R
Frequency of Ongoing Prenatal Care	National Committee for Quality Assurance (NCQA)	NQF 1391	P4P
Incidence of Episiotomy	Christiana Care Health System	NQF 0470	P4R
Long-Acting Reversible Contraception (LARC) Uptake <sup>2</sup>	United States Office of Population Affairs	NQF 2902	P4R
Low Birth Weight [Live births weighing less than 2,500 grams (preterm v. full term)]	Agency for Healthcare Research and Quality (AHRQ)	PQI 9	P4R
Percentage of Babies Who Were Exclusively Fed with Breast Milk During Stay	TJC	NQF 0480	P4R
Percentage of preterm births	NYS Department of Health (DOH)	-	P4R
Prenatal & Postpartum Care (PPC)—Timeliness of Prenatal Care & Postpartum Visits	NCQA	NQF 1517	P4P
Preventive Care and Screening: Screening for Clinical Depression and Follow-Up Plan	Centers for Medicare and Medicaid Services (CMS)	NQF 0418	P4R

<sup>2</sup> Long-Acting Reversible Contraception (LARC) Uptake is a two-part measure. The State recommends the Contraceptive Care - Postpartum measure be used.



## Category 2

The Category 2 table displays the Category 2 Maternity Measure set and includes measure title, measure steward, and the NQF number and/or other measure identifier (where applicable). All Category 2 measures are classified as P4R in MY 2017.

Measure	Measure Steward	Measure Identifier
Antenatal Hydroxyl Progesterone	New Measure	-
Antenatal Steroids	TJC	NQF 0476
Appropriate DVT Prophylaxis in Women Undergoing Cesarean Delivery	Hospital Corporation of America (HCA)	NQF 0473
Experience of Mother With Pregnancy Care	New Measure	-
Hepatitis B Vaccine Coverage Among All Live Newborn Infants Prior to Discharge	Centers for Disease Control and Prevention (CDC)	NQF 0475
Intrapartum Antibiotic Prophylaxis for Group B Streptococcus (GBS)	Massachusetts General Hospital	NQF 1746
Monitoring and reporting of NICU referral rates	New Measure	-
Postpartum Blood Pressure Monitoring	New Measure	-
Vaginal Birth After Cesarean (VBAC) Delivery Rate, Uncomplicated	AHRQ	IQI 22

Note: VBP Pilot contractors may include measures as outlined in the VBP Pilot webinar held on February 24, 2017. The measure, “Neonatal Mortality Rate” – AHRQ measure NQI# 2 was redacted from the Category 2 list subsequent to that presentation. VBP Pilot Contractors will not be held accountable for reporting this measure.



DECEMBER 2015

# Developing a State-based Quality Measurement Program Using an Episode-of-Care Framework: *Recommendations for State Purchasers*

Prepared by Francois de Brantes, MS, MBA, Health Care Incentives Improvement Institute (HCI<sup>3</sup>)

## Introduction

As the US health care system moves towards value-based payment, it becomes clearer that, while alternative payment models are important, the underlying information processes required to vivify these new payment models are equally critical to the success of the payment model. As much as Patient Centered Medical Homes, Accountable Care Organizations and episode-based payments matter conceptually, the real effort lies in reforming the nature of health care information, or these payment models will languish. Significant gaps in quality of care measurement continue, as do the means for capturing quality of care data and marrying them to cost of care data.<sup>1</sup>

As a system designed for fostering accountability, federal value-based purchasing (VBP) programs have focused on the clinical outcomes of care that rely on Medicare's Physician Quality Reporting System (PQRS)<sup>2</sup> and Hospital Inpatient Quality Reporting (IQR) System<sup>3</sup>, and in some instances, in concentrated local pilots. Both the PQRS and Hospital IQR systems are conveyed through different conduits as defined measures of care. CMS integrates and reports the data in comparative data sets on physician and hospital performance respectively, largely focused on measures of care for Medicare patients.

**Whether or not genuine transformation of the delivery system takes place through the use of new payment models will depend almost entirely on the ability of practicing physicians to have access to timely, reliable and actionable feedback loops on clinical and financial outcomes.** One area where this appears to be paying off in the Medicare program is the penalty for excessive hospital readmissions.<sup>4</sup> By aligning penalties for excessive readmissions with specific comparative reports on hospital performance, CMS has seen reductions in excessive admissions for Medicare patients. State-led efforts can take a cue from this success: incentives coupled to actionable feedback reporting have the potential to give frontline clinicians the tools they need to redesign care.

This Brief outlines action-oriented steps for state purchasers to develop a quality measurement program based on episodes-of-care that leverages existing information technology infrastructure and clinical registries. Specific suggestions for state purchasers include:

1. Start with an incremental approach to quality measurement and pilot, using manual processes.

## ABOUT STATE HEALTH AND VALUE STRATEGIES

State Health and Value Strategies, a program funded by the Robert Wood Johnson Foundation, provides technical assistance to support state efforts to enhance the value of health care by improving population health and reforming the delivery of health care services. The program is directed by Heather Howard at the Woodrow Wilson School of Public and International Affairs at Princeton University. For more informations, visit [statenetwork.org](http://statenetwork.org).

## ABOUT THE ROBERT WOOD JOHNSON FOUNDATION

For more than 40 years the Robert Wood Johnson Foundation has worked to improve the health and health care of all Americans. We are striving to build a national Culture of Health that will enable all Americans to live longer, healthier lives now and for generations to come. For more information visit [www.rwjf.org](http://www.rwjf.org). Follow the Foundation on Twitter at [www.rwjf.org/twitter](http://www.rwjf.org/twitter) or on Facebook at [www.rwjf.org/facebook](http://www.rwjf.org/facebook).

## ABOUT HEALTH CARE INCENTIVES IMPROVEMENT INSTITUTE

Health Care Incentives Improvement Institute (HCI<sup>3</sup>) is a not-for-profit company that designs and implements innovative payment and benefit plan design programs to motivate physicians, hospitals and consumer-patients to improve the quality and affordability of care.

2. Create and publish a master list of data elements required from selected quality measures to appropriately identify current data collection efforts and potential gaps in measurement.
3. Create a central database that leverages existing clinical data registries and utilizes direct provider submission.
4. Develop provider feedback loops that incorporate episode-of-care efficiency metrics, with episode-of-care outcome metrics and synthesize results in a transparent manner.

### State-led VBP: Works in Progress

For states leading the way in value-based purchasing, a “pardon our dust” sign should be considered, which is to say, a work in progress is just that. There is no need to leap to artificial or stopgap measures to give the appearance of completion. Indeed, by rushing towards badly-fashioned, readily and/or publicly available mechanisms that give the appearance of completion, states actually distort information or make it too remote and ambiguous for consumers and providers alike. States need to be frank about shortfalls in publicly reported measures and resist filling them in with measures that can lead to false positives and false negatives (classifying a hospital as being good at everything when it’s not or bad at everything when it’s not).

By emphasizing episode-of-care pathways, as the states of Arkansas, Ohio and Tennessee are doing,<sup>5</sup> gaps in quality measurement can be identified, and where need be, uniquely redesigned. States can address the gaps incrementally and make the most of limited resources by building episode-specific measure sets.

A case example for this incremental approach can be found in the work of Community Health Choice (CHC), a Houston-based Medicaid plan. CHC launched a “womb-to-crib” bundled payment program and tied all of the phases of pregnancy, delivery and newborn care into a single, severity-adjusted global fee. When the plan looked for available data on quality of maternity care, data available to CHC at the time were fragmented and limited. As a result, the plan created a maternity quality scorecard which requires input from clinical record data. Participating providers use manual processes to submit information from medical

records as an initial step. Once results are validated and found useful for clinicians, automated processes can be instituted. Ideally, over time manual processes such as these will get converted to automated data feeds using clinical registries<sup>6</sup> as discussed later in this Brief.

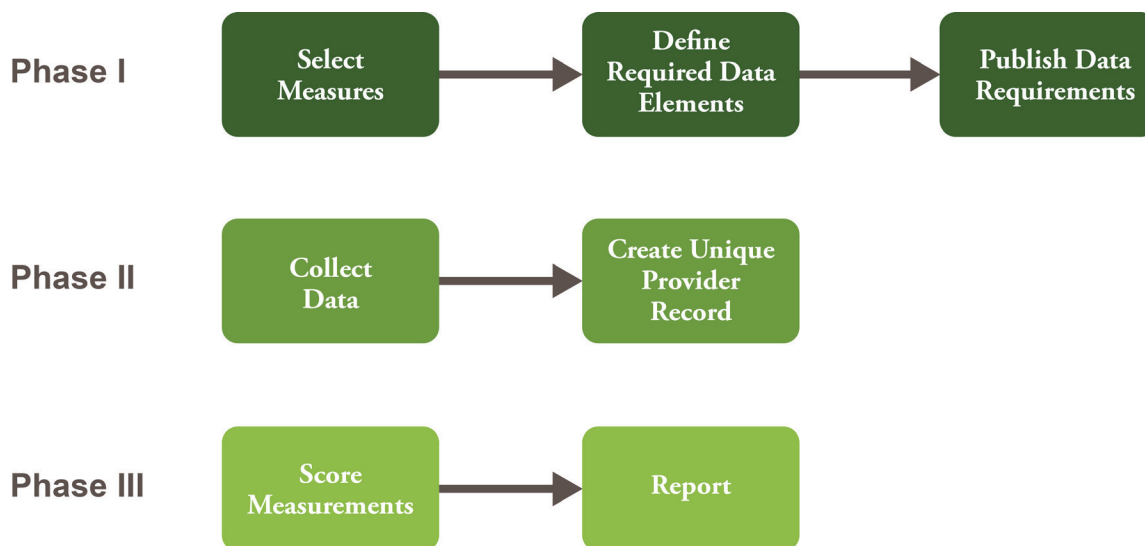
Designing a quality scorecard that matches the scope of the bundle is an essential feedback mechanism for clinicians, providing two critical views of the same episode of care: a financial view and a quality of care view. These views are within the clinicians’ line of sight and highly actionable, making care redesign and other process changes far more likely to happen faster.

Relying on manual processes to get started and ensure provider engagement, when registries are not available, is defensible to engage providers on quality performance. A manual process allows for refinement and modification, and requires only minimal capital investment and modest amounts of labor. Once methods are proven, scaling issues become important, but not until then.

Given the dearth of publicly available measures on the quality of most episodes of medical care,<sup>7</sup> states must roll up their sleeves, work with local provider advisory groups, and develop ad hoc protocols for data collection and reporting. While few meaningful measures are publicly reported, quality measures have been defined for a large number of medical episodes of care and a reasonable subset are being reported and collected through clinical data registries.<sup>8</sup> As part of the technical assistance for Tennessee Medicaid,<sup>9</sup> HCI<sup>3</sup> delineated the availability of measures and the corresponding registries collecting them related to episodes targeted by the state. **Appendix A** provides examples of clinical data registries (CDRs), including those qualified by Medicare, which align with certain episodes of medical care.

### Process for Quality Measuring and Reporting

The following section outlines a three-phase pathway (Figure 1) for establishing, measuring and reporting protocols that enable states to create timely clinical outcome feedback loops by leveraging existing data registries and providing alternative data submission protocols for providers who do not have access to or choose not to use available registries.

**Figure 1: Potential Data Sources and Approach for Quality Reporting**

### ***Phase 1: Selecting Performance Measures and Defining Data Elements***

As noted in *Figure 1*, the first step involves selecting quality performance measures. Using standardized measures and common measure sets reduces the administrative burden and sends a consistent message about provider performance accountability. For additional perspectives on selecting measures, states may wish to review a prior SHVS brief, *“Considerations for State Development of Performance Measure Sets.”*<sup>10</sup>

*Create and publish a master index of candidate data elements:* States should examine clinically related or proximate episodes to reduce potential duplication of data elements being measured. The process for developing performance measures begins with a) the element being measured, for instance, systolic blood pressure, and b) the patients that should be included (and excluded). Data elements for measure sets of related conditions may be used for multiple measures. For example, a measure set often includes measures of superior control (such as number of patients with systolic blood pressure below 120) and measures of poor control (such as number of patients with systolic blood pressure over 140). Noticeably, both of these examples are measuring the same clinical indicator: systolic blood pressure, which can then be used to create a number of quality measures across many episodes of medical care. **It’s essential to create a master index of candidate data elements to determine the overall quantity of such data elements and better indicate to physicians and hospitals the extent of the data collection process.** Publishing a master index helps

all involved with a state effort to determine which data elements they are currently collecting and to identify potential gaps. Gaps can be assessed both in terms of the extent to which those providers for whom the measures will be applied are collecting the data elements, and the number of data elements that need to be collected to create all agreed-upon measures. The result should enable stakeholders to prioritize data collection efforts.

Publishing the list of desired data elements also signals clinical data registries and Electronic Medical Record (EMR) vendors of upcoming demands from physicians and hospitals on extracting data from internal medical record databases. For the vast majority of existing clinical quality measures, required data elements reside, in some fashion, in existing and deployed EMR systems.<sup>11</sup> Our experience suggests that extracting needed data elements from practices, hospitals and health systems with an EMR is not a particularly big challenge. The key is to be clear on the data elements and any other specifications related to a measure for which the data element will apply, for example clinical exclusions. Alternatives to EMRs are discussed in the next section.

### ***Phase 2: Data Collection for Quality Reporting***

**Whether measurement data comes from established registries, directly from providers, or participating health plans, it should be subsumed into a master database and reconciled around single provider records.** Assembled data can then be analyzed to compare the effectiveness of treatments and reported out to providers in a consistent way, irrespective of payers to the extent feasible. This concept is

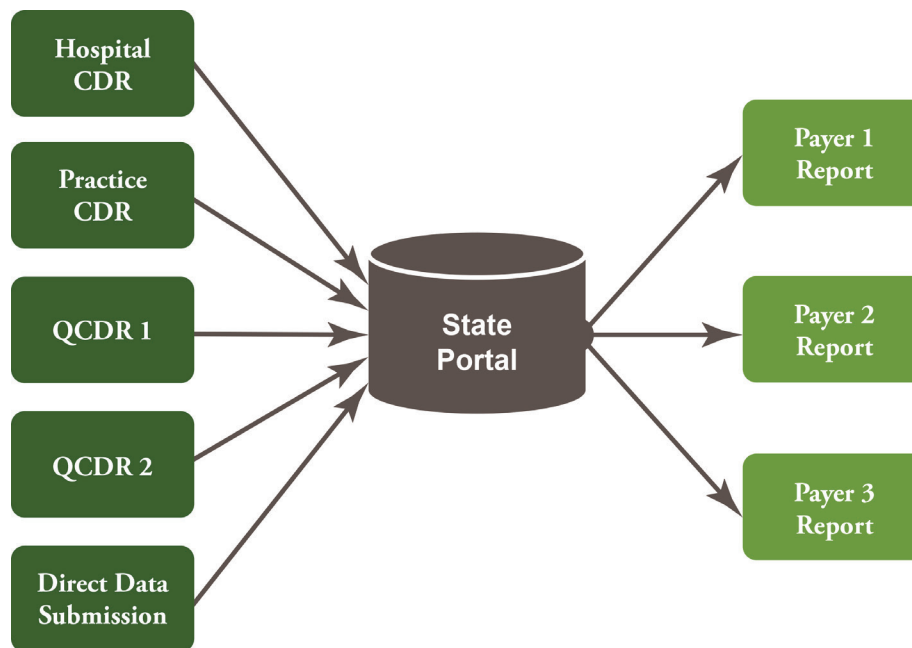


important because the traditional way in which provider performance measurement has been conducted is payer by payer. As a result, provider performance reporting has a tendency to vary by payer, creating confusion.

A centralized scoring mechanism across all of a provider's patients will ensure that feedback to the provider on the quality of care will be the same across all payers. As part

of technical assistance for Tennessee, HCI<sup>3</sup> suggested a data collection and reporting schema as depicted in Figure 2, where the inputs come from hospital and practice Clinical Data Registries (CDR), CMS-authorized CDRs (known as Qualified Clinical Data Registries (QCDR)), and/or direct data submissions from providers, and the outputs are reports to clinicians.

**Figure 2: Potential Data Sources and Approach for Quality Reporting**



QCDRs<sup>12</sup> are registries authorized by CMS to collect quality measures from physicians to satisfy reporting requirements of the Medicare Physician Quality Reporting System. As such, leveraging QCDRs can speed up the process of setting up a data collection infrastructure. Generally speaking, leveraging registries – whether native to an EMR in a provider organization, maintained by a medical specialty society, or qualified by CMS – is the more efficient and effective way of developing a central data collection system. Direct data submission by providers requires instituting a series of processes, including data validation and integrity, that have to be designed from scratch.

#### **Basic decisions for states relative to designing direct data submission portals include identifying:**

- *The purposes of the portal* – Data submission only or data submission and reporting;
- *The scope of the portal* – Whether direct submission will be accepted for all measures/programs or only some;
- *Whether access by parties other than the clinician managing the patient will be allowed* – Many physicians may elect to

have a practice administrator submit data on their behalf;

- *What auditing requirements for sampling of patients included in the direct data submission are necessary* – Typically, direct data submission entails drawing the data from a random sample of patients rather than reporting on all patients.

#### ***Phase III: Measure Scoring and Reporting***

No matter their good intentions, states getting into the process of scoring and reporting on performance should be aware that the physician community tends to view publicly reported clinical and financial performance with deep suspicion. In addition, two decades of measurement reporting have shown that those being measured gravitate towards emphasizing measures that are common with easily attainable thresholds. This has been true at both the federal and state level. Today, little usable physician and hospital quality information exists for the public at large.<sup>13</sup> **As a result, state purchasers should keep these important lessons regarding performance measurement and reporting in mind:**

1. *Measure what matters* – Scorecards should be concise and populated with high impact measures that have a direct relationship with patient outcomes.
2. *Encourage continuous performance* – All measures should be scored using the result of the numerator/denominator calculation, and that result should be applied against the total number of points allocated to each measure. Additions to numerators should yield additional points, so that clinicians have continuous incentives to improve the quality of care.
3. *Make results actionable* – Feedback should be timely and relevant. This means: (a) providing benchmark comparisons and best practice sharing; (b) making clinical reengineering experts available to frontline clinicians; (c) providing knowledge exchange mechanisms to facilitate peer-to-peer interactions (such as online forums).
4. *Make results and reports consistent* – Whenever feasible, states should assess quality of care across payers, not payer-by-payer. Assessing provider performance across all patients avoids a potential sample selection bias and the likelihood that a physician will have varying scores from one payer to another.

*Integrating, Not Reconciling Data Streams:* State agencies spearheading these efforts should be cognizant of the fact that there is a good chance discrepancies will appear between the clinical exclusions/inclusions of defined quality measures and the corresponding episodes of care definitions. For example, patients who have undertaken two-step therapies for controlling their blood pressure and who still have high blood pressure can, under certain circumstances, be excluded from a quality measure. However those patients will always be included in an episode of care for several reasons. First, by default, because there is no way to discern such an exclusion from claims data, and second because the quality measure is designed to measure the effectiveness of the physician’s treatment of the patient’s condition, while episodes of care cost accounting is designed to measure the efficiency with which a physician manages patients with a certain condition. For the latter exercise there is no rationale to exclude patients who have taken two therapies and can’t get their blood pressure under control. The payer still has to pay for the costs of care.

### Sustainable Feedback Loops: The Real Goal

Over the past decade or so, the use of the term “feedback loop” has increasingly entered health policy. A feedback loop from a quality measurement perspective is a way in which physicians can understand their performance, relative to a

benchmark. The underlying assumption of a quality measurement program is that the physician would change behavior to improve their own performance based on the feedback. In **Appendix B**, we outline necessary system parameters common to viable feedback loops that states should keep in mind when designing quality reporting mechanisms.

**Insofar as transparency is concerned, state purchasers should set up a performance reporting system that synthesizes cost (efficiency) and quality (effectiveness) in a way that concisely reveals value to payers, providers and consumers.** In developing a transparency approach, states should recognize that each of these stakeholders has different interests and levels of understanding. The value synthesis rests on combining efficiency calculations (total episode cost against benchmarks) and effectiveness calculations (episode-specific patient quality of care against benchmarks), and feeding back the resulting value synthesis to all providers and other stakeholders.

### Claims and Clinical Data

State purchasers can think of data drawn from claims data as Channel 1 (measuring efficiency), and non-claims, clinical data as Channel 2 (measuring effectiveness). Units of analysis for Channel 1 are patient-centered episodes of care, with an eye towards measuring variability in these episodes. Episode cost variability can come from several sources: the price of individual services, the use of services (either too many or too few), and the mix of services. Information on the contribution of each of these sources to the total variability in episode costs can help providers better understand how to improve the sum of the inputs used to manage an episode of medical care. The importance here is not simply in creating a feedback loop on a provider’s specific variability, but rather how that variability compares to that of others. For example, a provider who gets a report that shows her variability comes mostly from higher pricing of services will have a very different strategy than a provider getting a report indicating that his variability comes from a significantly higher use of certain services. As one might surmise, these reports should be payer specific, especially when analyzing variability based on price.

The units of analysis for Channel 2 (non-claims, clinical data) are all patients that have a specific medical episode, irrespective of the payer, and for two principal reasons. First, states should want to encourage providers to treat all patients with a certain condition as optimally as possible and not introduce a potential payer-specific bias. The central idea being that a single provider quality score cannot be manipulated by a payer to try and tilt that provider’s

attention preferentially towards that payer’s plan members. Second, insofar as transparency is concerned, states should set up a system that synthesizes cost and quality in a way that succinctly reveals value to payers, providers and consumers, with each having different interests and levels of understanding. **The value synthesis rests on combining efficiency calculations (total episode cost against benchmarks) and effectiveness calculations (episode-specific patient quality of care against benchmarks), and feeding back the resulting value synthesis to all providers.**

## Conclusion

While the concept of tying cost and quality of care into a timely, actionable and reliable report for physicians seems common sense enough, the general availability of data to create these reports is extremely low. As such, states that wish to accelerate the transformation of the existing delivery system into one that delivers high quality and affordable health care have to take action to develop a comprehensive

data collection and reporting mechanism.

This Brief suggests that such an approach be done using episodes of medical care – such as a chronic condition, an illness or a major treatment/procedure – as the central unit of measure because (a) quality measures are generally tied to specific episodes of medical care, and (b) acting on the cost of an episode of care is a lot easier to do for frontline clinicians than acting on a higher level of cost aggregation such as total cost of care. Of course, for states implementing bundled payment programs, the cost of the medical episode is simply the price of the bundled payment.

Further, this Brief outlines specific steps that can be taken by states to launch a data collection and reporting effort, perhaps with manual processes initially, and then to scale such an effort. The information technology infrastructure in place in the US today can be leveraged to rapidly scale a central data collection and reporting process and create highly relevant feedback loops for providers.

## Appendix A: Sample of Select Episodes of Care and Related Clinical Data Registries

Episode	Matching Qualified CDRs (QCDRs)*	Matching CDRs
Asthma acute exacerbation	American Academy of Allergy Asthma and Immunology (AAAAI) <a href="https://www.aaaai.org/home.aspx">https://www.aaaai.org/home.aspx</a>	
Bariatric surgery	Metabolic and Bariatric Surgery Accreditation and Quality Improvement Program	
Breast cancer	American Society of Breast Surgeons Mastery of Breast Surgery Program American College of Physicians (ACP) Genesis Registry <a href="https://www.medconcert.com/content/medconcert/Genesis/">https://www.medconcert.com/content/medconcert/Genesis/</a>	
Depression	ACP Genesis	
Diabetes acute exacerbation	ACP Genesis Chronic Disease Registry	
Female reproductive cancer	American Society of Clinical Oncology (ASCO) QOPI Oncology Nursing Society Quality Improvement Registry Oncology Quality Improvement Collaborative	
Lung cancer	ASCO QOPI Oncology Nursing Society Quality Improvement Registry Oncology Quality Improvement Collaborative	
Neonatal		Vermont Oxford VLBW Database <a href="https://public.vtoxford.org">https://public.vtoxford.org</a>
Perinatal		American Association of Birth Centers (AABC) Perinatal Data Registry <a href="http://www.birthcenters.org">www.birthcenters.org</a>
Spinal fusion	Anesthesia Quality Institute: National Anesthesia Clinical Outcomes Registry	
Tonsillectomy	Anesthesia Quality Institute: National Anesthesia Clinical Outcomes Registry	
Total joint replacement	American Joint Replacement Registry	

\* **Qualified CDRs:** clinical data registries authorized by CMS to collect quality measures from physicians to satisfy reporting requirements of the Medicare Physician Quality Reporting System. For a list of 2015 QCDRs see: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2015QCDRPosting.pdf>

## APPENDIX B: 7 Essential Questions That Identify Working Feedback Loops in Healthcare

1. *Where are the circuits of data and information connecting providers?* – Given the fragmented payer and provider institutional arrangements, siloed information systems, and inconsistent means of data collection, it is hard to discern the structured conduits connecting the relevant healthcare actors. The systemic “wires” must be in place.
2. *How is work (output) measured?* – This question would fall into two parts: a) the definition of work, and b) the salient contributors to work. There are so many parties, both governmental and private, creating inconsistent quality measures, the result can only be confusion and lack of uptake. Moreover, the two salient contributors to work, patients and providers, are treated as if they live on different planets. Almost all the measures place heavy emphasis on provider response, with little attention to patient response. In payment reform, this asymmetry begs for correction.
3. *What is the unit of analysis?* – Feedback systems operate on meaningful units of analysis, and thus the unit of analysis has great bearing on work measures. If the work measures are analyzed through inappropriate units, as with hospital-only measures, analysis and work output fall out of sync with each other. The appropriate unit of analysis in healthcare cannot, therefore, be institutional; it has to focus on the primary consumer of the work product: the patient.
4. *How much energy is consumed?* – Engineers are in a constant quest to lower the amount of energy required per unit of work; this is the definition of efficiency, and is often quantified in terms of wasted energy. Systems engineers would be staggered by how much energy is wasted in American healthcare, the current of work being dollars. Dollars, therefore, tie work measured and unit of analysis together as definable work products. FFS and TCOC are not defined healthcare products in dollar terms if the patient is the unit of analysis.
5. *Are the feedback mechanisms parsimonious?* – Not all metrics are equal. At some point, measuring every conceivable variance to the  $n^{\text{th}}$  degree and granting them equal weight creates more noise than signal. It turns out that most episodes of care have only a handful of meaningful metrics, that when controlled for, give the most amount of bang for the buck. This is what is meant by creating high signal to noise feedback loops. A parsimonious design gives relevant decision-makers the right amount of data points (signal) they need to optimize outcomes (work product), and weeds out extraneous information (noise).
6. *Is the feedback timely?* – This system parameter seems fairly obvious, in that outdated feedback is not only useless; it’s a nuisance. Actionable feedback must not only be parsimonious, it must be available at critical decision nodes where applying it has the most amount of potential to affect optimal Delta.
7. *Where are the control mechanisms?* – The means of making operational adjustments to bring actual performance to optimal performance (Delta) are either nowhere to be found (as with FFS), or posited in structures so large and ill defined (as with ACOs), as to conclude there are no controlling mechanisms, at least none that could qualify as actionable feedback systems. And this brings us to the heart of the matter: accountability. Since we’re not talking about feedback in machines, but rather, feedback within human networks and relationships, then accountability must be aligned with control, and that means getting the first 6 parameters right; otherwise, managerial spans of control, or “lines of sight,” become diffuse, chaotic and very difficult to coordinate.

## Endnotes

---

- <sup>1</sup> Report from the National Quality Forum: 2012 Measure Gap Analysis; [www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72981](http://www.qualityforum.org/WorkArea/linkit.aspx?LinkIdentifier=id&ItemID=72981).
- <sup>2</sup> <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/index.html?redirect=/pqri/>.
- <sup>3</sup> <https://www.cms.gov/medicare/quality-initiatives-patient-assessment-instruments/hospitalqualityinits/hospitalrhqdapu.html>.
- <sup>4</sup> See for example “Transitional Care Interventions Prevent Hospital Readmissions For Adults With Chronic Illnesses”, Kim J. Verhaegh et al. *Health Affairs* September 2014 vol. 33 no. 9 1531-1539.
- <sup>5</sup> Arkansas, Ohio and Tennessee have launched statewide bundled payment programs for specific episodes of care as the central focus of their Medicaid payment reform efforts. For additional detail on the Tennessee initiative, see: <https://www.tn.gov/hcfa/topic/episodes-of-care>.
- <sup>6</sup> Registries are databases containing specific information on patients and have been instituted by Medical Specialty Societies to help their members better monitor patient outcomes and understand the effectiveness of treatments. Some registries are also native to electronic medical records and are simply a subset of data stored in EMRs, making it easier for clinicians to extract information.
- <sup>7</sup> See HCI<sup>3</sup> report, “State Report Card on Transparency of Physician Quality Information,” December 16, 2014.” at <http://www.hci3.org/content/physician-quality-transparency-report-2014>.
- <sup>8</sup> Many clinical data registries exist and are often condition-specific. For example, the Oncology Quality Improvement Collaborative (<https://www.medconcert.com/content/medconcert/ONCQIR/>) measures and reports on outcomes in oncology and specialty care, whereas the Vermont Oxford Network hosts a database about the care and outcomes of high-risk newborn infants (<https://public.vtoxford.org/databases/very-low-birth-weight/>).
- <sup>9</sup> Support for this technical assistance work in Tennessee was provided through a grant from the Robert Wood Johnson Foundation’s State Health and Value Strategies program.
- <sup>10</sup> See <http://statenetwork.org/resource/considerations-for-state-development-of-performance-measure-sets/> accessed November 2015.
- <sup>11</sup> HCI<sup>3</sup>, through its Bridges To Excellence (BTE) programs, has been successfully collecting data elements for dozens of quality measures on common chronic conditions from various EMR systems for well over five years. For more BTE information see: [http://www.hci3.org/what\\_is\\_bte](http://www.hci3.org/what_is_bte).
- <sup>12</sup> For a list of 2015 QCDRs see: <https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2015QCDRPosting.pdf>.
- <sup>13</sup> See HCI<sup>3</sup> report, “State Report Card on Transparency of Physician Quality Information,” December 16, 2014.” at <http://www.hci3.org/content/physician-quality-transparency-report-2014>.



SEPTEMBER 2014

# Considerations for State Development of Performance Measure Sets

Prepared by Beth Waldman and Michael Bailit, Bailit Health Purchasing, LLC

## Introduction

As states play a more active role in health care delivery system and payment reform, Medicaid programs have joined other public and private sector purchasers in measuring performance as part of value-based purchasing initiatives. While essential to value-based purchasing, performance measurement can create a significant administrative burden for providers. This burden can grow significantly when individual payers (e.g., insurers, managed care plans, and third-party administrators) utilize different measures. There is a growing interest by Medicaid programs and other payers in developing common measure sets to reduce administrative burden on providers and send a common message to them about performance accountability.

This guide provides an overview of the steps states should take in developing a performance measure set—either on their own or in partnership with others—identifies critical considerations, and offers guidance in selecting measures.

## Key Initial Steps in Developing a Performance Measure Set

A number of basic questions must be answered in order to appropriately shape a discussion of what measures should be included in a measure set. It is essential to define early on whose performance is to be measured, for what purpose, and by whom. It is also important to decide who will participate in measure set development and how decisions will be made within the participant group.

### 1. Whose Performance is Being Measured?

States may choose to measure health plans and/or providers. Most current state measure set development work is focused on provider organizations, including one or more of the following: patient-centered medical homes (PCMHs), health homes, hospitals, and accountable care organizations (ACOs). In some cases, states are also developing general, procedure-specific, and condition-specific measures to support episode-based payment programs. There are also efforts to measure the performance of behavioral health and long-term services and supports providers. Measurement of ambulatory health care, however, is most common.

While this guide focuses on developing measure sets for providers, the processes described here are also applicable to health plan measure set development.

### 2. What is the Purpose of Measuring Performance?

There are a number of reasons why a state chooses to measure provider performance. Historically, states have measured provider performance as a component of a quality monitoring system, and have used performance results to inform selection of quality improvement initiatives. More recently, states have begun using performance measures to provide consumers with information about the performance of a provider and to inform discussions with contracted provider groups about their performance. State

### ABOUT BAILIT HEALTH PURCHASING

Bailit Health Purchasing is a health care consulting firm dedicated to ensuring insurer and provider performance accountability on behalf of public agencies and private purchasers.

### ABOUT STATE HEALTH AND VALUE STRATEGIES

State Health and Value Strategies, a program funded by the Robert Wood Johnson Foundation, provides technical assistance to support state efforts to enhance the value of health care by improving population health and reforming the delivery of health care services. The program is directed by Heather Howard at the Woodrow Wilson School of Public and International Affairs at Princeton University.

### ABOUT THE ROBERT WOOD JOHNSON FOUNDATION

For more than 40 years the Robert Wood Johnson Foundation has worked to improve the health and health care of all Americans. We are striving to build a national Culture of Health that will enable all Americans to live longer, healthier lives now and for generations to come. For more information visit [www.rwjf.org](http://www.rwjf.org). Follow the Foundation on Twitter at [www.rwjf.org/twitter](http://www.rwjf.org/twitter) or on Facebook at [www.rwjf.org/facebook](http://www.rwjf.org/facebook).

purchasers and their contracted health plans are also introducing new payment models that tie reimbursement to performance. Some state employer purchasers also use performance measures to tier a provider network or to identify a center of excellence.

It is not uncommon for states to use measures for more than one of the above purposes, or to use some measures for some purposes, and other measures for different purposes. For example, the state of Vermont organized a multistakeholder process to establish a performance measure set for a large ACO pilot. Some measures were selected for ACO reporting only, some for reporting and for influencing payment, and still others for measurement at the health plan level due to high baseline performance.

### 3. Is Measurement Specific to a State Program or Part of a Multipayer Initiative?

It is important to determine whether state programs will measure performance on their own or as part of a larger, multipayer initiative. For example, it is increasingly common for state Medicaid programs, state-operated insurance exchanges, and agencies charged with purchasing state employees' health coverage to use the same measure sets commercial payers use. When deciding if a single or multipayer measure set is desirable, state staff must determine the following:

- Is there a shared set of providers from whom services are being purchased?
- Are there common areas of measurement interest?
- Is there a shared purpose or intended use for the measures?

If the answer to each of these questions is “yes,” then it may make sense for a state agency to embark on a multipayer measure set initiative. Where feasible, there are advantages to both payers and providers. First, it offers a way to consistently assess performance across the entire health system within a state or geographic region. Second, depending upon the approach utilized, it can increase the measure denominator, resulting in greater ability to measure with statistical certainty. Third, it reduces the burden on providers of supplying data and attempting to improve across a large number of measures.\* Fourth, it gives providers a clear message on what aspects of care are most important to purchasers and payers, and encourages them to focus on those areas.

### 4. How Often Will Measurement Occur?

As part of the initial planning process, states and/or multipayer initiatives should consider whether measurement will be one-time or ongoing, and if ongoing, how often. In most cases, measurement occurs on an annual basis as many quality measures use 12-month measurement periods. There may be a desire to measure more frequently to track progress toward an established

goal (something the Oregon Health Authority does two to three times a year), or for certain types of measures, such as utilization (something which the Vermont Green Mountain Care Board does when it tracks ACO member service utilization on a year-to-date basis).

### 5. Who Participates in the Process and How Are Decisions Made?

When developing a measurement set internal to a state agency, it is important to include the right staff from across the organization to ensure appropriate consideration is given to the entirety of the agency's measurement goals, and that the appropriate decision-makers are in the room. There will be some difficult decisions about how to prioritize measures and whether the agency has sufficient resources to implement a particular measure or set of measures. At a minimum, an agency's quality, informatics, medical management, and finance departments should be represented, and there should be a clear decision prior to the start of the project as to who will own the project and serve as the ultimate decision-maker.

In addition, the participation of external stakeholders, such as affected providers, health plans, and consumer advocates can not only increase the likelihood of obtaining buy-in from key constituents, but also contribute to a better-reasoned and effective measure set.

If state agencies are participating in a multistakeholder effort to develop a measure set, it is important to have the right staff from all participating organizations actively engaged. Participants must be able to make decisions and commit their organization to an approach. Individuals who are neither technically informed (e.g., an insurer's regional sales manager) nor empowered will be unable to contribute to the process or ensure that the resulting measure set will be adopted by their organizations.

Multistakeholder initiatives must clearly delineate up front how decisions will be made within the group and how measures will be prioritized when there are differing goals or disagreement on how to move forward. At the start of the process, participants should lay out how decisions will be made and how disagreements will be addressed.

### Measure Selection

The first step in selecting measures is to set out selection criteria that allow for a consistent review of potential measures that is informed by the overall goals and desired outcomes for the measurement program.

Selection criteria typically address:

1. clinical and technical merits of the measure;
2. the relation of the measure to goals and improvement opportunities;

\* Cambridge Health Alliance (MA) reported having 546 payer-defined measures. (Somava Stout, personal communication, May 14, 2014).

3. operational considerations for generating the measure; and
4. the relation of the measure to other pre-existing measure sets of interest.

### Selection Criteria

There are a number of important questions to consider when selecting measures. States should leverage the *Buying Value Measure Selection Tool*, which provides both technical and programmatic criteria for each measure, and a set of criteria for the overall measure set. Further, the tool also provides examples of measure set criteria and can help states track whether the measures under consideration meet measure selection criteria adopted by the state. Such criteria can and should be applied both to individual measures and the entire set, the latter to ensure that the entirety of the measure set is balanced and complete. Examples of criteria commonly adopted include whether measures:

- are collectively consistent with the overall goals of those involved in measure set development;
- are valid and reliable;
- represent opportunities for performance improvement;
- measure the provider's performance in an area within the targeted providers' control;
- have been endorsed by a national body, such as the National Quality Forum (NQF) or the National Committee for Quality Assurance (NCQA);
- have sufficient denominators to produce reliable measurement, be they intended for assessment of statewide, multi-provider, practice site or individual practitioner performance;
- have relevant benchmarks;
- are focused on outcomes;
- are feasible to implement, and are not overly burdensome to generate, report, and if applicable, aggregate;
- are aligned with existing state measure sets and initiatives;
- are aligned with measures currently in use by health plans; and
- are aligned with national and federal measurement initiatives.

One potential criterion is the size of the set. It is often difficult to set a limit on size before knowing the types of measures to be adopted and their intended use. For example, a measure set that includes both physician and hospital measures, as well as access, quality, patient experience, and efficiency measures,

should be expected to be larger than one including only physician ambulatory care quality measures. Should the state desire to adopt a measure set size criterion, however, the number should not be set in stone, but should be used to help filter and prioritize potential measures.

### Use of New and Innovative Measures

As states look to develop measure sets, they often begin with a desire to look at outcome measures rather than process measures, and to focus on areas that may currently be under-measured, such as care integration, social determinants of health, and social supports. Such measures can pose implementation challenges. This is not to say that a state should not strive to innovate, or adopt “transformational measures,” but in so doing the state should ensure that implementation is feasible, recognizing that it will require significant time and resources to develop and/or implement such measures. The state may want to consider staging the implementation of innovative measures, piloting and testing them before using them for transparency or payment purposes.

### Designating Measures for Specific Uses and Specific Populations

As indicated above, measures may be selected for one or more uses. The Maine Health Management Coalition organized a multistakeholder measure selection process on behalf of the state with the specific purpose that the measures would be employed in both the state's and commercial insurers' contracts with ACOs. Other states, however, have designated different measures for distinct purposes, including performance monitoring, value-based payment, public reporting, and measure testing.

In addition, measures may be selected for use across populations or for a specific population. For example, Medicaid and commercial payers may agree that common measures of diabetes care are a priority for both of their populations. They may differ in opinion, however, when considering measures specific to persons with serious and persistent mental illness due to the greater prevalence of the condition in the Medicaid population. In such circumstances, the parties may agree to adopt a measure set that is common to commercial and Medicaid populations, but also allows for a limited number of Medicaid-only measures.

This measure designation process can occur during measure set development, or following initial development of the measure set.

### Identifying Populations, Performance Domains and Services for Measurement

To develop a comprehensive measure set, the state should include measures that comprehensively address patient populations, performance domains, and services. Table 1 provides a description of potential populations, domains, and clinical service areas. Not all of the categories are mutually exclusive.



**Table 1: Potential Measurement Categories**

Populations	Performance Domains	Service Areas
<ul style="list-style-type: none"> <li>▪ children</li> <li>▪ adolescents</li> <li>▪ non-disabled adults</li> <li>▪ adults with disabilities</li> <li>▪ pregnant women</li> <li>▪ seniors</li> </ul>	<ul style="list-style-type: none"> <li>▪ provider infrastructure</li> <li>▪ access</li> <li>▪ clinical process</li> <li>▪ clinical outcomes</li> <li>▪ health status</li> <li>▪ function</li> <li>▪ consumer experience</li> <li>▪ patient engagement</li> <li>▪ patient safety</li> <li>▪ cost</li> <li>▪ efficiency</li> <li>▪ utilization</li> <li>▪ overuse and misuse</li> </ul>	<ul style="list-style-type: none"> <li>▪ prevention</li> <li>▪ acute care</li> <li>▪ chronic illness care</li> <li>▪ dental care</li> <li>▪ behavioral health care</li> <li>▪ inpatient care</li> <li>▪ ambulatory care</li> <li>▪ long-term services and supports</li> <li>▪ care management</li> </ul>

States sometimes identify sub-populations, performance domains, and service areas of special interest to them. It is quite common for states to identify specific diseases that are prevalent within a population or program being measured. Where diabetes and asthma are common across populations, Medicaid programs might want to target care for behavioral health conditions, such as serious and persistent mental illness and substance use disorders. The specific conditions and/or procedures to be measured depend on the goals of the measurement program, the participants in the measure selection process, and the criteria that they adopt at the outset of their work.

**Resources for Locating Measures**

There are many sources that may be used to identify potential measures. In addition to the 700 NQF-endorsed measures, measure set developers should consider the following resources:

- Federal measure sets (partial list)
  - Consumer Assessment of Healthcare Providers and Systems (CAHPS®) surveys
  - Children’s Health Insurance Program Reauthorization Act (CHIPRA) core set
  - Center for Medicare and Medicaid Innovation (CMMI) core measures
  - Hospital Compare
  - Meaningful Use Clinical Quality measures\*
  - Medicare Advantage Stars Program measures

- Medicare Shared Savings Program measures
- Medicare-Medicaid Financial Alignment Model measures
- Medicaid adult core set
- Nursing Home Compare
- Pre-existing state measure sets (partial list, not applicable to all states)
  - Measure sets currently in place in state health plan and third-party administrators contracts
  - Measures sets currently in place in state ACO, PCMH, and health home contracts
  - Measure sets defined through state-facilitated processes for multipayer and provider use. For example, Massachusetts’ [Standard Quality Measure Set](#) and California’s [CalQualityCare.org](#).
- Pre-existing multistakeholder coalition measure sets, such as those developed by the Wisconsin Collaborative for Health Care Quality (WCHQ), Better Health *Greater* Cleveland, Minnesota Community Measurement, and the New Mexico Coalition for Healthcare Quality.
- Agency for Healthcare Research and Quality’s prevention quality indicators
- NCQA’s Healthcare Effectiveness Data and Information Set (HEDIS)
- Long-term services and supports scorecard

\* These measures are a subset of the larger Physicians Quality Reporting System and Physician Value Based Payment Modifier Program measure set.

Web links to several of the measure sets cited above are in the Buying Value Measure Selection Tool. The tool also includes a list of the most frequently used measures by domain. In addition, the tool provides a scoring template states can use to organize the measures in use or under consideration and apply their selection criteria. Through an automated crosswalk, the template lets states identify whether a measure is included within a federal measure set.

### Measure Set Fidelity

Most current measure set activity involving multiple payers is aimed at achieving true alignment, where each payer agrees to adopt the common measure set in full, with the exception of a few population-specific measures.

An alternative approach, however, involves the adoption of a common measure set from which each participating payer (or payer and provider) chooses which measures to use. While this “menu” approach reduces variation across payer measure sets, it leaves the door open to non-alignment.

## Producing the Measure Set

The process of collecting data and producing measures can be resource-intensive. It is important to understand the data that are needed to produce a particular measure, and to consider the ability of the state and/or its health plans to access, collect, and analyze such data prior to selecting a measure for use.

### Data Sources

A variety of data sources can be used to generate measures. For the most part, measures that use claims or encounter data are the easiest to produce, because they are readily available to the state and/or its health plans. Measures that require a consumer survey are also relatively easy to produce, particularly if the survey process is already in place.

More difficult to produce are measures that require a review of clinical records. If performed manually, reviewing clinical records is time-consuming and expensive for providers and states and/or health plans. If performed using electronic data sources, there are additional challenges, including:

- limitation in the numbers of providers able to capture and report the designated measures;
- inconsistent reporting across electronic health records (EHRs), creating problems in the reliability of reported data; and
- the inability of many health information exchanges to facilitate electronic measure reporting.

Despite the current difficulties associated with generating measures using clinical data sources, there is little question that current trends toward expanded EHR adoption and health information exchange development will result in increasing use

of clinical data-based measures over time. States should anticipate this trend and make provision for testing or including some clinical data-based measures in their measure set.

### Identifying Benchmarks

In addition to identifying data sources for measure generation, it is also important to identify benchmarks to which a provider’s performance will be compared. This is particularly true if the state anticipates using the measure set for quality improvement, public reporting, or adjusting payment. In all three applications, it is often necessary to assess performance relative to a benchmark to identify opportunities for improvement.

Unfortunately, there are limitations in the number of measures for which national benchmarks are available. Many states select NCQA’s HEDIS measures for their measure sets, because NCQA annually publishes Medicaid, Medicare, and commercially-insured population benchmarks for most of the HEDIS measures.<sup>1</sup> Yet, use of the HEDIS health plan measure benchmarks for provider performance can be troublesome. As reported by WCHQ at the *Buying Value* meeting in March 2014, differences in specifications necessary to make a health plan measure applicable to a provider entity can significantly impact the comparability of the two rates.

Other sources for national benchmarks exist, but these too have their limitations, as noted below:

- Health Resources Service Administration (HRSA): HRSA collects and reports on a number of clinical data-based measures. The rates are reported from the EHRs operated by federally-qualified health centers (FQHCs) and reflect FQHC performance only.
- Centers for Disease Control and Prevention (CDC): The CDC publishes the results of the Behavioral Risk Factor Surveillance System, the world’s largest, ongoing telephone health survey system. While research has shown the reliability of patient-reported measures to be good, states cannot be certain of the comparability of each measure relative to measures generated from other data sources.
- Medicare Hospital Compare: The CDC publishes benchmarks for hospitals using Medicare performance data, as well as for nursing homes (Nursing Home Compare).

States and state and regional quality improvement organizations have often created their own internal state benchmarks; these can also be a resource.

## Reviewing and Modifying the Measure Set

It will be important to develop a process for both ad hoc and regular periodic review of current measures to determine whether they should be retained or modified, or if new measures should be included based on changing circumstances or priorities.

Ad hoc measure review is necessary because changes in national clinical guidelines have direct impact on commonly used, nationally endorsed measures. For example, the new American College of Cardiology/American Heart Association guidelines issued in late 2013 on cholesterol management had significant impact on the LDL-C control measure employed in many measure sets.<sup>2</sup> As a result, many state and multi-payer/multistakeholder organization measure sets had to be modified based on the new guidelines.

Periodic measure set review should occur well in advance of the implementation of any measure set changes so that affected provider organizations will have adequate time to react. For example, the Oregon Health Authority created a calendar of planned measure review activities to inform affected provider organizations 60 days prior to their effective date. As with initial measure set development, a set of explicit criteria should be used to inform decision-making.

## Pitfalls in Performance Measurement

While there are important opportunities in performance measurement, it is also important to be mindful of the potential pitfalls. While performance measurement can serve to align goals and incentives, it has the potential to narrowly focus providers and health plans on aspects of care that are being measured, and especially so when the measure is tied a reward or penalty. This narrow focus could lead to unintended consequences, such as paying too little attention to other important health care components that are not being measured. One way to reduce this potential pitfall is to include both monitoring and incentive measures within a performance measurement set. Monitoring measures can be promoted to incentive measure status if performance slides.

As mentioned previously, the development of homegrown measures can be problematic for a number of reasons, including validity, reliability, and the inability to access a performance benchmark. As states try to measure social determinants of health as part of measurement initiatives, it is important to consider whether it is appropriate to hold health care providers accountable for things over which they have little or no control, such as education, environment, and poverty.

## Conclusion

In developing a performance measurement initiative, the state should consider how measurement can evolve over time. While there may be short-term limitations to the depth and breadth of measures that can be implemented, the consideration of a broader array of measures gives states a pathway for expanding their measurement set and increasing their options for incentives.

In addition to developing a measure set as part of a multipayer initiative—the state and its payer partners if in a multipayer initiative—should engage the participating providers to help them achieve success on these measures. While quality-based incentives offer providers extrinsic motivation to improve the quality of care and the health status of Medicaid beneficiaries, they are not sufficient. Providers must not only want to change, they must also know what and how to change in order to improve care. States and other payers will need to continue their efforts to actively manage health plans and providers, including setting strategic direction and providing ongoing performance review and support for quality improvement activities. They must also consider how to provide technical and data support to providers to ensure that measurement and other activity yield desired results.

## Endnotes

1. Bazinsky, Kate, and Michael Bailit. 2013. “The Significant Lack of Alignment Across State and Regional Health Measure Sets.” *Buying Value*. [www.buyingvalue.org/wp-content/uploads/2014/02/buying-value-common-measures-Bailit-State-Measure-Set-Brief-9-10-13-FINAL-FINAL.docx](http://www.buyingvalue.org/wp-content/uploads/2014/02/buying-value-common-measures-Bailit-State-Measure-Set-Brief-9-10-13-FINAL-FINAL.docx).
2. Stone, Neil, et al. 2013. “2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidance.” American Heart Association. doi: 10.1161/01.cir.0000437738.63853.7a.