

U.S. Department of Health and Human Services

REPORT TO CONGRESS:
Plan to Implement a Medicare Hospital
Value-Based Purchasing Program

November 21, 2007



Acknowledgements

Prepared by the CMS Hospital Value-Based Purchasing Workgroup
with Assistance from the RAND Corporation, Brandeis University,
Booz | Allen | Hamilton, and Boston University

This work was performed through a collaborative contracting effort with the Office of the
Assistant Secretary for Planning and Evaluation (ASPE), Department of Health and
Human Services.

Table of Contents

Acknowledgements.....	ii
Table of Contents.....	iii
Table of Figures.....	v
Table of Tables.....	vi
Executive Summary.....	1
The Medicare Hospital Value-Based Purchasing (VBP) Plan.....	2
Overview.....	2
The Performance Assessment Model.....	3
Translating the VBP Total Performance Score into Incentive Payment.....	5
VBP Measures.....	9
Transitioning from RHQDAPU to VBP.....	12
Redesign of Data Infrastructure and the Data Validation Process for the VBP Program.....	15
Public Reporting.....	20
VBP Program Monitoring and Evaluation.....	20
Appendices	
Appendix A: Background.....	22
The Context for the Hospital VBP Plan.....	22
Setting the Stage for Value-Based Purchasing.....	23
Process Used to Develop the Value-Based Purchasing Plan.....	25
Stakeholder Perspectives.....	26
Summary of Findings from the Hospital Pay-for-Performance Environmental Scan and Literature Review.....	28
Appendix A-1: List of Organizations Participating in the Environmental Scan.....	36
Appendix B: Details of the Performance Assessment Model.....	39
Scoring Clinical Process-of-Care Measures.....	39
Approach to Scoring HCAHPS Survey of Patients' Perspectives of Care.....	47
Creating the VBP Total Performance Score.....	56
Approach to Scoring 30-Day Mortality Measures.....	58
Appendix B-1: Example of Scoring Hospital Performance on Clinical Measures.....	59
Appendix B-2: CAHPS [®] Hospital Survey (HCAHPS) Fact Sheet.....	60
Appendix B-3: Example of Scoring Hospital Performance on HCAHPS Measures... ..	62
Appendix B-4: Example of Combined Clinical and HCAHPS Performance Assessment Score.....	64
Appendix C: Examples of Modifying DRG Payment with VBP Incentive Payment.....	65
Appendix D: Measure Development, Selection, and Modification.....	67
The Measure Development and Selection Process.....	67
VBP Measures at Program Start.....	69
Additional Measures Beyond Year 1.....	72
Small Numbers on Individual Performance Measures.....	74
Appendix D-1: 2005 Scores on Clinical Process-of-Care Measures.....	76
Appendix E: Enhancements to Data Infrastructure and Measure Validation.....	77

Enhanced User Support.....	77
Proposed Measure Validation Process.....	78
Exclusion of Data on Hospital Compare for Hospitals Failing Validation	81
Appendix E-1: Example to Illustrate Differences Between Data Element Match and Measure Match Rate Approaches	83
Appendix F: VBP Program Monitoring, Evaluation, and Support for Quality Improvement.....	85
VBP Program Monitoring.....	85
Monitoring the Impact of Value-Based Purchasing on Disparities	87
Design Elements of the VBP Program that Seek to Reduce Disparities	90
Quality Improvement Support for Hospitals.....	90
Appendix G: Glossary of Terms and Acronyms.....	93
Appendix H: CMS Hospital VBP Workgroup Membership	95
Appendix I: VBP Plan Support Contractors	97

Table of Figures

Figure 1: Nonlinear Exchange Function for Translating Total Performance Score into Percent of VBP Incentive Payment Earned	7
Figure 2: Proposed Process for Introducing Newly Developed Measures into VBP	10
Figure 3: Timeline Linking Baseline, Measurement and Performance Payment Periods	13
Figure 4: Detailed Timeline Linking Measurement and Performance Payment Periods .	15
Figure 5: Validation Process and Validation Appeals Timeline: Random and Targeted Audits	18
Figure B-1: Example of Hospital Earning Points by Exceeding Benchmark, Clinical Process-of-Care Measure Scoring Model	44
Figure B-2: Example of Hospital Earning Points by Attainment or Improvement, Clinical Process-of-Care Measure Scoring Model	45
Figure B-3: Example of Hospital Earning No Points, Clinical Process-of-Care Measure Scoring Model.....	46
Figure B-4: Example of Hospital Earning Points by Exceeding Benchmark, HCAHPS Measure Scoring Model.....	52
Figure B-5: Example of Hospital Earning Points by Attainment or Improvement, HCAHPS Measure Scoring Model	53
Figure B-6: Example of Hospital Earning Zero Points, HCAHPS Measure Scoring Model	54
Figure B-7: Examples of Hospitals Earning Points Based on Minimum Performance Across All Eight HCAHPS Dimensions.....	55
Figure D-1: Universe of VBP Measures.....	67
Figure E-1: Details of Validation and Appeals Timeline	80
Figure E-2: Timeline for Reporting Data on Hospital Compare	82

Table of Tables

Table 1: Potential Measures for the Financial Incentive at VBP Program Start	11
Table A-1: Summary of Empirical Studies Examining Hospital P4P Programs.....	30
Table B-1: Benchmarks and Attainment Thresholds for Clinical Process-of-Care Measures	40
Table B-2: HCAHPS Survey Performance Measures	48
Table C-1: Example for DRG 498	65
Table D-1: Evolution of VBP Measures over Time	73
Table E-1: RHQDAPU Validation: Element Level Example – Heart Failure	84
Table E-2: VBP Validation – Measure Level for Heart Failure	84

Executive Summary

The Congress, through the Deficit Reduction Act of 2005, Section 5001(b), authorized the Secretary of Health and Human Services to develop a plan to implement value-based purchasing (VBP) commencing Fiscal Year (FY) 2009 for Medicare hospital services provided by subsection (d) hospitals paid under the Inpatient Prospective Payment System (IPPS). By statute, the plan must include consideration of: (1) the development and selection of measures of quality and efficiency in inpatient settings; (2) reporting, collection, and validation of quality data; (3) the structure, size, and source of value-based payment adjustments; and (4) disclosure of information on hospital performance.

This report discusses options for a plan to implement a Medicare Hospital VBP program, which builds on Medicare's current Reporting Hospital Quality Data for Annual Payment Update (RHQDAPU) Program, which since FY 2005 has provided differential payments to hospitals that meet certain requirements, including publicly reporting their performance on a defined set of inpatient care performance measures. Building on the foundation of RHQDAPU, CMS recommends replacing the current quality reporting program with a new program that would include both public reporting and financial incentives for better performance as tools to drive improvements in clinical quality, patient-centeredness, and efficiency. A Medicare Hospital VBP Program should be implemented in a manner that does not increase Medicare spending.

This report contains the following key components: (1) a potential Performance Assessment Model that incorporates measures from different quality "domains" (clinical process of care, patient perspective of care, outcomes, etc.) to calculate a hospital's Total Performance Score; (2) options to translate that score into an incentive payment that makes a portion of the base DRG payment contingent on performance; (3) options for criteria to select performance measures for the financial incentive and candidate measures for FY 2009 and beyond; (4) a potential phased approach to transitioning from RHQDAPU to VBP; (5) a redesign of current data transmission and validation infrastructure to support VBP Program requirements; (6) potential enhancements to the Hospital Compare website to support expanded public reporting; and (7) an approach to monitoring VBP impacts.

The Medicare Hospital Value-Based Purchasing (VBP) Plan

Section 5001(b) of the Deficit Reduction Act (DRA) of 2005 (P.L.109-171) authorizes the Secretary of Health and Human Services to develop a plan to implement a value-based purchasing program for payments under the Medicare program to subsection (d) hospitals beginning with Fiscal Year 2009. By statute, the plan must include consideration of at least the following design issues:

1. The on-going development, selection, and modification process for measures of quality and efficiency in hospital inpatient settings.
2. The reporting, collection, and validation of quality data.
3. The structure of value-based payment adjustments, including the determination of thresholds or improvements in quality that would substantiate a payment adjustment, the size of such payments, and the source of funding for the value-based payments.
4. The disclosure of information on hospital performance.

This Report to Congress discusses options for a plan to Implement Medicare Hospital Value-Based Purchasing (VBP) and accompanying Appendices. The Appendices provide background on option development and details on various features of potential plan methodology.

Overview

Value-based purchasing (VBP), which links payment to performance, is a key policy mechanism that CMS is proposing to transform Medicare from a passive payer of claims to an active purchaser of care. The options in this report build on the foundation of the current Reporting Hospital Quality Data for Annual Payment Update (RHQDAPU) Program. Developed in response to Section 501(b) of the Medicare Modernization Act of 2003 and DRA Section 5001(a), RHQDAPU ties a portion of the Annual Payment Update under the Inpatient Prospective Payment System (IPPS) to a hospital's meeting certain requirements, including reporting on a defined set of inpatient quality measures. A VBP program would phase out RHQDAPU and would make a portion of hospital payment contingent on actual performance on

specified measures, rather than simply on a hospital's reporting data for these measures. Under VBP, payments to high-performing hospitals would be larger than those to lower performing hospitals, for the first time using the IPPS to provide financial incentives to drive improvements in clinical quality, patient-centeredness, and efficiency. Public reporting of performance on Medicare's Hospital Compare website, a key component of RHQDAPU, would remain an essential component of VBP.

The VBP Plan could include the following basic components and enhancements of current data submission and public reporting:

1. A Performance Assessment Model that is used to score a hospital's performance on a specified set of measures, generating a Total Performance Score for each hospital.
2. Translation of the VBP Total Performance Score into an incentive payment.
3. A measure development process, including selection criteria for choosing performance measures for the financial incentive, and candidate measures for VBP Program start.
4. A phased approach to transition from RHQDAPU to VBP.
5. Redesigned data submission and validation infrastructure to support the VBP Program requirements.
6. Enhancements to the Hospital Compare website to support expanded public reporting of performance results.
7. An approach to monitoring VBP impacts, including potential impacts on health disparities.

This report presents options and considerations for the essential characteristics of each design component. Further details are included in Appendices to this Plan.

The Performance Assessment Model

The Performance Assessment Model is the methodology that could be used for scoring hospital performance and computing each hospital's VBP Total Performance Score that then would be translated into a specific level of incentive payment. Each hospital's performance would be assessed annually using the methodology. The potential model discussed below combines scores on individual measures across different performance domains (clinical process-of-care, patient perspectives of care, and 30-day mortality outcomes at VBP Program start) to compute a

hospital's VBP Total Performance Score, which is then used to determine the percentage of the VBP incentive payment earned by the hospital.

The potential model evaluates a hospital's performance on each measure based on the higher of an "attainment score" in the measurement period or an "improvement score" determined by comparing the hospital's current measure score with its prior-period baseline performance. This approach would encourage a broad range of hospitals to engage in value-based purchasing—even those that begin with a low absolute level of performance.

Overview of a Potential Performance Assessment Model

- A hospital must submit data for all VBP measures that apply to its patient population and service mix. The measures could be used for incentive payment, public reporting, or measure development.
- A hospital receives a performance score on each measure for incentive payment for which it has a minimum number of cases.
- Measures are grouped into "domains"—*e.g.* clinical process-of-care measures, HCAHPS patient perspectives of care survey, efficiency measures—and a score is calculated for each domain by combining the measure scores within that domain, weighting each measure equally. The score reflects the percentage of points earned out of the total possible points for which a hospital is eligible.
- A hospital's VBP Total Performance Score is determined by aggregating the scores across all domains. Domains could be weighted equally or unequally.
- The Total Performance Score is translated into the percentage of VBP incentive payment earned using an "exchange function," which aligns payments with desired policy goals.

Other models considered would have selected one of these dimensions, but not the other, for incentive payment under VBP. To base VBP incentives exclusively on attainment would emphasize the goal of rewarding relatively good current performance. Alternatively, VBP could focus incentive payments on improvement, thereby shifting more resources to under-performing hospitals. However, these models could engage a smaller range of hospitals because they convey opportunities only for either high or low performers. Rewarding only the best performance leaves hospitals needing the greatest improvement with little opportunity to earn incentives, which could cause low-performing hospitals to stagnate or fall even further behind in relative performance. Alternatively, rewarding only improvement would provide little or no recognition for hospitals that have already attained exemplary performance. The potential model blends these objectives with the intention of mobilizing all hospitals to improve quality.

For each measure a *benchmark* and an *attainment threshold* would need to be determined annually. There are several options for determining the thresholds, as considered in Appendix B. One option is the distribution of national hospital performance on that measure during the previous 12-month baseline reporting period. Because these scoring “cut points” would be determined from actual hospital performance nationally, they would provide realistic markers of performance expectations. These parameters, along with the *attainment range* and *improvement range* they define, would be used to determine a hospital’s score on each measure. Appendix B explains a methodology that could be used to determine these parameters and detailed descriptions for scoring measures, for determining a hospital’s performance score for each domain, weighting each measure equally within a domain; and then for combining the domain scores, to determine a hospital’s score.

**Performance Assessment Model
Terminology**

For each measure:

Benchmark: reference point used to define a high level of performance

Attainment threshold: minimum level of performance required to receive attainment points

Attainment range: scale between the attainment threshold and the benchmark

Improvement range: scale between the hospital’s prior-year score (baseline) on the measure and the benchmark

Translating the VBP Total Performance Score into Incentive Payment

Translating a hospital’s score into an incentive payment requires the following three elements:

- Specification of an exchange function used to translate the VBP Total Performance Score into the percent of the VBP incentive payment earned and of the benchmark performance level for a hospital to obtain its full incentive amount.
- Identification of the funding source for the incentive payments
- Determination of how to allocate the pool of funds that would be created because not all hospitals would earn the full incentive payment.

The Exchange Rate and Benchmark Performance Level

Figure 1 presents an example of a nonlinear exchange function for translating each hospital’s Total Performance Scores into the percent of VBP incentive payment earned. The horizontal axis represents hospital VBP Total Performance Scores ranging from 0 to 100 percent. The vertical axis represents the percentage of the incentive payment earned conditional on quality

performance. The benchmark level of performance for a hospital to receive its full incentive is set at 85 percent in this example.

By modifying the benchmark level of performance and the slope of the exchange function, policymakers could determine how difficult it would be for a hospital to qualify for a particular level of incentive payment. As shown in Figure 1, steeper slopes represent proportionally greater returns to the performance score, and flatter slopes represent proportionally lower returns. The steep curve for lower VBP performance scores assumes higher initial fixed costs associated with launching a significant quality improvement program within a hospital, followed possibly by lower incremental costs associated with ongoing quality improvement efforts. In this example, setting benchmark level at 85 percent is estimated to ensure that at least 20 percent of hospitals would earn back their full VBP incentive in any given year. The table below the graph shows the percentage of the incentive payment earned by the hospitals in the scoring examples presented in Appendix B-4.

Figure 1: Nonlinear Exchange Function for Translating Total Performance Score into Percent of VBP Incentive Payment Earned

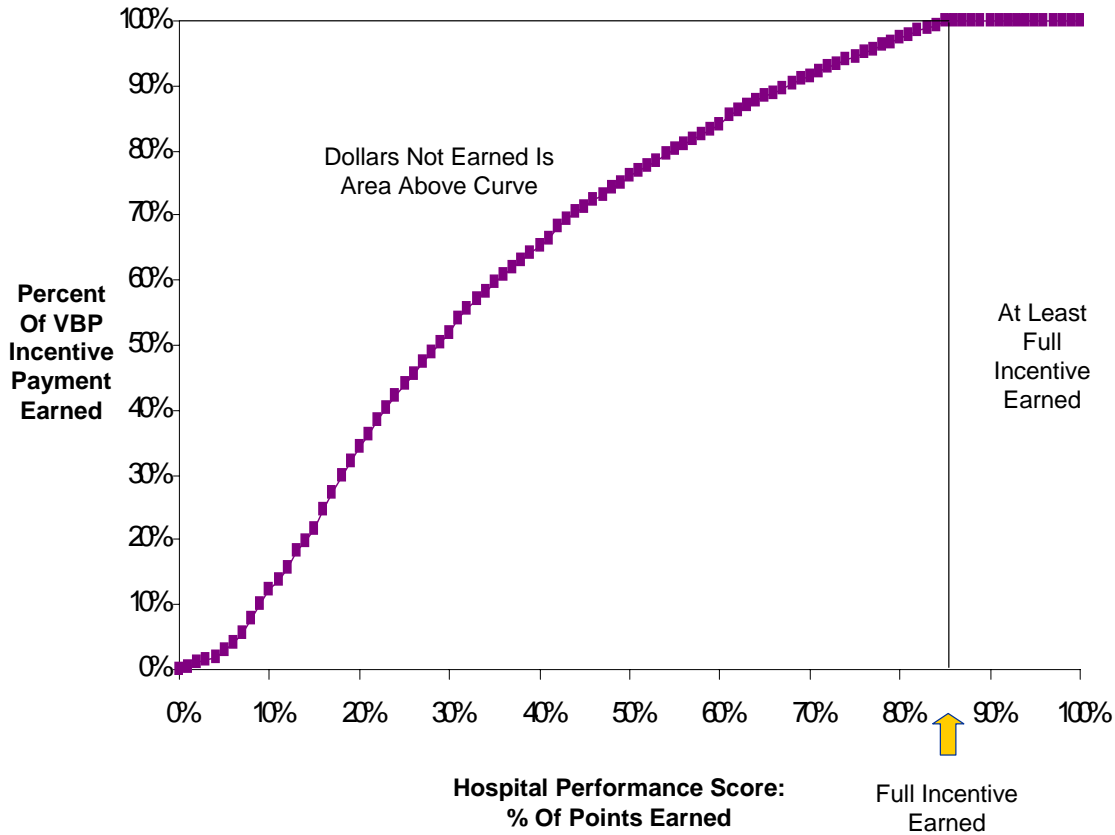


Illustration of Nonlinear Exchange Function

Hospital	Total Performance Score (% of total points achieved)	Incentive Payment (% of payment earned)
Hospital B	94	100
Hospital A	57	80
Hospital I	62	83
Hospital L	11	14

In this example, any hospital receiving a Total Performance Score above 0 would receive incentive dollars.

Basis of Incentive Payments

There are a range of options for the components of the IPPS payment that could provide an appropriate basis for the incentive payment. The incentive could be a percentage of the base operating DRG payment only (limited to the geographic and DRG relative weight adjustments).

This approach would link the incentive payment most directly to the clinical services provided during a patient stay. Alternatively, the incentive could be based on additional components of the IPPS payment, including:

- Capital costs
- Disproportionate Share Hospital (DSH) payments
- Indirect Medical Education (IME) payments
- Outlier payments for unusually costs cases

This approach could include components of IPPS payments that may be less directly related to the policy objectives of the VBP Program. Regardless, the percentage of the base that is allocated to the VBP incentive payment could be established annually. Hence, no additional funding would be required to provide a source for the VBP incentive.

The size of the incentive payment and the base to which it applies could influence hospitals' decisions on whether to undertake or continue investments to earn incentive payments. No definitive body of research exists that indicates the optimal payment policy parameters for achieving the goals of the VBP program. A range of 2 – 5 percent could be considered, and CMS expects that experience under the VBP program will guide modifications to the payment parameters to meet the objectives of the VBP program over time.

The VBP incentive could apply to all Medicare discharges or only to those for which performance measures are included in the incentive measure set. Applying the incentive to all DRG's, not just to those associated with the clinical process-of-care measures, could encourage hospitals to address delivery system problems that transcend these specific conditions. Medicare wants to encourage this system improvement, rather than having hospitals focus narrowly only on the conditions being measured. Furthermore, the HCAHPS patient experience-of-care survey applies hospital-wide and is not DRG-specific. Adding other measures that transcend specific conditions would be a high priority as the VBP program evolves.

An illustration of applying the VBP incentive payment to modify the base DRG is presented in Appendix C. In this example, DRG payments would include geographic and DRG relative

weight adjustments, but payments for capital costs, IME (indirect medical education), and DSH (disproportionate share hospital) would not be part of the basis of the incentive payment.

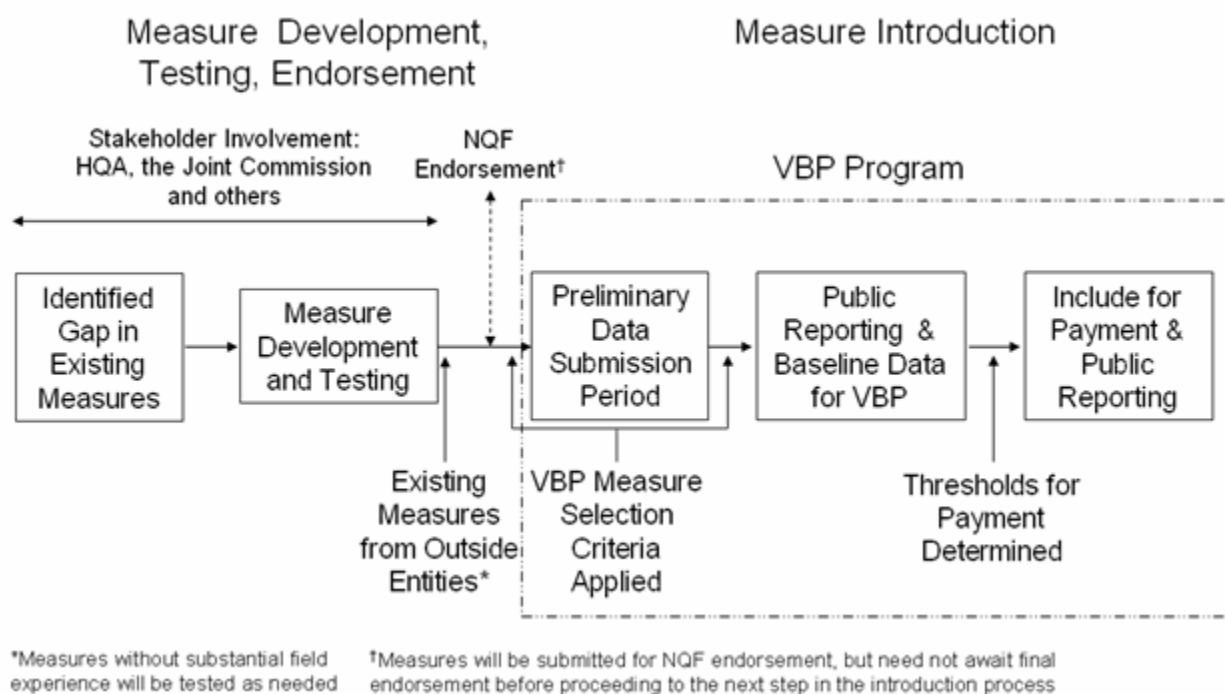
Allocation of Unearned VBP Incentive Dollars

As the exchange rate curve illustrates, many hospitals would not earn the full VBP incentive payment, creating a pool of unallocated funds. The unallocated funds could be distributed to hospitals in whole or part as an additional quality incentive. As discussed later in the report, unearned incentives in the first year would be expected to be zero or negligible, if the first year is based only on reporting quality measures.

VBP Measures

The VBP Program could use measures for a number of purposes, including: public reporting, financial incentive, and development of new measures to support ongoing expansion of the measure set. As depicted in Figure 2, a staged approach to measure introduction could be used to appropriately test and have hospitals submit data for new measures before these measures would be included in the set used for public reporting and then for the financial incentive. CMS would apply screening criteria against candidate measures to determine their suitability for inclusion in the VBP Program and specifically for the financial incentive. To qualify for the financial incentive, a hospital would have to report on all measures relevant to its service mix, including new measures in the testing stage for possible introduction into the VBP Program, measures being publicly reported but not included in the measure set for the financial incentive, and those measures used for determining the financial incentive. Appendix D describes in more detail the stages of this process and the potential approach to measure selection.

Figure 2: Proposed Process for Introducing Newly Developed Measures into VBP



Examples of measures that could be used for the VBP program include a subset of the RHQDAPU clinical process-of-care measures, the HCAHPS patient perspectives-of-care survey, and the initial two clinical outcome measures. Data on the RHQDAPU clinical process-of-care measured used in this report are from 2004 and 2005. Baseline and follow-up data for the other measures are not yet available. Hospitals will be required to report baseline HCAHPS patient perspectives-of-care survey beginning in 2008, and data on clinical outcome measures were first reported on Hospital Compare in June 2007 (Table 1).

Table 1: Potential Measures for the Financial Incentive at VBP Program Start

Clinical Quality – Process-of-Care Measures		Initial Hospital Compare Inclusion
<i>Acute Myocardial Infarction (AMI)</i>		
AMI-1	Aspirin at arrival*	4/2005
AMI-2	Aspirin prescribed at discharge*	4/2005
AMI-3	ACE inhibitor (ACE-I) or Angiotensin receptor blocker (ARBs) for left ventricular systolic dysfunction*	4/2005
AMI-4	Adult smoking cessation advice/counseling*	4/2005
AMI-5	Beta blocker prescribed at discharge*	4/2005
AMI-7a	Fibrinolytic (thrombolytic) agent received within 30 minutes of hospital arrival	4/2005
AMI-8a	Primary percutaneous coronary intervention (PCI) received within 120 minutes of hospital arrival	4/2005
<i>Heart Failure (HF)</i>		
HF-1	Discharge instructions	4/2005
HF-3	ACE inhibitor (ACE-I) or Angiotensin receptor blocker (ARBs) for left ventricular systolic dysfunction	4/2005
HF-4	Adult smoking cessation advice/counseling*	4/2005
<i>Pneumonia (PN)</i>		
PN-2	Pneumococcal vaccination status	4/2005
PN-3b	Blood culture performed before first antibiotic received in hospital	4/2005
PN-4	Adult smoking cessation advice/counseling	4/2005
PN-6	Appropriate antibiotic selection	9/2005
PN-7	Influenza vaccination status	1/2006
<i>Surgical Care Improvement / Surgical Infection Prevention (SCIP/SIP)</i>		
SCIP-Inf-1	Prophylactic antibiotic received within 1 hour prior to surgical incision	9/2005
SCIP-Inf-3	Prophylactic antibiotics discontinued within 24 hours after surgery end time	9/2005
Clinical Quality – Outcome Measures		
	30-day AMI mortality	6/2007
	30-day HF mortality	6/2007
Patient-Centered Care Measures		
	HCAHPS	3/2008

Transitioning from RHQDAPU to VBP

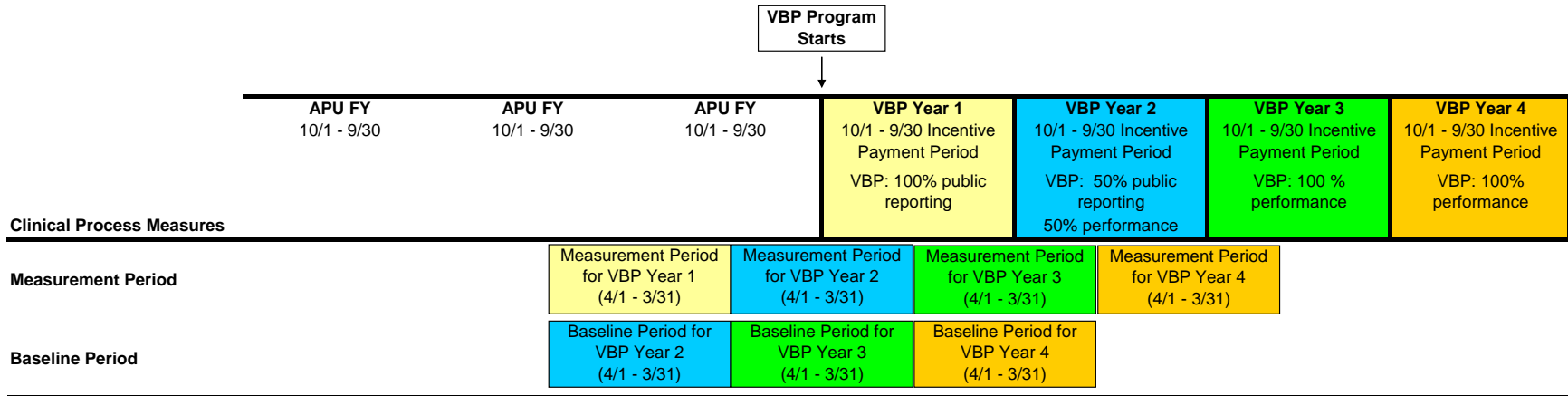
A phased approach for implementing the VBP Program would allow the time needed to: (1) provide hospitals with adequate notice about the set of measures and performance thresholds and benchmarks to be used for the financial incentive payment, (2) accrue baseline performance data on all VBP measures required for determining improvement scores, and (3) establish benchmarks and thresholds for computing attainment scores.

Figure 3 illustrates a timeline linking the “baseline period,” “measurement period,” and fiscal year in which the incentive payment would apply. Fiscal Years 1, 2 and 3 under the VBP Program are color-coded for ease in understanding these relationships. The “measurement period” could start on April 1st and end the following March 31st to tie financial incentives as closely as possible to hospitals’ recent performance, allow four full quarters of data to be used as the basis for determining incentive payments, and still provide sufficient time to receive data from the end of the reporting period needed by CMS for the calculation of performance scores. The “baseline period” would be the 12-month period preceding the “measurement period.” For each individual hospital, its performance scores on the subset of measures selected for the VBP financial incentive during the baseline period could provide the reference point against which year-to-year “improvement” would be determined during the measurement period.

As a second option, improvement could be determined based on performance during the first baseline year for a measure, so that a hospital would not receive credit if performance declined below the first baseline year’s performance. Alternatively, improvement could be measured against the highest level of performance during any prior year, so that a hospital could not receive credit if performance remained below prior performance levels. The performance of all hospitals during this baseline period could be used to establish the program-wide thresholds and benchmarks used for determining each hospital’s “attainment” during the measurement period.

Figure 3 also highlights the lags that are inherent in the proposed VBP Program. These lags necessitate a phased transition from “pay for reporting” on measures included in the RHQDAPU Program to “pay for performance” on the measures selected for the VBP financial incentive. The proposed approach for transitioning from RHQDAPU to the VBP Program is as follows:

Figure 3: Timeline Linking Baseline, Measurement and Performance Payment Periods



- VBP Year 1
 - The incentive payment could be based 100 percent on “pay for reporting” for all VBP measures in order to provide hospitals adequate notice of the incentive measures, thresholds, and benchmarks.
 - The financial incentive could be based on 2 - 5 percent of the base DRG payment, not on the Annual Payment Update used in the RHQDAPU Program.
- VBP Year 2
 - The incentive payment could be based 50 percent on reporting and 50 percent on performance on the VBP incentive payment measures.
 - The Performance Assessment Model would be used to calculate a hospital’s VBP Total Performance Score and translate it into the percent of the financial incentive earned.
- VBP Year 3
 - The incentive payment would be based 100 percent on performance.
- VBP Year 4 and beyond
 - The incentive payment would be based 100 percent on performance.

Hospitals would submit data on all VBP Program measures appropriate to their patient population and service mix, including new measures being introduced, as a requirement for participation in the incentive Program. As illustrated in Figure 2, the design of the VBP Program should ensure that all new performance measures have a preliminary data submission period (with no public reporting) to allow hospitals and their data support vendors to become familiar with the data specifications. Additionally, new VBP measures should be publicly reported for a period of time prior to being included in the financial incentive portion of the Program.

The annual IPPS rule-making process, which currently provides the mechanism for notifying hospitals of and receiving comments on the measures for the RHQDAPU Program, could be used as well to establish measures for the VBP Program.

Redesign of Data Infrastructure and the Data Validation Process for the VBP Program

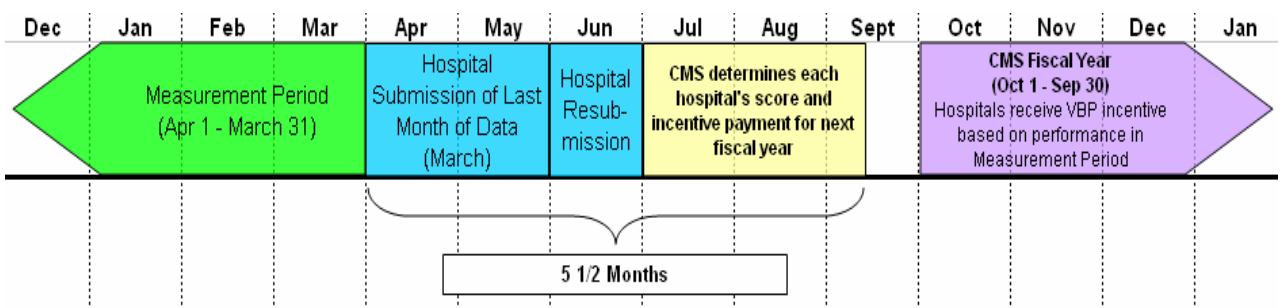
Although the VBP Program is proposed to be built on the foundation of the existing RHQDAPU Program, CMS recognizes that the current data infrastructure may need to be modified and strengthened to fully support VBP Program requirements. In particular, it would be important to ensure the accuracy and improve the timeliness of data used for incentive payment determinations and public reporting. The VBP financial incentive would be applied for each Inpatient Prospective Payment System Fiscal Year, which starts October 1. CMS seeks to have the data used to determine the financial incentive as current as possible.

Streamlined and Improved Data Submission Process

Several improvements for the data submission process could minimize time lags associated with data submission and validation.

- Compress the Data Submission Period:** The structure of the RHQDAPU Program currently allows hospitals 4.5 months (135 days) to submit data following the close of each quarter. This timetable would hamper CMS' ability to make timely determinations for the VBP Program financial incentive and to provide hospitals timely feedback for quality improvement purposes. Under the VBP Program, the submission period could be decreased to 60 days following the close of the reporting period, and hospitals could be encouraged, but not required, to submit monthly (Figure 4). Both steps would significantly improve CMS' ability to allow more recent data to be used for public reporting and incentive payment determinations and to provide more timely feedback.

Figure 4: Detailed Timeline Linking Measurement and Performance Payment Periods



- *Allow Data Resubmissions:* CMS understands that hospitals and their vendors make occasional errors during the submission process. For the VBP Program, hospitals could be allowed up to 30 days after the close of each data submission period (prior to the lock down of the data to determine incentive payments) to resubmit their data. Resubmissions would not be allowed once the deadline for the annual payment determination deadline is reached (*i.e.*, as of July 1 each year).
- *Improve Data Submission Feedback Reports:* Feedback reports and tools could be redesigned for the VBP Program to provide hospitals with a simple, real-time scorecard showing the completeness of their data submissions and whether or not they have met their annual VBP reporting requirements. The scorecard would display preliminary measure rates with comparisons to benchmarks and to previous performance.
- *Enhance User Support:* The existing data submission infrastructure could be expanded to provide full-level user support during business hours in all time zones (7 a.m. Eastern to 7 p.m. Hawaiian). The expanded CMS user support would provide a single source of consistent, accurate, and timely information in response to specific abstraction, measures, and submission questions from hospitals and data vendors. Further details regarding assistance to be provided are presented in Appendix E.

Strengthening Data Validation

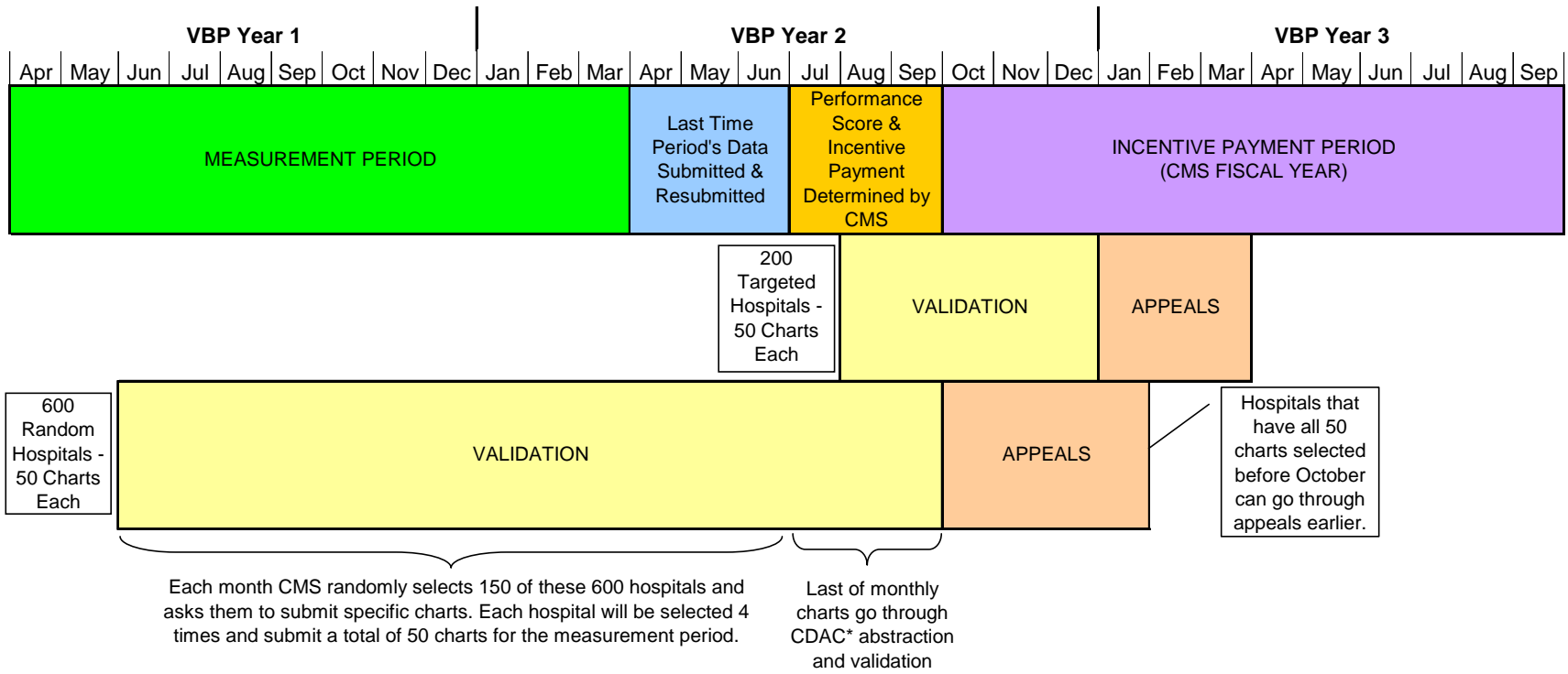
Under the RHQDAPU Program, CMS conducts validation at the data-element level on a small number of records for every participating hospital. This current validation approach involves reviewing a random sample of five charts per quarter per hospital (20 randomly selected charts per year), which is too small a number to assess the accuracy of the performance measure rate.

CMS considers this level of validation insufficient for the VBP Program. The current RHQDAPU data element match approach assesses the abstraction accuracy of the elements but does not assess the impact of combining an inaccurate element with other elements to calculate the measure rate. When tying VBP payment to performance rates, CMS should take additional steps to assure that the rates are correct, not just the data elements used to calculate the rates. The existing validation methodology could be revised to focus on assessing the accuracy of performance measure rates. For Year 1 of the proposed VBP Program, the RHQDAPU

validation process could continue in order to allow CMS to transition to the new approach. Starting with VBP Year 2, the proposed audit strategy could be used. Appendix E provides details that expand upon the following overview of a potential validation program:

- *Select hospitals for validation on both a “random” and “targeted” basis:* CMS could select approximately 600 hospitals each year for the random component of the validation and approximately 200 hospitals for targeted audit. For each hospital selected for validation, CMS would review approximately 50 charts per year. The dual audit selection strategy would serve two functions: (1) enable CMS to assess the overall quality of data submissions and (2) minimize gaming.
- *Focus validation on the accuracy of abstraction to calculate measure rates:* The current RHQDAPU approach focuses on accuracy of the many individual elements used to construct a measure rather than determining the accuracy of a measure rate. Requiring accuracy at the measure level is a stricter definition that is more relevant to the use of measures to determine a financial incentive.
- *Conduct the validation and appeals process post-payment:* This approach would avoid having the validation process delay incentive payment determinations and public reporting. Figure 5 presents an overview of the submission and appeals timeline for the random and targeted samples of hospitals.

Figure 5: Validation Process and Validation Appeals Timeline: Random and Targeted Audits



600 Random Hospitals - 50 Charts Each

VALIDATION

APPEALS

Hospitals that have all 50 charts selected before October can go through appeals earlier.

Each month CMS randomly selects 150 of these 600 hospitals and asks them to submit specific charts. Each hospital will be selected 4 times and submit a total of 50 charts for the measurement period.

Last of monthly charts go through CDAC* abstraction and validation

***Clinical Data Abstraction Center (CDAC)** is the contractor used by CMS to carry out the process of validating data collected from medical records for the RHQDAPU Program and would continue to be used for the VBP Program.

Consequences of Failing Validation

As shown in Figure 5, the validation appeals process for a measurement year would be finalized partway through the incentive payment period (*i.e.*, fiscal year). The first time a hospital fails the validation audit, payment adjustments could be applied prospectively, resulting in an adjusted incentive payment for the *remainder* of that fiscal year. The hospital's performance results on clinical process measures could also be suppressed on Hospital Compare for one quarter of the measurement period. In addition, the hospital could be included in the targeted-audit sample for the next round of validation audits. If a hospital failed validation in consecutive years, CMS would recoup the overpayment.

Strengthening CMS' Ability to Compute Stable Performance Rates

To improve the stability and reliability of performance estimates, the minimum required sample size for each measure could be increased under the VBP Program. In the current RHQDAPU Program, the effective samples of a substantial number of hospitals are small (fewer than 25 cases per condition). This small size may be a function of hospitals either failing to submit enough cases to comply with sampling requirements or applying exclusion criteria too broadly. As part of the VBP Program, the following changes could be made:

- Increase the minimum required effective sample size;
- Educate hospitals about this requirement;
- Require that hospitals submit aggregate population and sample counts so that compliance with data submission and sampling requirements could be assessed; and
- Penalize hospitals for not complying with sampling requirements, as determined in the context of data validation.

Public Reporting

Public reporting is an essential tool for motivating hospitals to improve quality of care and for helping Medicare beneficiaries choose quality providers. The existing Hospital Compare website could serve as the platform for displaying performance results. As a key component of the VBP Program, public reporting could be enhanced to:

- Address the needs of multiple stakeholder audiences,
- Employ display methods and/or decision supports that facilitate fair and accurate decision-making, and
- Ensure consumer understanding of performance data displays.

Measures that can meaningfully contribute to informed consumer decision-making should be part of the VBP Program and publicly reported on the Hospital Compare website, regardless of whether they are included in determining VBP financial incentive payments. CMS would work to modify data displays so that Medicare beneficiaries could more easily interpret performance results.

VBP Program Monitoring and Evaluation

Assessments of the early experience with the VBP Program would allow for timely corrective action and for building the evidence base essential for guiding the design of future CMS VBP programs in other settings of care. These ongoing monitoring and evaluation efforts could be part of CMS' larger efforts to promote an active learning system, both within CMS and between CMS and the hospitals in the VBP Program, to help drive improvements in quality and efficiency. Support for assessment is critical because there is a very limited research base documenting the effect of pay-for-performance schemes, as documented in Appendix A.

To ensure that VBP Program effects are fully understood, a program monitoring methodology should be developed, initiated prior to Program start, and then supported annually. Among the key areas to monitor are:

- *Programmatic impact:* Has VBP (both the financial incentives and public reporting components) improved quality and efficiency within the Medicare program?

- *Distribution of payments:* Does VBP provide similar incentives to all hospitals to improve or maintain their performance, or do some hospitals consistently face challenges in improving or maintaining their performance that VBP may inadvertently exacerbate?
- *Budget neutrality:* Does the VBP Program cause hospitals to increase their volume of services to offset potential losses in income from being at risk for performance?
- *Implementation:* Are there aspects of the VBP Program infrastructure that could be strengthened to help hospitals participate more easily in the Program or to facilitate quality improvement?
- *Best practices:* What best practices are being implemented by high-performing hospitals? How can these strategies be shared with other hospitals to improve care nationally?
- *Unintended consequences:* Does the VBP Program, as a function of its design, result in:
 - Hospitals “teaching to the test” (and potentially reducing quality in unmeasured areas)?
 - Hospitals dropping or avoiding caring for patients who are sicker or more difficult to manage?
 - Increased disparities in care (by region, race/ethnicity, etc.)?
 - Gaming of the data by hospitals to secure incentives?

Appendix F considers these issues in more detail.

Appendix A: Background

This Appendix provides background information on the following topics important to the development of a Medicare Hospital Value-Based Purchasing (VBP) Plan:

- The context for developing a Plan
- Setting the stage for Hospital Value-Based Purchasing
- The process used in Plan development
- Key stakeholder perspectives expressed at two Listening Sessions and presented in written comments regarding the Issues Paper and Options Paper developed for these sessions
- Key findings from the Environment Scan conducted by RAND to support Plan development

The Context for the Hospital VBP Plan

On August 22, 2006, President Bush issued an Executive Order, “Promoting Quality and Efficient Health Care in Federal Government Administered or Sponsored Health Care Programs,” which requires the Federal Government, to the extent permitted by law, to:

- Ensure that Federal health care programs promote quality and efficient delivery of health care using interoperable health information technology, transparency regarding health care quality and price, and better incentives for program beneficiaries, enrollees, and providers.
- Make relevant information available to these beneficiaries, enrollees, and providers in a readily useable manner and in collaboration with similar initiatives in the private sector and non-Federal public sector.

To support this mandate, Department of Health and Human Services (DHHS) Secretary Michael Leavitt has embraced “four cornerstones” for building a value-driven health care system:

- 1) Connecting the health system through the use of interoperable health information technology;

- 2) Measuring and publishing information about quality;
- 3) Measuring and publishing information about price; and
- 4) Using incentives to promote high-quality and cost-effective care.

Building on these four cornerstones, the Centers for Medicare & Medicaid Services (CMS) has articulated a vision for health care—*the right care, for every person, every time*. To achieve this vision, CMS seeks to implement policies that will promote the delivery of care that is safe, effective, timely, patient centered, efficient, and equitable.

Current Medicare hospital payment policies generally reward the quantity rather than the quality of care delivered and provide neither incentive nor support for improving quality of care. Today, hospitals are usually paid the same for services rendered regardless of the quality of care they provide, and in some cases, hospitals may even receive additional payment for treatment of avoidable complications. Value-based purchasing (VBP), which links payment more directly to performance, is a key policy mechanism that would transform Medicare from a passive payer for services to an active purchaser of care for millions of Medicare beneficiaries. CMS would focus on purchasing value for the Medicare program, which means that hospitals would receive differential payments depending on their performance. VBP is a key policy mechanism to achieve desired programmatic goals.

Goals for CMS' VBP Initiatives

- Improve clinical quality
- Address problems of underuse, overuse, and misuse of services
- Encourage patient-centered care
- Reduce adverse events and improve patient safety
- Avoid unnecessary costs in the delivery of care
- Stimulate investments in structural components and the re-engineering of care processes system-wide
- Make performance results transparent to and useable by consumers
- Avoid creating additional disparities in health care and work to reduce existing disparities

Setting the Stage for Value-Based Purchasing

The Medicare Prescription Drug Improvement and Modernization Act of 2003 (MMA),¹ required that Medicare establish the Reporting Hospital Quality Data for Annual Payment

¹ Public Law 108-173, December 8, 2003.

Update (RHQDAPU) Program, a quality pay-for-reporting (P4R) program for hospitals reimbursed under the Inpatient Prospective Payment System (IPPS). Beginning in August 2004, hospitals submitted data on a defined set of performance measures as a requirement for receiving a specified percentage of the IPPS Annual Payment Update (APU). Initially, 0.4 percentage points of the APU was at risk for reporting on a set of ten quality measures. For FY 2005, virtually every IPPS hospital in the country (98.3 percent) submitted the required data, and approximately 96 percent of all participating hospitals met the program requirements and received the full APU. The performance data are made available to Medicare beneficiaries and the public through the Hospital Compare website (<http://www.hospitalcompare.hhs.gov>).

For FY 2007 and subsequent years, Section 5001(a) of the Deficit Reduction Act extended hospital P4R by requiring the Secretary of DHHS to expand the set of P4R performance measures and to increase the differential payment from 0.4 to 2 percentage points. For the FY 2007 APU, the Secretary was directed to begin adopting the baseline set of hospital performance measures specified by the Institute of Medicine (IOM) report released in December 2005,² which includes 22 Hospital Quality Alliance (HQA) measures and the HCAHPS survey of patients' perspectives of care. For the FY 2007 APU, hospitals have submitted 11 clinical process-of-care measures in addition to the original starter set of ten measures, for a total of 21 measures. For the FY 2008 APU, six additional measures have been adopted for IPPS hospitals per the Calendar Year (CY) 2007 Outpatient Prospective Payment System (OPPS) Final Rule³: 1) HCAHPS survey of patients' perspectives of care; 2) three surgical care infection prevention measures; and 3) two 30-day mortality measures for acute myocardial infarction (AMI) and heart failure (HF). Thus, submitting data on a total of 27 measures will serve as the basis on which hospitals qualify for the full FY 2008 APU.

² Institute of Medicine, "Performance Measurement: Accelerating Improvement." December 1, 2005, available at <http://www.iom.edu>.

³ 71 FR 68201

Process Used to Develop the Value-Based Purchasing Plan

CMS created an internal Hospital VBP Workgroup with responsibility for developing the VBP plan for Medicare hospital services. The Workgroup was organized into four Subgroups, each of which addressed one of the required planning issues:

- Measures,
- Data infrastructure and validation,
- Incentive structure, and
- Public reporting.

In collaboration with the DHHS Office of the Assistant Secretary for Planning and Evaluation (ASPE), CMS contracted with RAND to support the Workgroup as a whole. To provide detailed support to the development of measures and an approach to scoring performance, CMS also contracted with a team led by Brandeis University with subcontract to Booz-Allen-Hamilton and Boston University.

CMS commenced the VBP planning process in April 2006 by seeking public feedback during the FY 2007 IPPS rule-making process, which referenced the DRA mandate and the planning process.⁴ Development of the VBP Plan occurred between September 2006 and June 2007. During this period, the CMS Hospital VBP Workgroup and its support contractors:

- Conducted a review of the hospital P4P literature,
- Conducted an environmental scan of the existing P4P landscape to inform CMS' consideration of the array of design options,
- Prepared a set of design issues, presented in *Medicare Hospital Value-Based Purchasing Plan Development: Issues Paper* (available on the CMS website at <http://www.cms.hhs.gov/AcuteInpatientPPS/downloads/hospitalVBPplanissuespaper.pdf>)
- Held an initial public Listening Session on January 17, 2007, to consult with affected stakeholders on various design issues (Listening Session 1),

⁴ Revision to Hospital Inpatient Prospective Payment Systems and Fiscal Year 2007 Rates. 71 Fed Reg 47870 (Aug 18, 2006) (amending 42 CFR Parts 409, 410, 412, 413, 414, 424, 485, 489, and 505), available at: <http://www.cms.hhs.gov/QuarterlyProviderUpdates/Downloads/CMS1488F.pdf>.

- Narrowed the set of design issues and prepared an options paper, *Medicare Hospital Value-Based Purchasing: Options Paper*, for Listening Session 2 (<http://www.cms.hhs.gov/AcuteInpatientPPS/downloads/HospitalVBPOptions.pdf>),
- Held a second public Listening Session on April 12, 2007, to solicit stakeholder feedback on the design options, including the strategy CMS was considering for scoring hospital performance, and
- Prepared a Plan for the Medicare Hospital VBP Program.

Stakeholder Perspectives

During the process used to develop the VBP Plan, CMS conducted active outreach efforts to understand the perspectives of affected stakeholders. CMS hosted two public Listening Sessions during which stakeholders were asked to present and/or submit in writing comments based on first an Issues Paper and then an Options Paper prepared as part of VBP Plan development. Comments from more than 100 stakeholders revealed broad support for the following principles that have guided VBP Plan design:

Incentives

- Provide incentives for both improvement and attainment.
 - “Raise all boats”—do not pick winners and losers.
 - Encourage excellence, but also spread payments broadly to engage and incentivize more hospitals.
- Tie incentive payments to only the base Diagnosis-Related Group (DRG) portion of a hospital’s payment. Exclude capital costs, disproportionate share and medical education payments from the amount that is held “at risk” based on performance.

Measures

- Use absolute performance thresholds, specified in advance, so that hospitals know what target they must hit to secure the incentive payment.

- Do not retire “topped-out” measures prematurely—hospitals need positive feedback about things they are doing well, as well as constructive feedback on areas needing improvement.
- Create a single VBP Program in which rural and small hospitals can participate.
 - Include measures that address services that small and rural hospitals provide—such as transfers and care coordination.
- Expand the set of performance measures to achieve a more comprehensive assessment of hospital performance.
- Work with consensus bodies and other reporting entities to align measures and minimize the burden of reporting.

Data Infrastructure and Validation

- Develop data resubmission process so that hospitals can correct errors in data.
- Improve the current validation process by using a combination of random and targeted audits and by drawing larger samples to assess data validity.
- Conduct validation work post hoc so as not to delay payments and public reporting of results.
- Provide support to hospitals at all stages of the process, particularly in the areas of data submission and ways in which to improve their performance.

Public Reporting

- Simplify the Hospital Compare website for ease of use by consumers.
 - Provide decision support tools that will help consumers use the information more effectively.
 - Develop composite or summary measures, such as at the condition level.
- Disclose uncertainty and variability in scores based on small numbers, but avoid negative labeling when data are suppressed because of small numbers.

Evaluation and Ongoing Monitoring

- Because VBP is a new concept and “best practice” is not yet known, develop an ongoing monitoring system that will support program adjustments.

- Monitor for unintended consequences, such as reduced access to care for vulnerable patient populations, and adjust the program accordingly.

Summary of Findings from the Hospital Pay-for-Performance Environmental Scan and Literature Review

In the fall of 2006, RAND conducted an environmental scan to summarize what is known about hospital P4P. The environmental scan consisted of two components:

- A literature review of empirical studies examining the impact of hospital P4P and
- Discussions with P4P program sponsors and hospitals to understand real-world experiences with P4P and P4R.

Discussions were held with private and public sector hospital P4P program sponsors, hospitals participating in P4P or P4R, hospital associations, and data vendors. The emphasis in these discussions was on defining the design components of currently operating hospital P4P programs, understanding experiences with P4R and P4P, identifying key lessons learned to inform the development of the Plan for a Medicare Hospital VBP Program.

Review of the Published Literature

There is little formal evaluation occurring of existing hospital P4P programs. As of February 2007, there were only seven published studies addressing the impact of three different hospital P4P programs:

- 1) Hawaii Medical Service Association P4P Program,
- 2) Michigan Blue Cross Blue Shield Program, and
- 3) The CMS-Premier Hospital P4P Program (PHQID).

All three of the programs examined by these studies target only the inpatient setting with no focus on the hospital outpatient setting.

A high-level summary of the study designs and results is presented in Table A-1. Most of the published studies have significant methodological limitations that hamper our ability to understand the impact of these programs and how various design features influence the observed results. Five of the seven studies did not use control groups,

which is a critical issue in evaluating the impact of a P4P program. One study that used a control group only included 10 hospitals, and it is unclear whether the controls used were appropriate. A variety of quality improvement efforts in addition to the specific P4P program being evaluated, including efforts by the Joint Commission and CMS to measure and improve quality, the RHQDAPU Program, and other P4P programs, could cause increased performance on quality metrics. Documented national temporal trends towards increasing performance on many hospital quality metrics indicate the need to use control groups in P4P program evaluation to be able to distinguish between the program's effects and these national trends.

The five articles that examined changes over time observed improvements in at least some of the hospital performance measures or condition-specific composites included in the study. The two studies that included control groups saw very modest improvements in performance associated with P4P compared to improvements accomplished through public reporting.^{5, 6} There is even less evidence of the effect of P4P on patient outcomes. Berthiaume and colleagues⁷ found improvements in complication rates for obstetrical and surgical patients in an uncontrolled study, but did not report whether the improvements were statistically significant. The studies on PHQID did not analyze the patient outcome measures for the program separately, but included them in condition-specific composites. Furthermore, there is no literature assessing the relationship between specific design elements and programmatic impact.

Another limitation in the ability to draw conclusions from these few studies is that these P4P programs generally focused on a small set of process measures covering a handful of diagnoses. Therefore, it is unclear whether their findings generalize to other kinds of measures, such as patient experience, which was included in only one of the three programs examined and on which none of the articles reported. It is also unknown what

⁵ Grossbart, S. R. What's the return? Assessing the effect of "pay-for-performance" initiatives on the quality of care delivery. *Med Care Res Rev* 2006;63(1 Suppl): 29S-48S.

⁶ Lindenauer, P. K., et al. Public reporting and pay for performance in hospital quality improvement. *N Engl J Med* 2007;356(5): 486-96.

⁷ Berthiaume, J. T., et al. Aligning financial incentives with quality of care in the hospital setting. *J Healthc Qual* 2006;28(2): 36-44, 51.

happens to care for unmeasured areas of care if the focus is shifted to a limited set of measures for selected aspects of care.

Table A-1: Summary of Empirical Studies Examining Hospital P4P Programs

P4P Program	Article	Type of Study	Change in Performance	Control Group
Michigan Blue Cross Blue Shield	Reiter et al. (2006) ⁸	Survey of participating hospitals to track behavioral responses	No	No
	Nahra et al. (2006) ⁹	Cost-effectiveness analysis	Yes	No
Hawaii Medical Service Association	Berthiaume et al. (2004) ¹⁰	Describes uptake of one component of program and how many dollars were dispensed	No	No
	Berthiaume et al. (2006) ¹¹	Describes trends in measures	Yes	No
Premier Hospital Quality Improvement Demonstration (PHQID)	Premier White Paper ¹²	Describes improvements in quality measures	Yes	No
	Grossbart (2006) ¹³	Evaluates improvements in quality versus a “matched” control group	Yes	Yes
	Lindenauer et al. (2006) ¹⁴	Evaluates improvements in quality versus a “matched” control group	Yes	Yes

⁸ Reiter, K. L., et al. Hospital responses to pay-for-performance incentives. *Health Serv Manage Res* 2006;19(2): 123-34.

⁹ Nahra, T. A., et al. Cost-effectiveness of hospital pay-for-performance incentives. *Med Care Res Rev* 2006;63(1 Suppl): 49S-72S.

¹⁰ Berthiaume, J. T., et al. Aligning financial incentives with "Get with the Guidelines" to improve cardiovascular care. *Am J Manag Care* 2004;10(7 Pt 2): 501-4.

¹¹ Berthiaume JT, Chung RS, Ryskina KL, Walsh J, Legorreta AP. (2006) Aligning Financial Incentives with Quality of Care in the Hospital Setting. *Journal for Healthcare Quality* 28(2):36-44, 51.

¹² Premier, Inc. (2006) *Centers for Medicare and Medicaid Services (CMS)/Premier Hospital Quality Incentive Demonstration Project: Project Overview and Findings from Year One*. Charlotte, NC.

¹³ Grossbart SR. (2006) What's the Return? Assessing the Effect of "Pay-for-Performance" Initiatives on the Quality of Care Delivery. *Medical Care Research and Review* 63(1 Suppl):29S-48S.

¹⁴ Lindenauer PK, Remus D, Roman S, Rothberg MB, Benjamin EM, Ma A, Bratzler DW. (2007) Public Reporting and Pay for Performance in Hospital Quality Improvement. *New England Journal of Medicine* 356(5):486-496.

Summary of Discussions with Hospital P4P Program Sponsors, Hospitals Participating in Those Programs, Hospital Associations, and Data Vendors

Given the absence of a strong evidence base showing the impact of pay-for-reporting and pay-for-performance programs as well as the implications of various designs, RAND held discussions with a variety of organizations to understand their experiences with P4P and P4R. Between October 2006 and March 2007, discussions were held with the following types of organizations:

- 23 P4P program sponsors and 3 organizations without established P4P programs that are leaders in public reporting and other incentives for providers;
- 27 hospitals in 5 categories, including:
 - Hospitals eligible for the PHQID demonstration (7 participating and 5 non-participating),
 - 5 hospitals participating in a private sector P4P program,
 - 6 small or critical access hospitals (CAHs) that submitted data for RHQDAPU and appear on the Hospital Compare website,
 - 3 hospitals that failed RHQDAPU data validation, and
 - 1 hospital that elected not to participate in RHQDAPU in FY 2006;
- 8 hospital associations;
- 6 data vendors; and
- 7 organizations with expertise in rural health issues.

Appendix A-1 provides a list of the organizations with which discussions were held.

Summary of Discussions with Hospital P4P Program Sponsors

Improving the quality of care delivered to patients was overwhelmingly the primary goal of the P4P program sponsors we talked to as part of the environmental scan (n = 21). Improving efficiency (n = 6) and patient safety (n = 5) were also commonly mentioned as goals. All of the programs included measures of clinical quality, most commonly the Joint Commission core measures (n = 10) or RHQDAPU measures (n = 12). Measures of patient safety (n = 16); resource use (n = 11); structure, such as use of information technology (n = 9); patient experience (n = 9); and quality improvement (n = 8) were also common. Criteria that program sponsors used in measure selection included, consistency

with other reporting activities, minimizing hospital burden, evidence-based or endorsed by known organizations, availability of data, and ease of data collection. Currently, there is minimal inclusion of hospital outpatient measures in private-sector P4P programs, though there was general agreement that outpatient services is an important area that should be addressed. Two reasons cited for not including the outpatient setting in hospital P4P programs are 1) the absence of readily available measures that are applicable to the hospital outpatient setting, and 2) some concerns that the outpatient setting is not within the hospital's locus of control. When included, outpatient measures focused on emergency room satisfaction, or simple utilization or cost measures.

The P4P programs were often voluntary, and often implemented through contract negotiations. However, most hospitals approached by the program sponsors signed up for the incentive programs. Specialty hospitals and small or critical access hospitals were not typically included in these programs.

Most hospital P4P programs were not conducting formal evaluations. There was no tracking of unintended consequences by the program sponsors we spoke with, but they recognize that evaluation is an important activity. Many program sponsors wanted to know the return on investment of their programs, but they had not determined the return within their own programs.

The financial incentives were usually structured as either a bonus (n = 10) or a differential update in future year's payment increase (either per diem or DRG) (n = 9). To determine the financial incentives, program sponsors most frequently used improvements over time (n = 12), relative performance thresholds (n = 10), or absolute thresholds, which were frequently based on national percentile ranking from the prior year. Eight sponsors reported using several of these methods in a single program. Many programs used weighted measure domains to determine the financial incentive. These weights were often tailored to individual hospitals in small programs. Many of the programs also included non-financial incentives, including public reporting (n = 12), peer comparisons reported back to hospitals (n = 11), public recognition (n = 5), or tiering (n = 2).

Most of the program sponsors involved hospitals in program design (n = 15). The majority provided hospitals with some type of assistance (n = 13), most often in the form of education or technical assistance. Most of the programs were evolving over time with the most common change reported being the expansion or modification of the measure set. Programs were working to include more measures of outcomes, patient experience, cost and a broader set of clinical measures. Program sponsors were trying to make the programs broader and deeper without overburdening hospitals. Other changes program sponsors were making include increasing the number of participants, increasing the size of the incentive, and an increasing focus on consumer engagement.

Summary of Discussions with Hospitals and Hospital Association on P4P and P4R

Hospitals universally noted that RHQDAPU and PHQID have allowed them to prepare for VBP. Many hospitals believe that that VBP is inevitable and they are making efforts to prepare. Most hospitals were positive about their experiences participating in P4P and RHQDAPU, that the programs are addressing important areas, and that program sponsors are, for the most part, using measures for which hospitals should be accountable. However, hospitals acknowledged the challenge of changing physician behavior without aligned incentives for physicians. Hospitals reported that RHQDAPU has caused important changes in their organizations. There is a more proactive focus on quality improvement efforts in hospitals, and hospital boards now focus on quality performance in addition to financial and management issues. The boards frequently review data from the Hospital Compare website to see their progress on the indicators and how their hospital stands relative to their peer hospitals.

Hospitals have, however, found that participating in P4P and P4R is resource intensive due to the combined needs of data collection, submitting data for the validation process, and continuous quality improvement efforts. Hospitals noted that the initial financial incentive for RHQDAPU (0.4 percent of the APU) did not match the level of investment needed to participate in the program. Some hospitals thought that the increased RHQDAPU financial incentive for FY 2007 (2.0 percent of APU) might cover the costs of participation. Hospitals and hospital associations viewed the coordination and

alignment of a CMS VBP Program with other P4P programs and hospital reporting requirements as critical for keeping the burden of participation manageable for hospitals.

Small hospitals experienced more challenges participating in P4P and RHQDAPU than did larger hospitals. In general, they had fewer resources to collect data. They also faced the problem of having only a few cases to report for the clinical conditions of interest or the areas of focus for the program were not services that they provided.

Hospitals and hospital associations felt strongly that measures should be evidence based and endorsed by consensus organizations, such as NQF and HQA. They also felt strongly that data collection burden should be a measure selection criterion. Hospitals were hesitant about the use of outcome or patient experience measures as the basis for financial incentives. They were concerned about the ability of risk-adjustment to adequately account for differences in patient populations, but felt its use was necessary to minimize the possible unintended consequence of risk selection. Pilot testing of new measures was viewed as critical to work out the measure specifications and allow hospitals to become familiar with the measures.

Hospitals in PHQID and those exposed to private sector P4P programs were asked to consider various P4P payment options expressed dislike for the use of relative thresholds to determine payments. Relative thresholds were viewed as penalizing hospitals with high, but not top, performance, particularly when performance scores are compressed at the top end of the range of scores. There was strong support by program sponsors and hospital associations for tying incentives to improvement in performance, but there was less support among hospitals for this approach. Hospitals strongly preferred the use of absolute thresholds, so that all hospitals that perform well can receive quality incentive payments.

Physician engagement was a challenge experienced by some hospitals. Hospitals stated CMS should directly incentivize the physicians and align physician measurement with hospital measurement. An alternative suggestion was changing restrictions on

gainsharing so that hospitals themselves can provide financial incentives to their doctors to achieve alignment.

Appendix A-1: List of Organizations Participating in the Environmental Scan

Hospital Pay-for-Performance and Public Reporting Program Sponsors

Anthem – National office	Kaiser Permanente – National,
Anthem VA	Northern, and Southern CA offices
BCBS HI	Leapfrog Group (Hospital Rewards
BCBS IL	program that organizations can
BCBS MA	license)
BCBS MI	Maine Health Management Coalition
Blue Shield Northeastern NY	PacifiCare/United Healthcare
Employers’ Coalition on Health	Premier Health System
Excellus/Univera	Priority Health
Fallon Community Health Plan	Providence Health Plan
Harvard Pilgrim Health Plan	Regence Blue Shield
Health Partners	The Employer Healthcare Alliance
Highmark BCBS	Cooperative ("The Alliance")
Horizon BCBS NJ	Tufts Health Plan
Independent Health	The Veterans Administration
	Anonymous program sponsor (1)

Hospitals and Health Systems

Amsterdam Memorial Hospital – Amsterdam, NY	McLeod Medical Center – Florence, SC
Baptist Health System of East TN – Knoxville, TN	Mercy Medical Center – Centerville, IA
Bleckley Memorial Hospital – Cochran, GA	Park Nicollet – St. Louis Park, MN
Crenshaw Community Hospital – Luverne, AL	Rice County District One Hospital – Faribault, MN
Fairchild Medical Center – Yreka, CA	San Luis Valley Regional Medical Center – Alamosa, CO
Foote Memorial Hospital – Jackson, MI	South Central Regional Medical Center – Laurel, MS
Franklin Medical Center – Greenfield, MA	Southwestern General Hospital – El Paso, TX
Geisinger Health System – Danville, PA	Spruce Pine Community Hospital – Spruce Pine, NC
Hackensack University Medical Center – Hackensack, NJ	St. John Health System – Warren, MI
Henry Ford Health System – Detroit, MI	St. Joseph Hospital – Polson, MT
Hopi Health Care Center – Polacca, AZ	St. Jude Medical Center – Fullerton, CA
Kaiser Permanente – CA	Trinity Health System – 20 hospitals in 7 states
	Walla Walla General Hospital – Walla Walla, WA

White River Medical Center –
Batesville, AR

William Beaumont Hospital – Royal
Oak, MI
Anonymous hospitals (2)

Hospital Associations

American Hospital Association
Association of American Medical
Colleges
Catholic Health Association
Federation of American Hospitals

National Association of Children’s
Hospitals & Related Institutions
North Carolina Hospital Association
South Dakota Hospital Association
Voluntary Hospital Association

Data Vendors

Hospital Corporation of America
Illinois Hospital Association
Maryland Hospital Association

Premier Health System
Quantros
Thomson Healthcare

Other Organizations

Cypress Healthcare
Kansas Department of Health and
Environment, Office of Local and Rural
Health
Health Resources and Services
Administration, Office of Rural Health
Policy
National Rural Health Association
Stratis Health (Minnesota QIO)
Stroudwater Associates
Upper Midwest Rural Health Research
Center

Appendix B: Details of the Performance Assessment Model

Appendix B provides essential details on the following aspects of the potential Performance Assessment Model:

- Scoring the clinical process-of-care measures
- Scoring the HCAHPS patient perception-of-care survey
- Creating the VBP Total Performance Score
- Scoring the 30-day mortality measures

Scoring Clinical Process-of-Care Measures

Setting Benchmarks and Attainment Thresholds

To provide an empirical basis for designing the Performance Assessment Model and for testing the scoring of the clinical process-of-care measures, a database was created containing the 20 process-of-care measures reported on Hospital Compare by more than 3,000 IPPS hospitals for 2004 and 2005 (the most recent data available). Analyses were conducted to explore different statistical approaches to establishing attainment thresholds and benchmarks and to examine issues associated with hospitals having only a small number of cases to report on individual measures and/or being able to report on only a few measures. The attainment thresholds and benchmarks used in the clinical measure examples presented below to illustrate the potential Performance Assessment Model were developed based on analysis of this database.

Analysis of the 2004–2005 Hospital Compare data demonstrated that the approach used to set benchmarks and attainment thresholds for measures on which hospital performance is broadly distributed would not suffice for “topped-out” measures (*i.e.*, measures whose value for the 75th percentile is not statistically different from that for the 90th percentile) and that a new approach was needed. Table 1 displays potential clinical process-of-care benchmarks and attainment thresholds used for the approach analyzed in this report.

Table B-1: Benchmarks and Attainment Thresholds for Clinical Process-of-Care Measures

Measure Designation	Benchmark	Attainment Threshold
Standard method for measures with a broad distribution of performance scores	Mean of the top decile	50 th percentile
Method for topped-out measures	90% performance	60% performance

For clinical measures with a broad distribution of performance scores, the standard method for establishing the benchmark that could be used: the benchmark, which represents exemplary performance, could be set at the mean value of the top-performing 10 percent of all hospitals in the previous reporting period. The attainment threshold could be defined as the performance of the median hospital (50th percentile performance) in the previous reporting period. These parameters would be used to judge performance in the current performance period. Therefore, in this example, hospitals that perform at least as well in the current period as the mid-performing hospital did in the previous period would earn points for attainment.

As all hospitals improve their performance over time on any given measure, variation in performance would decrease, and the distribution of hospital scores for that measure would concentrate at high values. As hospital performance on a measure improves, the values of both the benchmark and attainment threshold would increase. Where hospital performance is very concentrated at high values, a measure has topped out.

Scoring a topped-out measure presents a number of challenges. First, requiring hospitals to meet or exceed an empirical benchmark that is statistically indistinguishable from a perfect score in order to earn all ten points could result in unintended consequences as hospitals strive to achieve the top tail of the distribution. Examples of unintended consequences may include, but are not limited to, inappropriate delivery of a service to some patients (such as delivery of antibiotics to patients without a confirmed diagnosis of pneumonia), unduly conservative decisions on whether to exclude some patients from being counted toward the measure, and a focus on achieving a perfect score at the expense of real improvements in quality or patient outcomes. Second, for

topped-out measures, it is more difficult to differentiate among hospitals performing above the median in a meaningful clinical or practical way.

However, topped-out measures could still be considered valid indicators of appropriate care and should be included in the VBP Program. It is important to reinforce high standards of care and to convey to Medicare beneficiaries the excellence in care they can expect to receive. For those instances in which a measure has topped out, CMS could use a method that sets the clinical process-of-care benchmark at .90, which represents a 90 percent performance rate. This would allow a great number of high-performing hospitals to reach the benchmark and earn all ten points for the measure. The attainment threshold for topped-out measures could be set at .60, which represents a 60 percent performance rate. This defines a fairly large range over which hospitals could earn points for attainment. Hospitals could also earn points for improvement on topped-out measures. Indeed, part of the rationale for continuing to incentivize topped-out measures is to drive quality improvement among the subset of hospitals that have not yet achieved a high level of performance.

The empirically determined benchmark and attainment thresholds used in Figures B-1-B-3 are for illustrative purposes only. CMS would empirically establish benchmarks and attainment thresholds for any given period using national data from the previous reporting period. The actual benchmarks and attainment thresholds Year 1 of the VBP Program would be established using the most recent data available at the start of Program implementation.

CMS considered several approaches for determining thresholds and benchmarks. Based on empirical analysis, demonstrably high and achievable standards of excellence were selected. This evidence-based approach supports setting the benchmark at the mean of the top decile, and the attainment threshold at the 50th percentile.

Another key decision was whether the benchmarks and thresholds should be based on prior experience (*e.g.*, last year) or current experience (*e.g.*, the ranking model used for the first phase of the Premier Hospital Quality Incentive Demonstration). However, findings from the environmental scan and listening sessions indicated that hospitals overwhelmingly prefer to know in advance the performance standards that define the attainment range. This knowledge

would allow them to set their priorities and to gauge their own achievable targets for improvement. Accordingly, the benchmarks and thresholds could be based on the distribution of national hospital performance during the previous 12-month baseline reporting period. This approach balances the policy objectives of defining meritorious performance on achievable standards of excellence and hospitals' need to plan and implement activities that respond most appropriately to the VBP incentive structure.

Scoring Performance Based on Attainment

In the approach analyzed for this report, for each VBP measure that counts toward the incentive payment, a hospital would earn 0–10 points for attainment based on where its score for the measure falls relative to the attainment threshold and the benchmark. All attainment points would be rounded to the nearest whole number (*e.g.*, 9.6 points would be rounded to 10). If a hospital's score was:

- Equal to or greater than the benchmark, the hospital receives 10 points for attainment.
- Within the attainment range (*i.e.*, greater than the attainment threshold but below the benchmark), the hospital receives a score of 1–9 based on a discrete linear scale established for the attainment range.
- Equal to or less than the attainment threshold (*i.e.*, the lower bound of the attainment range), the hospital receives 0 points for attainment.

Scoring Performance Based on Improvement

In the approach analyzed for this report, for each VBP measure tied to the incentive payment, a hospital would earn 0–9 points based on how much its performance on the measure has improved since the previous period.¹⁵ A unique improvement range for each hospital on each VBP measure could be established to define the distance between a hospital's prior-period score on a measure and the national benchmark for the measure. All improvement points would be rounded to the nearest whole number. If a hospital's score was:

¹⁵ If a hospital's performance meets or exceeds the benchmark, no improvement score would need to be calculated.

- Between its previous-period score and the benchmark, within the improvement range, then the hospital would receive a score of 0–9 based on the discrete linear scale that defines the improvement range.
- Equal to or lower than its previous-period score on the measure, then the hospital would receive 0 points for improvement.

Another approach is one that assesses improvement for a given hospital by referencing performance to the highest percentile previously achieved for that hospital on that measure, rather than by comparing performance only to previous-period performance on a measure. Such an approach would improve the reliability of improvement scores and increase the standard by which improvement points are awarded. In particular, this approach would reduce the improvement points that result for a hospital that alternately improves and then regresses to a previous performance level. While there would be no penalty for regression in performance scores, there would also be no points awarded for surpassing the previous period’s performance on a given measure while remaining below a level of performance achieved two or more periods prior.

Examples to Illustrate Clinical Process of Care Measure Scoring

Three examples are presented here to illustrate how the potential Performance Assessment Model could be applied in the context of clinical process-of-care measures. The hospitals in these examples were selected from the empirical database created from the 2004–2005 data to support model development, and all performance scores would be calculated for the pneumonia measure “patients assessed and given pneumococcal vaccine.” (Appendix B further illustrates the potential Model’s application by demonstrating how four hospitals would be scored on five clinical measures and how the resulting performance scores would be calculated.)

Figure B-1 shows the scoring for Hospital B. The benchmark calculated for the pneumonia measure in this case was 0.87 (mean value of the top decile in 2004), and the attainment threshold was 0.47 (performance of the median of 50th percentile hospital in 2004). Hospital B’s

2005 performance rate of 0.91 on this measure¹⁶ exceeds the benchmark, so Hospital B would earn 10 (the maximum) points for attainment. Because Hospital B has earned the maximum number of points possible for this measure, its improvement score would be irrelevant.

Figure B-1: Example of Hospital Earning Points by Exceeding Benchmark, Clinical Process-of-Care Measure Scoring Model

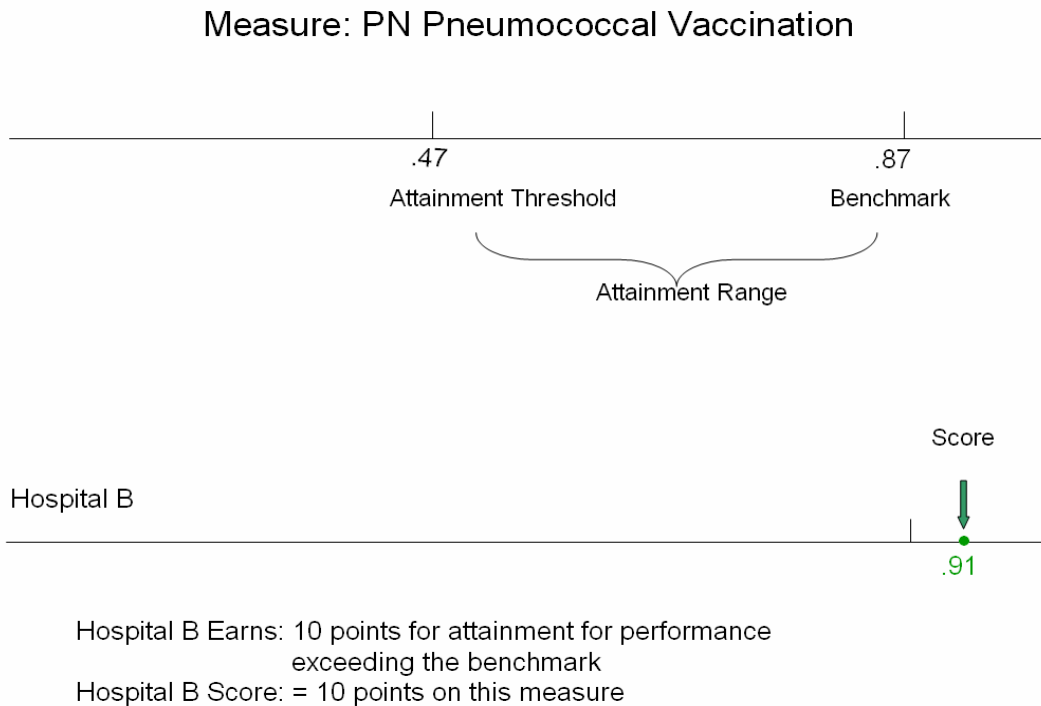
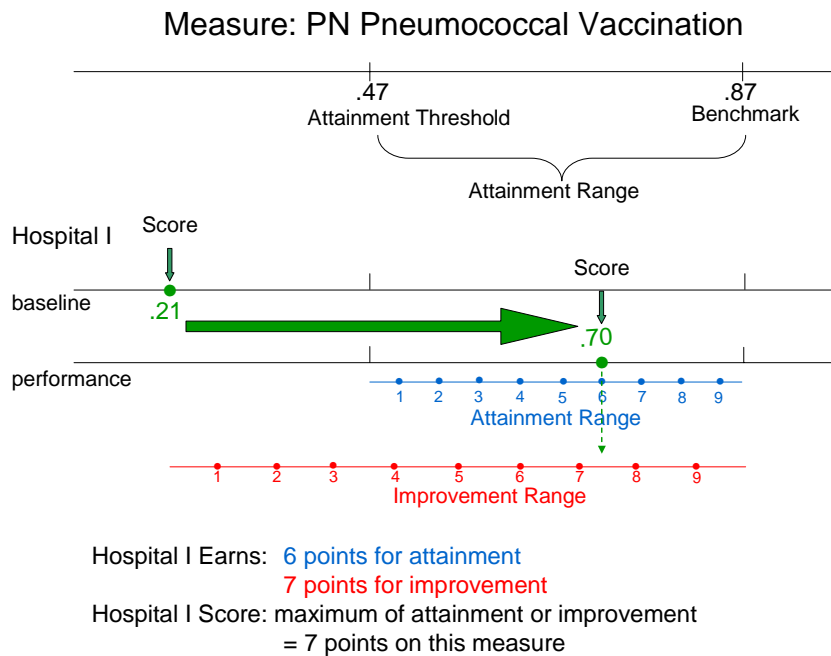


Figure B-2 shows the scoring for another hospital, Hospital I. As can be seen, the hospital’s performance on this measure went from 0.21 (below the attainment threshold) in the previous period to 0.70 (above the attainment threshold) in the current period. Applying the attainment scale, Hospital I would earn 6 points for this measure. However, because Hospital I’s current-period performance is also greater than its previous-period performance, it would be scored based on improvement as well. According to the improvement scale, Hospital I’s period-to-period improvement, from 0.21 to 0.70, Hospital I would earn 7 points. Because the higher of

¹⁶ A hospital’s performance rate on a measure is expressed as a decimal. In the illustration, Hospital B’s performance rate of 0.91 means that 91 percent of applicable patients admitted for pneumonia were assessed and given the pneumococcal vaccine.

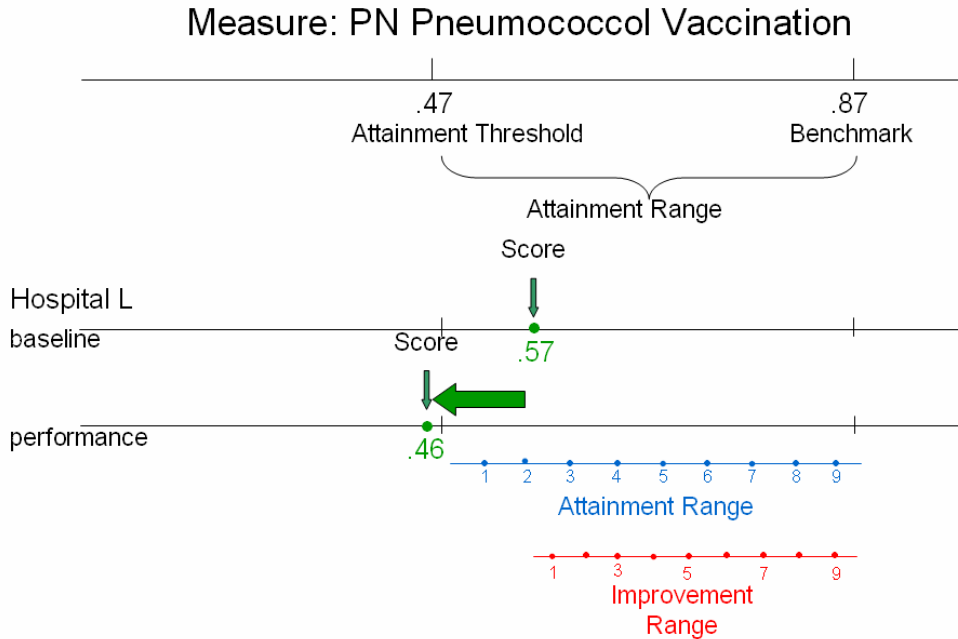
the two scores is used for determining the measure score, Hospital I receives 7 points for this measure (rounded to the nearest whole number).

Figure B-2: Example of Hospital Earning Points by Attainment or Improvement, Clinical Process-of-Care Measure Scoring Model



In Figure B-3, Hospital L's performance on the pneumonia measure drops from 0.57 to 0.46 (a decline of 0.11 points). Because this hospital's performance is lower than the threshold of 0.47, it receives 0 points based on attainment. It would also receive 0 points for improvement, because its current-period performance is lower than its prior-period performance. In this example, Hospital L would receive 0 points for the measure.

Figure B-3: Example of Hospital Earning No Points, Clinical Process-of-Care Measure Scoring Model



Hospital L Earns: 0 points for attainment
 0 points for improvement
 Hospital L Score: maximum of attainment or improvement
 = 0 points on this measure

Calculation of the Total Clinical Process-of-Care Performance Score

A hospital's overall clinical performance score would be based on all measures that count toward the financial incentive for which the hospital submitted data and for which it had a sufficient number of cases. The number of measures for each hospital would vary, of course, depending on the services the hospital provides (*e.g.*, some hospitals may not perform percutaneous coronary intervention; therefore, this measure would not apply to them). As described above, for each applicable measure, a hospital would receive a score of 0–10 based on the higher of its attainment and improvement scores.

The points earned for each clinical process measure are summed to determine the *total earned points*:

Total earned points = Sum of points earned across all reported measures

Each hospital also would have a corresponding universe of *total possible points* for the clinical process measures, calculated as follows:

Total possible points = Total number of measures reported by hospital x 10 points

The hospital's *clinical process-of-care performance score* is a percentage, calculated as follows:

Clinical process-of-care measures performance score = Total earned points / Total possible points x 100%

Because the performance score would be based only on the measures for which a hospital could report, which depend on the hospital's patient population and service mix, the scores would be normalized across hospitals that report different numbers of measures. (See Appendix B-1 for scoring examples.)

Approach to Scoring HCAHPS Survey of Patients' Perspectives of Care

As part of the expansion of the FY 2007 RHQDAPU measure set, the HCAHPS survey of patients' perspectives of care was added to the set of performance measures on which hospitals are required to report to receive their full APU for FY 2008. (See Appendix B-2 for a brief description of the HCAHPS survey.) With the transition from the RHQDAPU to the VBP, the HCAHPS measures could become part of the VBP measure set.

Setting Benchmarks and Attainment Thresholds

To provide an empirical basis for designing the Performance Assessment Model and testing the scoring of the HCAHPS survey measures, CMS analyzed HCAHPS data from the National CAHPS Benchmarking Database (NCBD), which represented discharges between December 2005 and September 2006. Of the 1,018 submitting hospitals, the analysis focused on 526 IPPS hospitals (representing 161,141 patients) that had a minimum of 100 completed HCAHPS surveys. Only one year of data was available when this analysis was conducted, which precluded CMS from computing improvement scores.

The potential scoring approach for HCAHPS performance measures used in the approach analyzed in this report captures eight HCAHPS outcome dimensions (seven composites and one

global rating of care) and would seek to incentivize hospitals to improve each of the eight dimensions of patient experience (See Table B-2).

Table B-2: HCAHPS Survey Performance Measures

Eight Measure Dimensions	
Nurse communication	Cleanliness and quiet
Doctor communication	Responsiveness of hospital staff
Pain management	Discharge information
Communication about medications	Overall rating of hospital

The HCAHPS measures could be scored using an approach that parallels the one used to score the clinical process measures, using an attainment score of 0–10 and an improvement score of 0–9, with the final score on each HCAHPS dimension being the higher of the two scores.

However, in contrast to the reporting for clinical process measures, in which different hospitals report on different numbers of measures because of different service mixes, all hospitals supplying HCAHPS data would be expected to report on all eight HCAHPS dimensions.

As with the clinical process measures, attainment thresholds and benchmarks used to judge performance in the current period would be established using data from the previous reporting period. Thus, a hospital’s attainment score would be based on a fixed standard rather than its current standing relative to its peers in the approach analyzed in this report. The attainment threshold for each HCAHPS dimension would correspond to median performance in the baseline period (50th percentile performance). Therefore, hospitals would earn points for attainment if they performed at least as well in the current period as the mid-performing hospital did in the previous period. The benchmark corresponds to excellent performance observed in the baseline period and would be set such that full attainment points are awarded at the 95th percentile performance of the baseline period. The actual benchmarks and attainment thresholds for the VBP Program would be established using the most recent data available at the start of Program implementation.

Similar to the clinical process measures, each of the eight HCAHPS dimensions could be given equal weight in calculating the overall HCAHPS score. The clinical process of care measures assess the rate at which an event occurs, such as aspirin at arrival for heart attack, and are scored

on a 0-1 scale (*e.g.*, a score of 0.76 means that the event occurred 76% of the time). The HCAHPS dimensions are scored on a different scale and were converted to percentiles of adjusted hospital-level scores.¹⁷ This approach would ensure comparability across both process-of-care measures and HCAHPS dimensions being scored in the model.

Scoring HCAHPS Performance Based on Attainment

In the approach analyzed for this report, for each of the eight HCAHPS dimensions, a hospital could earn from 0–10 points for attainment based on where its score for the HCAHPS measure fell relative to the attainment threshold and the benchmark. All attainment points would be rounded up to the nearest whole number, and the number of points awarded would be proportionate to the number of percentiles a given measure is above the baseline median. If the hospital's HCAHPS score on a dimension was:

- Equal to or greater than the benchmark (*i.e.*, the baseline 95th percentile performance), the hospital would receive 10 points for attainment.
- Within the attainment range (greater than the attainment threshold of 50th percentile performance but below the benchmark of 95th percentile performance), the hospital would receive a score of 1–9, based on a discrete linear scale established for the attainment range. For example, if performance on a given measure is at the 60th percentile of baseline (which is 20 percent of the way from the attainment threshold of 50th percentile to the possible 100th percentile), 2 (of 10) attainment points would be awarded.
- Equal to or less than the attainment threshold for the dimension (*i.e.*, 50th percentile performance, which defines the lower bound of the attainment range), the hospital would receive 0 points for attainment.

Scoring HCAHPS Performance Based on Improvement

In the approach analyzed for this report, for each HCAHPS dimension, a hospital could earn from 0-9 improvement points for improving its score on the dimension compared to its prior

¹⁷ The scoring of HCAHPS would adjust for survey mode (phone, mail, Interactive Voice Response or mail with phone follow-up), service line (medical, surgical, maternity), patient characteristics (age, education, self-reported health status, primary language not English), emergency room source, and age by service line interactions to ensure equitable comparisons among hospitals.

period's score.¹⁸ This approach would recognize and encourage all incremental improvement for each of the eight HCAHPS dimensions. A unique improvement range for each hospital on each HCAHPS dimension would be established to define the distance between the hospital's prior period score on a dimension and the national benchmark for the dimension. Improvement points would be awarded proportionately to the positive improvement from the hospital's baseline performance and would be rounded up to the nearest whole number. If the hospital's score was:

- Between its performance in the baseline period and the benchmark, within the improvement range, then the hospital would receive between 0 and 9 improvement points based on the discrete linear scale that defines the improvement range.
- Equal to or lower than its previous period's score on the dimension, the hospital would receive 0 points for improvement.

Thus, for example, if a hospital were to improve from the 10th percentile to the 40th percentile for a given measure, it would achieve $30/90 = 1/3$ of the possible improvement and would be awarded 3 of 9 possible improvement points for that measure. Similarly, if a hospital were to improve from the 20th percentile to the 60th for a measure, it would achieve $40/80 = 1/2$ of the possible improvement and would be awarded 5 (4.5 rounded up) of 9 possible improvement points.

Scoring HCAHPS Performance Based on Achieving Minimum Performance Across All HCAHPS Dimensions

In the approach analyzed for this report, to ensure at least adequate performance across all dimensions, hospitals could also earn points for having all eight dimensions above a minimum threshold. The purpose of the "minimum performance" score is to convey to hospitals that all HCAHPS dimensions should be improved and to provide an incentive to hospitals to bring lagging scores up to at least the attainment threshold. Providing incentives for an entire group of measures is consistent with promoting wider systems changes within hospitals to improve quality, and this approach is consistent with the direction the industry is moving. While the minimum performance score is not currently part of the clinical process measures due to the absence of robust composite measures of performance, a minimum performance score for clinical process measures could be added as the measure set evolves.

¹⁸ If a hospital's performance meets or exceeds the benchmark, no improvement score would need to be calculated.

In the approach analyzed for this report, a total of 20 minimum performance points would be awarded proportionately, based on the *lowest* percentile of eight HCAHPS dimensions, up to the 50th percentile attainment threshold, at which point the hospital would earn all 20 points. Minimum performance points would be rounded up to the nearest whole number (*e.g.*, 9.6 minimum performance points would be rounded up to 10 points). Points would be awarded proportionately to the number of percentiles the lowest dimension is between the 0th and 50th percentile performance.

- If all eight dimensions were at or below the baseline 0th percentile, then 0 minimum performance points would be awarded.
- If the lowest percentile across eight HCAHPS dimensions were the 10th percentile of baseline, then this is 20 percent of the way from the 0th percentile to the 50th, and so 4 minimum performance points (20 percent of 20) would be awarded. If the lowest percentile across eight HCAHPS dimensions were the 25th percentile of baseline, then this is 50 percent of the way, and so 10 minimum performance points (50 percent of 20) would be awarded.
- If all eight HCAHPS dimensions were at or above the baseline median/50th percentile,¹⁹ then all 20 minimum performance points would be awarded. Thus, a hospital would receive all 20 minimum performance points when all eight dimensions were equal to or exceeded the performance of half of the hospitals at the baseline.

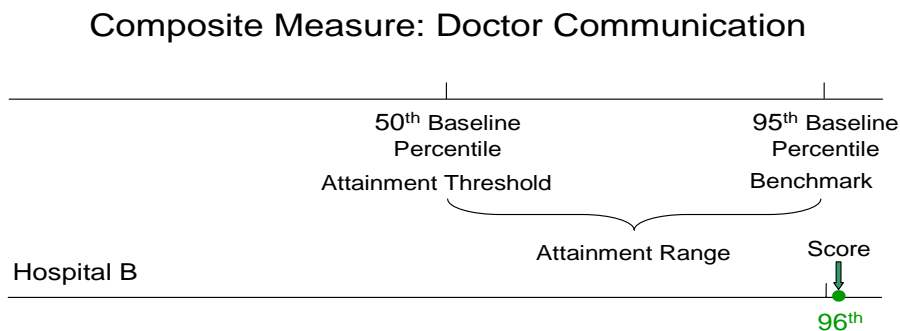
Examples to Illustrate HCAHPS Measure Scoring Model

Examples are presented here to illustrate how the potential Performance Assessment Model would be applied in the context of scoring the HCAHPS dimensions. The hospitals in these examples were selected from the empirical database of 526 IPPS hospitals created from NCBD HCAHPS data used to support model development. The dimension used is doctor communication. Appendix B-3 demonstrates how the model is applied to score four hospitals. Figure 4 shows Hospital B's scoring on the doctor communication dimension. It was placed at the 96th percentile, which exceeded the benchmark. Thus, Hospital B would earn the maximum, 10 points, for attainment. Because this is the highest number of points the hospital could attain

¹⁹ This is actually the 47.5th percentile, rounded.

for this measure, its improvement from its previous period's score on this measure would not be relevant.

Figure B-4: Example of Hospital Earning Points by Exceeding Benchmark, HCAHPS Measure Scoring Model



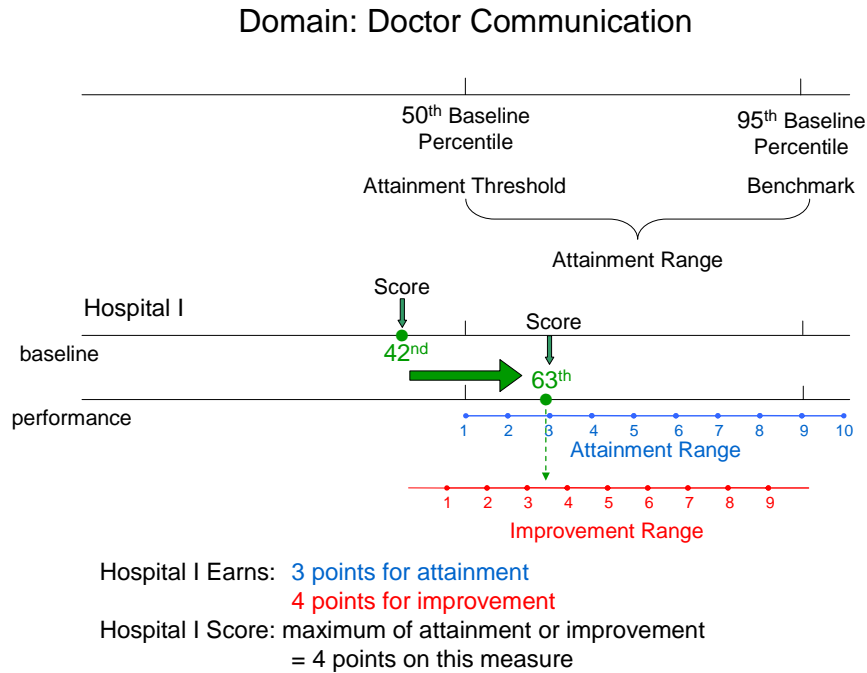
Hospital B's performance in measurement period equates to the 96th percentile in the baseline period

Hospital B Earns: 10 points for attainment for performance exceeding the benchmark

Hospital B Score: = 10 points on this domain

Figure B-5 shows that Hospital I's performance on the doctor communication performance dimension rose from the 42nd percentile to the 63rd percentile. Because the current period's performance exceeds the attainment threshold of the 50th percentile, it would lie in the attainment range. According to the attainment scale, Hospital I would earn 3 points. However, in this case, this period's performance is greater than that of the previous period, so Hospital I would be scored based on improvement as well as attainment. Applying the improvement scale, Hospital I's period-to-period improvement from the 42nd to 63rd percentile would earn it 4 points. Using the greater of the two scores, Hospital I would receive 4 points for this dimension (rounded up to the nearest integer).

Figure B-5: Example of Hospital Earning Points by Attainment or Improvement, HCAHPS Measure Scoring Model



In Figure B-6, Hospital L's performance in the previous period was at the 11th percentile, and its performance declines in the current period to the 6th percentile. Because Hospital L's performance is lower than the attainment threshold of the 50th percentile, it would receive 0 points based on attainment. Hospital L would also receive 0 points for improvement, because its performance is lower than its prior period's performance.

Figure B-6: Example of Hospital Earning Zero Points, HCAHPS Measure Scoring Model

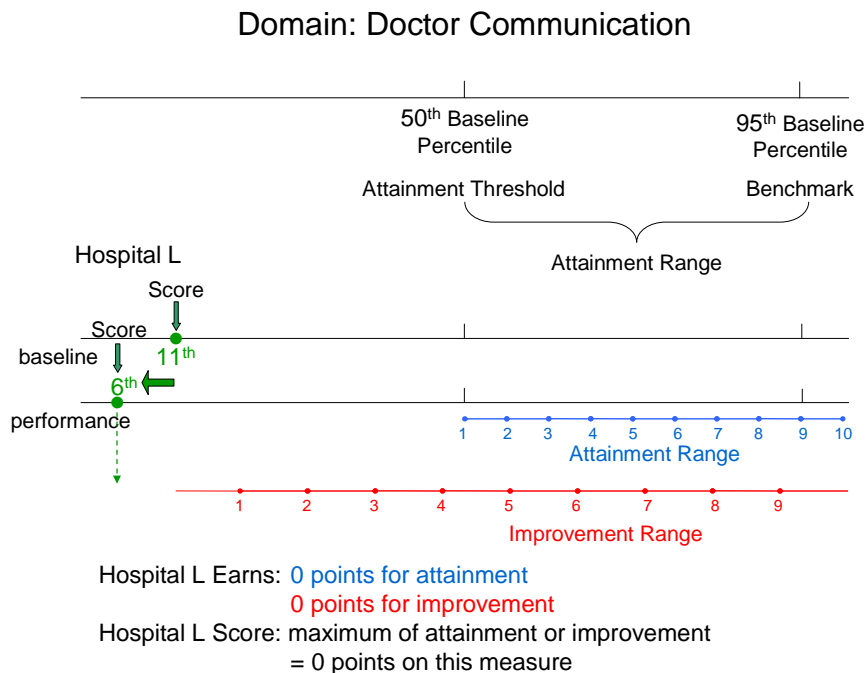
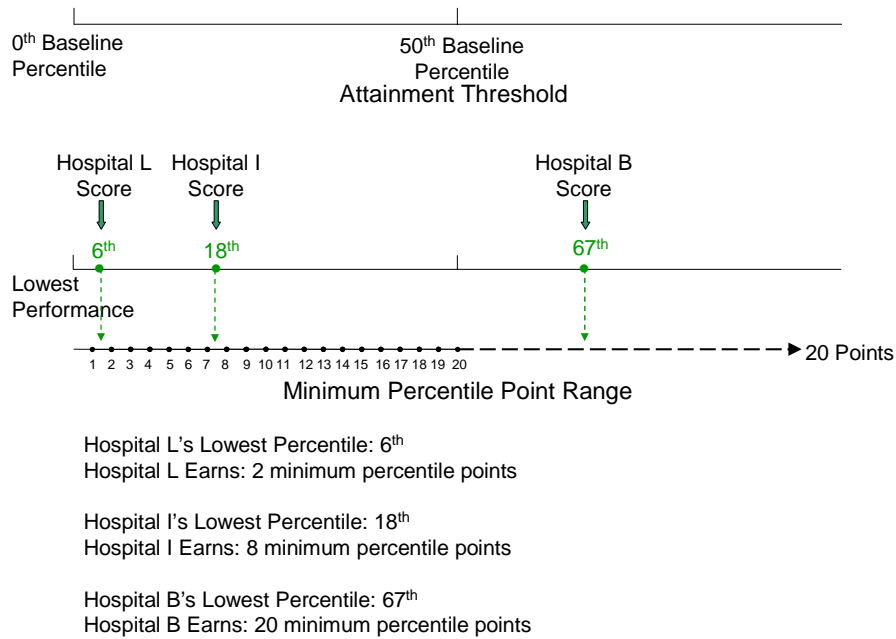


Figure B-7 illustrates the assignment of minimum performance points. Hospital B's minimum performance across the eight HCAHPS dimensions was at the 67th percentile. Because Hospital B's performance on all eight dimensions exceeded the attainment threshold, it would earn all 20 points. Hospital I's minimum performance across all eight dimensions was at the 18th percentile, which is 36 percent of the way from the 0th percentile to the 50th. Thus, Hospital I would earn 36 percent of the minimum performance points, or 7 minimum performance points rounded to the nearest whole number. Hospital L's minimum performance was at the 6th percentile, which is 12 percent of the way from the 0th percentile to the 50th, and so it would earn 2 minimum performance points.

Figure B-7: Examples of Hospitals Earning Points Based on Minimum Performance Across All Eight HCAHPS Dimensions



Calculation of the Overall HCAHPS Performance Score

In the approach analyzed for this report, similar to how clinical process-of-care measures would be scored, the score for a given HCAHPS dimension would be the larger of the attainment points and improvement points for each of the eight measures. Additionally, 0–20 minimum performance points would be awarded. Combined, there would be a total of 100 possible points under the HCAHPS scoring approach. The points earned for each of the HCAHPS dimensions and the minimum performance points would be summed to determine *total earned points*:

Total earned points = Sum of points earned across all dimensions + Minimum performance points earned

Each hospital also would have a corresponding universe of *total possible points* for the HCAHPS measures. The total number of possible points would not differ by hospital, since all participating hospitals would report for all eight HCAHPS dimensions and the 20 minimum performance points, for a total of 100 possible points.

Total possible points = 100

The hospital's *HCAHPS performance score* would be a percentage computed as follows:

$$\text{HCAHPS performance score} = \text{Total earned points} / 100 \times 100\%$$

To further illustrate the application of the potential Performance Assessment Model to HCAHPS measures, Appendix B-3 provides hospital scoring examples.

Creating the VBP Total Performance Score

After scoring each of the broad categories of measures in the VBP Program (*e.g.*, clinical process, HCAHPS), the final scoring step is to combine the different domains to produce a VBP Total Performance Score. Once constructed, this score would be translated into an incentive payment, as described in the following section.

In developing the Performance Assessment Model, CMS considered three options for how to weight the relative contribution of HCAHPS and clinical process scores.

- *Equally by constituent measure:* In this approach, measures are not rolled up to broader domains (clinical process or HCAHPS) before combination. Instead, final scores are a function of the proportion of all possible points achieved based on measure eligibility. This approach appears simple but may result in arbitrary weights between the two domains. For example, in this approach HCAHPS would get more weight in hospitals eligible for fewer clinical process measures (generally the smaller or more specialized hospitals).
- *Unequally by performance domain:* In this approach, measures are rolled up to domains (clinical process or HCAHPS) before combination, with unequal weights for the domains solicited from expert judgment or policy judgments.²⁰
- *Equally by domain:* In this approach, measures are rolled up to domains (clinical process or HCAHPS) before combination, with equal weights applied to the two domains. This approach does not require revision when measures are added.

Several approaches to combining the clinical process scores with the HCAHPS scores on a common set of 504 hospitals were evaluated.²¹ The analysis considered the degree of

²⁰ In practice, it is difficult to achieve expert consensus on exactly what the unequal weights should be.

association between clinical process and HCAHPS scores and the consequences of applying a range of different relative weights.

This analysis led to the conclusion that a performance score should be calculated for each domain, weighting the measures equally within that domain. The scores for the domains would then be combined, with differential weighting applied at the level of the performance domain, based on policy objectives, to determine the VBP Total Performance Score. CMS tested three different weights by performance domain over a range of weights that might be considered by policymakers:

- 60 percent clinical process + 40 percent HCAHPS.
- 70 percent clinical process + 30 percent HCAHPS.
- 80 percent clinical process + 20 percent HCAHPS.

Appendix B-4 provides an example of how the clinical process performance score and HCAHPS performance score are combined into the total performance score using the weighting 70 percent clinical process and 30 percent HCAHPS for four hospitals.

The analysis found an insensitivity to where weights are set within a reasonable range of values, which provides policymakers latitude in setting weights. It also found that the two performance dimensions are only weakly associated and thus appear to assess largely different dimensions of quality.

At VBP Program start, CMS proposes that the actual weighting among domains depend upon empirical analysis conducted as part of implementation planning to permit inclusion of the “clinical outcome” domain initially represented by the two 30-day mortality measures. The weights assigned to new domains would consider the influence the domain has on the variance of the combined score, among other factors.

²¹ The approach to combining different domains for constructing an overall performance score was developed using 2004 and 2005 clinical process-of-care measure data collected through the RHQDAPU Program and 2007 HCAHPS data from 504 hospitals with at least 100 responses to the survey (from the NCBD). At the time of this analysis, neither 30-day mortality data nor HCAHPS improvement scores were available for inclusion in the analysis.

Approach to Scoring 30-Day Mortality Measures

The RHQDAPU Program currently includes two 30-day mortality measures: (1) acute myocardial infarction (AMI) mortality and (2) heart failure (HF) mortality. CMS posted hospital performance on mortality measures on the HHS Hospital Compare website beginning June 2007. Data for the two mortality measures were not yet available for modeling in the context of creating the Performance Assessment Model. CMS will obtain and analyze the mortality data over the course of the next year to determine how these measures could be integrated into the Performance Assessment Model. CMS will also conduct analyses to establish appropriate benchmarks and attainment thresholds required to score these types of measures for a VBP financial incentive.

Appendix B-1: Example of Scoring Hospital Performance on Clinical Measures

Measure	Benchmark*	Attainment Threshold*	Hospital A					Hospital B					Hospital I					Hospital L				
			Performance*		Points			Performance*		Points			Performance*		Points			Performance*		Points		
			Baseline	Current	Attainment	Improvement	Awarded	Baseline	Current	Attainment	Improvement	Awarded	Baseline	Current	Attainment	Improvement	Awarded	Baseline	Current	Attainment	Improvement	Awarded
^Patients Given Aspirin at Arrival-AMI	.90	.60	.77	.75	5	0	5	.99	.99	10	na	10	.40	.78	6	7	7	.78	.60	0	0	0
^Patients Given ACE Inhibitor for LVSD-AMI	.90	.60	.78	.77	6	0	6	.67	1.0	10	na	10	Too few patients to include in calculation of performance score					.33	.38	0	1	1
^Smoking Cessation-HF	.90	.60	.81	.73	4	0	4	.65	.99	10	na	10	.11	.75	5	8	8	.60	.62	1	1	1
Patients Assessed and Given Pneumococcal Vaccination-PN	.87	.47	.69	.67	6	0	6	.88	.91	10	na	10	.21	.70	6	7	7	.57	.46	0	0	0
Surgery Patients who Received Antibiotic(s) one Hour Before Incision-Surgery	.96	.77	.96	.92	8	0	8	.96	.97	10	na	10	.25	.85	4	8	8	.81	.79	1	0	1
Potential Points			50					50					40					50				
Points Earned			29					50					30					3				
Performance Score			58%					100%					75%					6%				

*Benchmarks, Attainment Thresholds and Baseline Performance based on 2004 Hospital Compare Data; "Current" Performance based on 2005 Hospital Compare Data. ^Denotes measures that are topped out.

Appendix B-2: CAHPS[®] Hospital Survey (HCAHPS) Fact Sheet

Overview

The intent of the CAHPS^{®22} Hospital Survey, also known as Hospital CAHPS[®] or HCAHPS, is to provide a standardized survey instrument and data collection methodology for measuring patients' perspectives of hospital care.

Three broad goals have shaped the HCAHPS survey. First, the survey is designed to produce comparable data on patients' perspectives of care that allows objective and meaningful comparisons among hospitals on domains that are important to consumers. Second, public reporting of the survey results is designed to create incentives for hospitals to improve quality of care. Third, public reporting will serve to enhance public accountability in health care by increasing the transparency of the quality of hospital care provided in return for the public investment. With these goals in mind, the HCAHPS project has taken substantial steps to assure that the survey will be credible, useful, and practical. This methodology and the information it generates will be made available to the public.

HCAHPS Development

The Centers for Medicare & Medicaid Services (CMS) partnered with the Agency for Healthcare Research and Quality (AHRQ), another Agency in the Department of Health and Human Services, to develop HCAHPS. The HCAHPS survey is composed of 27 items: 18 substantive items that encompass critical aspects of the hospital experience (communication with doctors, communication with nurses, responsiveness of hospital staff, cleanliness and quietness of hospital environment, pain management, communication about medicines, discharge information, overall rating of hospital, and recommendation of hospital); four items to skip patients to appropriate questions; three items to adjust for the mix of patients across hospitals; and two items to support congressionally-mandated reports.

²² CAHPS[®] (Consumer Assessment of Healthcare Providers and Systems) is a registered trademark of the Agency for Healthcare Research and Quality, a U.S. Government Agency.

In May 2005, the HCAHPS survey was endorsed by the National Quality Forum (NQF), which represents the consensus of many healthcare providers, consumer groups, professional associations, purchasers, federal agencies, and research and quality organizations. In December 2005, the federal Office of Management and Budget gave its final approval for the national implementation of HCAHPS for public reporting purposes. The HCAHPS has also endorsed HCAHPS. More information about the development and background of HCAHPS can be found at: www.cms.hhs.gov/HospitalQualityInits/.

National Implementation

Voluntary collection of HCAHPS data for public reporting began in October 2006. The first public reporting of HCAHPS results, which will encompass eligible discharges from October 2006 through June 2007, is slated for early 2008. Hospitals are required to submit HCAHPS data under the RHQDAPU Program beginning with July 2007 discharges. HCAHPS results will be posted on the Hospital Compare website, found at www.hospitalcompare.hhs.gov, or through a link on www.medicare.gov.

**Appendix B-3: Example of Scoring Hospital Performance on HCAHPS Measures
Baseline and Current Percentile for Four Hospitals, A, B, I and L**

HCAHPS Dimension	Hospital A					Hospital B					Hospital I					Hospital L				
	Percentile of Baseline			Improvement		Percentile of Baseline			Improvement		Percentile of Baseline			Improvement		Percentile of Baseline			Improvement	
	Baseline	Current	Points Above 50th	Actual	Possible	Baseline	Current	Points Above 50th	Actual	Possible	Baseline	Current	Points Above 50th	Actual	Possible	Baseline	Current	Points Above 50th	Actual	Possible
RN Comm	75	81	31	6	25	88	89	39	1	12	13	19	0	6	87	3	10	0	7	97
MD Comm	63	64	14	1	37	81	96	46	15	19	42	63	13	21	58	11	6	0	0	89
Clean/Quiet	74	67	17	0	26	59	75	25	16	41	85	90	40	5	15	28	54	4	26	72
Staff Respon	64	76	26	12	36	93	91	41	0	7	38	64	14	26	62	14	33	0	19	86
Comm re: New Meds	57	63	13	6	43	71	81	31	10	29	5	23	0	18	95	59	70	20	11	41
Pain	65	63	13	0	35	96	98	48	2	4	27	26	0	0	73	14	45	0	31	86
Discharge Info	50	63	13	13	50	59	67	17	8	41	14	18	0	4	86	2	14	0	12	98
Hosp Rating	74	72	22	0	26	57	79	29	22	43	17	33	0	16	83	6	42	0	36	94
Lowest baseline percentile (current)	63					67					18					6				

*Benchmarks, Attainment Thresholds, and Baseline Performance based on 2007 NCBD Data from 526 hospitals with at least 100 completed surveys. Current performance data are based on simulated improvement.

Total HCAHPS Points Awarded for Four Hospitals

HCAHPS Dimension	Hospital A				Hospital B				Hospital I				Hospital L			
	Attainment Points	Improvement Fraction	Improvement Points	Measure Points	Attainment Points	Improvement Fraction	Improvement Points	Measure Points	Attainment Points	Improvement Fraction	Improvement Points	Measure Points	Attainment Points	Improvement Fraction	Improvement Points	Measure Points
RN Comm	7	24%	3	7	8	8%	1	8	0	7%	1	1	0	7%	1	1
MD Comm	3	3%	1	3	10	79%	8	10	3	36%	4	4	0	0%	0	0
Clean/Quiet	4	0%	0	4	5	39%	4	5	8	33%	3	8	1	36%	4	4
Staff Respon	6	33%	3	6	9	0%	0	9	3	42%	4	4	0	22%	2	2
Comm re: New Meds	3	14%	2	3	7	34%	4	7	0	19%	2	2	4	27%	3	4
Pain	3	0%	0	3	10	50%	5	10	0	0%	0	0	0	36%	4	4
Discharge Info	3	26%	3	3	4	20%	2	4	0	5%	1	1	0	12%	2	2
Hosp Rating	5	0%	0	5	6	50%	5	6	0	19%	2	2	0	38%	4	4
Minimum percentile score (0-20)				20				20				8				3
Total points (0-100)				54				79				30				24

Appendix B-4: Example of Combined Clinical and HCAHPS Performance Assessment Score

	Hospital A	Hospital B	Hospital I	Hospital L
Performance Score on Process Measures (PSPM)	58%	100%	75%	6%
Performance Score on HCAHPS (PSH)	54%	79%	30%	24%
Total Performance Score (OVP) (.7*PSPM) + (.3*PSH)	57%	94%	62%	11%

Appendix C: Examples of Modifying DRG Payment with VBP Incentive Payment

Here we provide examples to demonstrate how the VBP incentive payment would be calculated for DRG 498 based on hospitals' Total Performance Scores using the approach analyzed for this report. For the purpose of this example, the VBP incentive payment amount is set at 5 percent of the base operating DRG payment, clinical process measures are weighted at 0.7 and HCAHPS is weighted at 0.3 in calculating the Total Performance Scores, and we assume that Hospital B and Hospital A both have a wage index of 1. The hospital-specific incentive payment earned by each hospital is taken from the Figure 1 table in the VBP Plan. Hospital B earns 100 percent of the incentive payment; as a result, its VBP payment for DRG 498 is the same as the base operating payment for the DRG. Hospital A earns 82 percent of the incentive payment; as a result, its VBP payment for DRG 498 is \$132.42 less than the base operating payment for the DRG.

Table C-1: Example for DRG 498

	<u>Hospital B</u>	<u>Hospital A</u>
Payment for DRG 498	\$14,713.85	\$14,713.85
At-Risk VBP Portion of DRG Payment (5%)	\$735.69	\$735.69
% of VBP Incentive Payment Earned	100.0%	82.0%
Hospital-Specific Earned VBP Portion	\$735.69	\$603.27
VBP Payment for DRG 498	\$14,713.85	\$14,581.43

Note on the calculation of non-VBP related payment amounts:

In the approach analyzed in this report, the calculated payments for the non-VBP related portion of the IPPS payment (IME, DSH, outliers, etc.) would remain unchanged and would be calculated using the full base DRG; the financial incentive withhold would not apply in these calculations.

Distribution of unearned incentive dollars:

As shown in the example of Hospital A, not all hospitals would earn the full VBP incentive payment, creating a pool of funds that could be in whole or in part redistributed to hospitals as additional incentive. For example, if the pool to be redistributed is large enough that all hospitals earning the financial incentive would have their incentive increased by 15 percent, Hospital B would earn 115.0 percent of the financial incentive, while Hospital A would earn 94.3 percent of the financial incentive. Thus Hospital B's modified VBP DRG payment would be \$14,824.20, while Hospital A's modified VBP DRG payment would be \$14,671.92.

Appendix D: Measure Development, Selection, and Modification

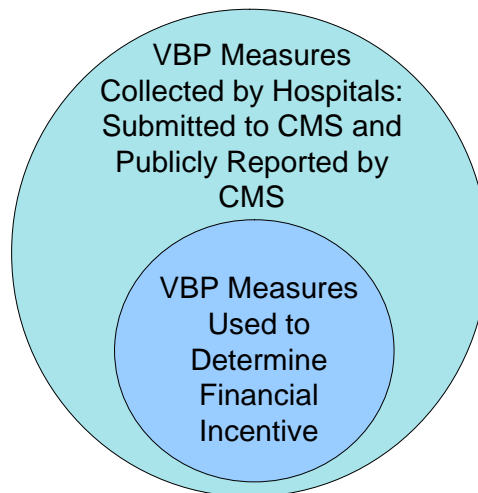
Appendix D provides details on:

- The measure development and selection process
- Potential VBP measures at program start
- Additional measures beyond VBP Year 1
- The challenge of small numbers

The Measure Development and Selection Process

Under the VBP Program, financial incentives for better performance would be coupled with public reporting as tools to drive improvements in clinical quality, patient centeredness, and efficiency. Thus, VBP measures would be used both for public reporting and for determining incentive payments. However, as Figure D-1 illustrates, only a subset of the publicly reported measures would be included for incentive payment purposes.

Figure D-1: Universe of VBP Measures



There are a number of reasons why the measures used for the VBP financial incentive would be only a subset of those used for public reporting. As shown in Figure 2 of the Plan, a staged approach for introducing measures is proposed in order to:

- Determine an empirical baseline performance on which to base future incentives,
- Refine processes for data collection, submission, validation, and formatting,
- Maintain the policy option to select a subset of publicly reported measures for payment incentives in order to focus improvement on particular areas (*e.g.*, to reinforce Quality Improvement Organization [QIO] or other national priorities) without constraining the development of a menu of measures for public reporting, and
- Decrease the possibility of unintended consequences that might arise as a result of financial incentives being tied to actual performance values, wherever those consequences may be less likely via simply public reporting of values.

In the *first* stage, measures meeting the basic criteria of an available data-collecting mechanism and consensus endorsement would undergo a “preliminary data submission period” without public reporting or the application of performance-based incentives. This period would allow submitted data to accumulate and hospitals to become familiar with the measure specifications and data submission requirements. Hospitals would be required to submit data to CMS during the preliminary data submission period to be eligible for the VBP financial incentive. Once sufficient data had been accumulated and any measures or data collection issues resolved, the measures would move to the second stage.

In the *second* stage, measures would be publicly reported and evaluated for their suitability for VBP performance-based incentive payment. Measures that fail to meet the VBP inclusion criteria for incentive payment could be retained for public reporting if they continue to have value for that purpose. The criteria listed below could be used in evaluating the suitability of measures for VBP incentive payment. This list builds on the set of criteria used by the National Quality Forum (NQF) for public reporting and quality improvement. Over time, the criteria may be augmented or modified based on experience with the VBP Program.

- Importance,
- Scientific acceptability,
- Feasibility,
- Usability,
- Improvability,
- Controllability,

- Potential for unintended consequences, and
- Contribution to comprehensiveness.

In the *third* stage, measures that meet the criteria for the VBP financial incentive would be expected to progress from public reporting to inclusion in the payment incentive. The data collected in stage 2 would be used to determine each new measure's national benchmark and attainment threshold and would establish each hospital's baseline score for determining improvement over time.

CMS anticipates that some measures introduced into VBP through the process may be temporarily withdrawn or permanently retired from the VBP Program for a variety of reasons, including changes in scientific evidence or policy objectives. Withdrawing or retiring VBP measures would be based on application of the same criteria used to screen measures for inclusion. Similarly, CMS may determine that a measure should no longer be used for the financial incentive even though performance data for that measure could continue to be publicly reported. In this case, hospitals would still be required to report data on the measure in order to be eligible for VBP financial incentives.

In summary, measures included for the financial incentive would be expected to meet a stringent set of criteria and would be continually evaluated for their appropriateness. A subset of measures would continue to be reported by hospitals to support the development of new measures and for public reporting.

VBP Measures at Program Start

To illustrate how the measures selected for the incentive payment are likely to be more limited than are those for public reporting alone, the potential VBP selection criteria were applied to the 20 process-of-care measures for which CMS had 2004–2005 data from Hospital Compare and two additional measures for which data were not yet available: (1) prophylactic antibiotic selection for surgical patients (SCIP-2) and (2) influenza vaccination status for pneumonia patients. All of these measures, except for SCIP-2, were included in the FY 2007 RHQDAPU Program.

Based on application of the selection criteria, only some process-of-care measures in the FY 2007 RHQDAPU measure set would be included as initial VBP program financial incentive measures. Several measures would not be recommended for inclusion in the incentive measure set for a number of different reasons. For example, the performance rates for one of the measures, “oxygenation assessment for pneumonia (PN-1),” have topped out²³ so completely across virtually every hospital that there is little opportunity for improvement. Therefore, this measure does not meet the improvability criterion and hence does not require financial incentives for improvement. Nonetheless, CMS is including this measure among those required for public reporting for informational purposes and to sustain performance. Three other measures not recommended for inclusion in the financial incentive at this time are:

- Beta blocker at arrival for acute myocardial infarction (AMI),
- Left ventricular ejection fraction (LVEF) assessment for heart failure,
- Initial antibiotic received within 4 hours of arrival for pneumonia, and

These measures either: (1) are under active consideration for possible specification changes, (2) overlap with another measure included for the financial incentive, or (3) have complex or changing clinical guidelines.

These examples illustrate that performance measure specifications are dynamic. This is particularly so for process-of-care measures, which depend on scientific evidence that is subject to change. Accordingly, measures that may not be currently recommended for performance-based financial incentives may later have a different recommendation based on re-specifications or other changes. Alternatively, it may be determined that currently recommended measures should be withdrawn or retired completely from the VBP Program.

Based on hospital scores for 2005 (Appendix D-1), six of the recommended measures have been designated as topped-out measures. It is anticipated that as hospitals’ performance on the RHQDAPU measures continues to improve, more of the potential measures could be designated as topped-out prior to any implementation of a Hospital VBP Program.

²³ A measure is considered topped out if the performance of reporting hospitals is very close to a perfect score, such that the value for the 75th percentile is not statistically different from the value for the 90th percentile.

In addition to the process-of-care measures in the FY 2007 RHQDAPU measure set, HCAHPS, AMI 30-day mortality, and HF 30-day mortality from the FY 2008 RHQDAPU measure set could be included for the financial incentive. Two new FY 2008 process-of-care measures could also be considered for inclusion:

- SCIP-VTE 1: Venous thromboembolism (VTE) prophylaxis ordered for surgery patient.
- SCIP-VTE 2: Venous thromboembolism (VTE) prophylaxis within 24 hours pre/post surgery.

Also considered would be SCIP-2: Prophylactic antibiotic selection for surgical patients because it was adopted for inclusion in the FY 2008 RHQDAPU measure set. SCIP-2 has previously been a measure voluntarily reported by some hospitals but not included in the measure set considered for the Annual Payment update because of changing measure specifications.

With the addition of measures of patients' perspectives of care, outcome measures, and surgical process-of-care measures (which are directly related to complications and hence to patient safety), CMS is actively taking steps to broaden the measures portfolio that could serve as the basis of the proposed initial VBP Program.

Beyond these candidate measures for VBP incentive payment, the measure set for public reporting would also include hospital outpatient measures currently under development, as required by Section 109 of the Tax Relief and Health Care Act of 2006.

In the FY 2008 IPPS Final Rule that appeared in the Federal Register on August 22, 2007, CMS added the Pneumonia—30-day mortality (Medicare patients) measure for the FY 2009 APU determination. Two new inpatient measures were also added in the CY 2008 OPSS final rule issued November 1, 2007: SCIP-Inf-4 Cardiac Surgery Patients with Controlled 6 A.M. Postoperative Blood Glucose and SCIP-Inf-6 Surgery Patients with Appropriate Hair Removal. New measures, such as these, would initially be used only for public reporting purposes in the VBP Program until CMS could apply the potential measure selection criteria for the VBP financial incentive. If the measures were deemed suitable, CMS would migrate them into the incentive payment portion of the Program, but only after sufficient baseline data for determining thresholds and benchmarks had been collected.

Additional Measures Beyond Year 1

The VBP Program is predicated on the assumption that it will build on the accomplishments of the RHQDAPU Program and its associated infrastructure. Some of the FY 2007 RHQDAPU measures provide a solid foundation for the initial VBP financial incentive measure set. These measures of clinically effective processes-of-care have high credibility within the hospital community and are well recognized as meeting the potential inclusion criteria discussed earlier in this Appendix for VBP financial incentive measures.

CMS recognizes the need for measurement to evolve rapidly beyond the current scope in order to ensure comprehensive measurement of hospital quality under the VBP Program. CMS agrees with the IOM recommendation that the initial focus of the VBP Program should be measures of clinical quality, patient centeredness, and efficiency:

- *Clinical Quality:* CMS would seek particularly to expand outcome measures and measures of patient safety. Such measures are more complex to develop than are process measures, because they require both risk adjustment to control for differences in the patient illness levels across hospitals and greater attention to the potential for unintended consequences, such as decreasing access to care for patients with greater outcome risk. However, such measures are potentially more stable over time, as desirable outcomes are less likely to change than are processes of care. Outcome measures also are less provider-specific and thus tend to recognize the continuum of care that extends within and beyond the hospital, decrease fragmentation, and promote coordination of care across settings when applied broadly.
- *Patient-Centered Care:* The inclusion of HCAHPS in the FY 2008 RHQDAPU Program advances the measurement of patient-centered care significantly, and CMS anticipates that this composite measure set is mature enough to be included in a VBP Program financial incentive.
- *Efficiency:* The importance of efficiency measures is widely recognized, but such measures pose major development challenges. Measurement of the use of services alone does not constitute a measure of efficiency. By expanding outcome measures as

previously discussed, there will be a greater opportunity to relate use of services to outcomes forming the basis for measurements of efficiency.

Recognizing the need to evolve the measure set over time, CMS could develop and introduce the additional potential measure categories shown in Table D-1. Broadening and deepening the set of measures would enhance CMS’ ability to achieve its VBP programmatic goals. CMS could also establish a systematic approach that identifies and prioritizes areas for longer-term measure development. CMS is developing new measures that will promote coordination of care across settings, patient centeredness, longitudinal assessment, and shared provider accountability, in addition to clinical quality.

Table D-1: Potential Evolution of VBP Measures over Time

FY 2010–FY 2011	FY 2012 and Beyond
<ul style="list-style-type: none"> • Outcome measures • Patient safety measures • Care coordination measures • Emergency care measures • Efficiency measures 	<ul style="list-style-type: none"> • Performance areas where gaps are identified

CMS understands that introducing new measures into VBP beyond the initial years will require the development and/or thorough evaluation of additional measures, and that the development, review, and maintenance of measures are ongoing programmatic functions. Because the performance measures provide the foundation for a VBP Program, it is essential to fill the measure pipeline for future application to the VBP Program.

CMS is committed to working with consensus organizations and other stakeholders to identify gaps in measures, prioritize areas for measure development, and identify existing measures for inclusion in VBP. All newly developed measures would be tested prior to introduction into VBP. New measures also would be submitted to NQF for endorsement. CMS would continue to collaborate with the Joint Commission to align measure specifications for those measures that CMS has in common with the Joint Commission and would maintain measures such that

specifications are consistent with scientific evidence and coding systems. Updates to measure specifications and the technical manual would continue to be maintained in the public domain.

Small Numbers on Individual Performance Measures

Under the current RHQDAPU Program, many hospitals report a small number of cases in the measure denominator for one or more of the clinical process measures that could be used in the VBP financial incentive. Small numbers on individual measures occur for a variety of reasons, including low patient volume, the use of sampling rather than a census of all cases (to minimize hospital abstraction burden), and the use of discretionary exclusions for patients otherwise eligible for a measure. Very low numbers of cases can provide only limited approximations of the true, underlying performance of a hospital.

For hospitals reporting a small number of cases for a given performance measure, the resulting measure rate could vary substantially from period to period. If performance is measured with error, such as sampling error from insufficient sample sizes, hospitals would be exposed to a combination of valid and random components to their incentive payment. If an incentive scheme has substantial measurement error, hospitals will experience large variations in their scores and incentive payments over time, even when their own behavior does not change. Such a situation would likely undermine support for the Hospital VBP Program. A related problem occurs in publicly reported scores/rank-based measures when sample sizes vary substantially by institution: the smallest institutions are more likely to be extreme by chance alone, beyond any true differences between smaller and larger hospitals.

CMS has explored several approaches to increase measure reliability and thereby address the problem of small numbers:

- Composite measures that combine information across related performance measures within the same hospital. Composite measures combine individual measures according to selected topics or themes, such as specific conditions, patient safety, or patient experience (*e.g.*, HCAHPS). In addition to other advantages, composite measures can bring more stability to quality measurement for hospitals with small case volumes. Composite measures resulting from these techniques can garner information from available individual measures, even though small case

volumes give rise to significant measurement error and missing values for many individual measures.

- Rolling up data within the same hospital over longer time periods. Some hospitals serve few patients overall or few patients for a given condition. Using a longer time period would acknowledge that discerning the quality of performance for such hospitals requires more data than can be obtained within a 12-month period and would allow longer periods for sufficient data to accumulate. For example, aggregating data over two years would effectively double the number of observations.
- The use of shrinkage methods, which incorporate information across hospitals.²⁴ There are established and widely used statistical methods for adjusting observed or raw scores by blending them with averages or expectations borrowed from other entities, in this context, other hospitals. However, this method conflicts with the policy goals of VBP to provide reliable public reporting and financial incentives based on a hospital's individual performance. In this context, this method may lead to misleading information, replacing either low or high scores for hospitals with scores closer to the average.

CMS would prefer the first two of these approaches because they are more transparent to hospitals and use only a hospital's own performance data to improve reliability. Even using these approaches, however, a small number of hospitals may still not have sufficient numbers of cases for reliable scoring of the selected process-of-care measures. In these instances, CMS would need to consider the possibility of a modified process for determining the financial incentive.

²⁴ Elliott MN, Zaslavsky AM, Cleary PD (2006). "Are Finite Population Corrections Appropriate when Profiling Institutions?" *Health Services and Outcomes Research Methodology*, 6(304): 153–156.

Appendix D-1: 2005 Scores on Clinical Process-of-Care Measures

All Hospitals Reporting in Both 2004 and 2005

Condition	Measure Name	n*	5th pcntl	10th pcntl	25th pcntl	Median	75th pcntl	90th pcntl	95th pcntl
Heart Attack	Patients Given ACE Inhibitor or ARB for Left Ventricular Systolic Dysfunction (LVSD)	2781	0.50	0.57	0.74	0.85	0.95	1.00	1.00
	Patients Given Adult Smoking Cessation Advice/Counseling	2200	0.34	0.50	0.81	0.94	1.00	1.00	1.00
	Patients Given Aspirin at Arrival	3204	0.75	0.84	0.91	0.96	0.98	1.00	1.00
	Patients Given Aspirin at Discharge	3137	0.63	0.73	0.87	0.95	0.99	1.00	1.00
	Patients Given Beta Blocker at Arrival	3198	0.59	0.71	0.84	0.92	0.97	1.00	1.00
	Patients Given Beta Blocker at Discharge	3146	0.60	0.71	0.86	0.94	0.98	1.00	1.00
	Patients Given PCI Within 120 Minutes Of Arrival	1092	0.29	0.40	0.56	0.69	0.80	0.88	0.91
	Patients Given Thrombolytic Medication Within 30 Minutes Of Arrival	1107	0.00	0.00	0.00	0.33	0.50	0.73	1.00
	Heart Failure	Patients Given ACE Inhibitor or ARB for Left Ventricular Systolic Dysfunction (LVSD)	3218	0.58	0.65	0.75	0.83	0.91	0.97
Patients Given Adult Smoking Cessation Advice/Counseling		2688	0.38	0.50	0.71	0.86	0.96	1.00	1.00
Patients Given Assessment of Left Ventricular Function (LVF)		3305	0.49	0.66	0.81	0.90	0.96	0.98	0.99
Patients Given Discharge Instructions		2705	0.10	0.20	0.40	0.59	0.77	0.89	0.95
Pneumonia		Patients Assessed and Given Pneumococcal Vaccination	3287	0.13	0.23	0.44	0.62	0.77	0.88
	Patients Given Adult Smoking Cessation Advice/Counseling	2686	0.38	0.48	0.65	0.82	0.93	0.99	1.00
	Patients Given Initial Antibiotic(s) within 4 Hours After Arrival	3290	0.53	0.61	0.69	0.77	0.85	0.90	0.92
	Patients Given Oxygenation Assessment	3302	0.95	0.97	0.99	1.00	1.00	1.00	1.00
	Patients Given the Most Appropriate Initial Antibiotic(s)	2661	0.61	0.67	0.76	0.82	0.86	0.90	0.92
	Patients Having a Blood Culture Performed Prior to First Antibiotic Received in Hospital	2723	0.67	0.72	0.78	0.84	0.89	0.92	0.94
	Surgical Infection Prevention	Surgery Patients Who Received Preventative Antibiotic(s) One Hour Before Incision	644	0.45	0.59	0.74	0.84	0.91	0.94
Surgery Patients Whose Preventative Antibiotic(s) Are Stopped Within 24 Hours After Surgery		601	0.29	0.42	0.58	0.73	0.84	0.91	0.95

* Number of hospitals reporting out of 3,518 hospitals. A subset of these measures were not in RHQDAPU but were voluntarily reported in 2004 and 2005.

Appendix E: Enhancements to Data Infrastructure and Measure Validation

Appendix E presents further details on:

- Enhanced user support as part of the improved data infrastructure for the VBP Program
- The potential measure validation process
- Exclusion of data on Hospital Compare for hospitals failing validation

Enhanced User Support

To support the VBP Program, CMS could expand its existing data submission infrastructure to provide full-level user support during business hours in all time zones (7 a.m. Eastern to 7 p.m. Hawaiian). The expanded CMS user support would provide a single source of consistent, accurate, and timely information in response to specific abstraction, measures, and submission questions from hospitals and data vendors. The user support would allow hospitals to submit all administrative information directly to a single source. The expanded infrastructure could replace much of the hands-on assistance currently performed by Quality Improvement Organizations (QIOs) to assist hospitals in abstracting, submitting, and publicly reporting hospital quality data to CMS. The types of assistance provided could include the following:

- Helping support contractors respond to specific abstraction and measure definition questions from hospitals;
- Collecting and processing administrative forms from hospitals (*e.g.*, HQA participation forms);
- Calling hospitals to remind them to submit a complete set of data to comply with sampling requirements; and
- Educating hospitals to interpret QIO clinical warehouse and QualityNet submission and validation reports.

Proposed Measure Validation Process

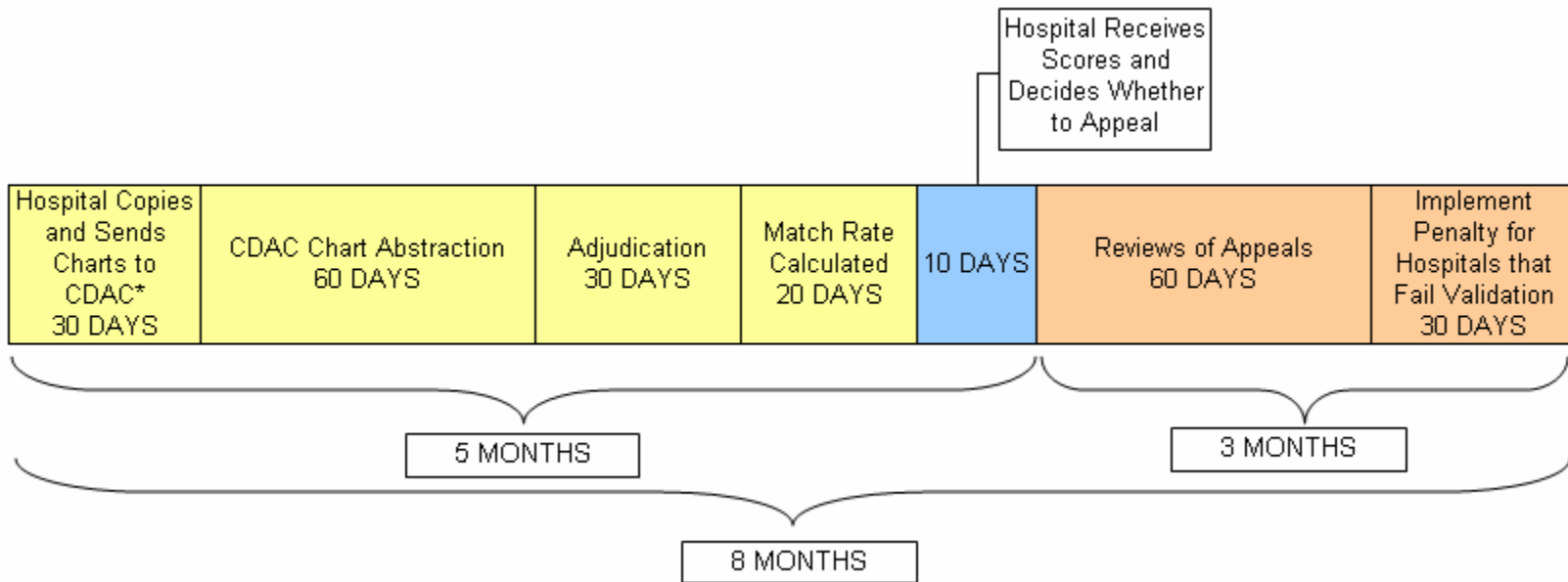
Under the RHQDAPU Program, CMS conducts validation at the data-element level for 20 randomly selected charts each year for every participating hospital, which is too small a number to assess the accuracy of the performance measure rates. CMS considers the validation process used by the RHQDAPU Process insufficient for the VBP Program. The RHQDAPU data element match-based approach assesses abstraction accuracy of the elements but does not assess the impact of inaccuracy for an element as it is combined with other elements to calculate the measure rate used for public reporting and VBP incentive payment. When scoring performance and tying payment to it, CMS must assure that the rates are correct, not just the data elements used to calculate these rates. The existing validation methodology could be revised to focus on assessing the accuracy of performance measure rates. For Year 1 of the proposed VBP Program, the RHQDAPU validation process would continue in order to allow CMS to transition to the new approach. However, starting with VBP Year 2, the following potential audit strategy could be used:

- *Select hospitals for validation on both a “random” and “targeted” basis:* For each hospital selected for validation, CMS would review approximately 50 charts per year. The dual audit selection strategy for auditing would serve two functions: (1) enable CMS to assess more completely the overall quality of data submissions and (2) minimize gaming.
 - Random audit: The annual random audit would include approximately 600 hospitals. In any given year, each hospital would have roughly a 1-in-5 chance of being reviewed by random audit.
 - Targeted audit: A hospital would be selected for targeted audit if it had unusual data patterns, such as an abnormally high rate of exclusions; if it had submitted data that would likely result in an unusually high incentive payment; or if it had previously failed data validation. Approximately 200 hospitals would be selected annually for the targeted audit.
- *Focus the validation on the accuracy of abstraction to calculate measure rates:* The current RHQDAPU approach focuses on accuracy of the many individual elements used to construct a measure rather than determining the accuracy of a measure rate. Requiring

accuracy at the measure level is a stricter definition. The accuracy of the denominator status (included versus excluded cases) would also be assessed. The pass/fail threshold for the validation score would be determined from empirical analyses, but is expected to be approximately a 60-70 percent accuracy rate for both random-audit and targeted-audit hospitals. Appendix E-1 presents an example of this scoring.

- *Conduct the validation and appeals process post-payment:* This approach would avoid having the validation process delay incentive payment determinations and public reporting. Figure E-1 presents a detailed timeline for data submission for the validation and appeal processes.

Figure E-1: Details of Validation and Appeals Timeline



***Clinical Data Abstraction Center (CDAC)** is the contractor used by CMS to carry out the process of validating data collected from medical records for the RHQDAPU Program and would continue to be used for the VBP Program.

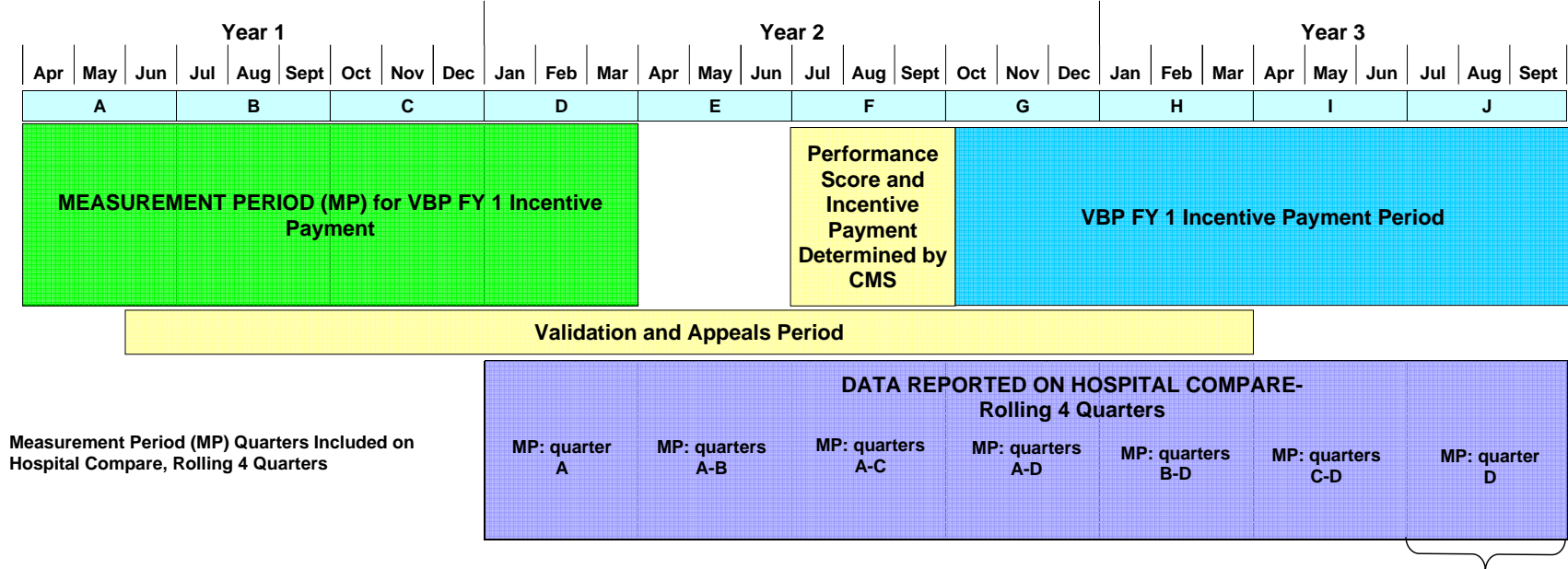
Exclusion of Data on Hospital Compare for Hospitals Failing Validation

Because CMS is proposing to conduct the validation and appeals process post hoc in order to more closely link the measurement year to the incentive payment determination and reporting year, the timetable for having the validation work completed and for knowing whether a hospital failed would occur after the majority of the “failed” performance data would have been posted to the Hospital Compare website. (See Figure E-2) Starting with the January update of Hospital Compare, first one, then the second and third quarters of the measurement year would become part of the rolling four quarters of data that are posted on Hospital Compare. The October posting includes all four quarters of the measurement year, while validation would not be completed until April of the following year.

As soon as CMS becomes aware that a hospital has failed data validation, it would move to exclude affected quarters of data from Hospital Compare. Given the proposed timetable for submissions, validation, and appeals, CMS would only be able to exclude one quarter of data, which would occur for the July update of Hospital Compare. This posting would bear a footnote explaining that the hospital had only three quarters of data included because one quarter failed validation.

Because of the proposed change in the timing of validation, a note would be included on the Hospital Compare website that until validation is completed, data displayed are preliminary. After validation and appeals were completed, the final VBP data report for the annual measurement period would be posted on the main CMS website (www.cms.hhs.gov). This proposed approach would allow CMS to provide timely release of performance data to consumers and still appropriately identify validation failures in the final data report.

Figure E-2: Timeline for Reporting Data on Hospital Compare



For hospitals that fail validation, data for the last quarter of the measurement period (Quarter D) would be removed from Hospital Compare, but data for Quarters E - G would appear and a descriptive footnote provided.

***Clinical Data Abstraction Center (CDAC)** is the contractor used by CMS to carry out the process of validating data collected from medical records for the RHQDAPU Program and would continue to be used for the VBP Program.

Appendix E-1: Example to Illustrate Differences Between Data Element Match and Measure Match Rate Approaches

Four heart failure process-of-care measures were selected to illustrate the difference between the data element match and measure match approaches. There are 16 data elements abstracted and used in calculating these measures. Table E-1 presents whether a hospital's data abstraction and the CDAC assessment match for the 16 elements for two cases. We use the "smoking cessation" measure to illustrate the measure match process.

Of the 16 heart failure data elements, there are 5 data elements that define the smoking cessation measure: principal diagnosis code, comfort measures only, discharge status, adult smoking history, and adult smoking counseling. For this measure, all five elements must match for the measure to match. For Case 1, all five data elements match. (See Table E-1.) Therefore, the measure matches, as is shown in Table E-2. For Case 2, the "adult smoking counseling" element did not match; therefore, the measure does not match. (See Table E-2.) The information in Table E-2 can be used to calculate the measure match rate for all of the heart failure measures. Case 1 had four out of four measures match, while Case 2 had two out of four measures matches, for a total of 6 measure matches out of 8 measures, or a 75 percent overall match rate. Similar measure matches would be determined for the example hospital's other 48 cases sampled for validation. The hospital's overall *validation score* would be calculated from all measures and all cases sampled as shown below.

$$\text{Validation Score} = (\text{total measure matches}/\text{total measures}) \times 100\%$$

Table E-1: RHQDAPU Validation: Element Level Example – Heart Failure

Data Element	Hospital / CDAC Assessment		Data Element	Hospital / CDAC Assessment	
	Case 1	Case 2		Case 1	Case 2
Principal diagnosis code	Match	Match	Discharge instructions address weight monitoring	Match	Match
Comfort measures only	Match	Match	LVF assessment	Match	Match
Discharge status	Match	Match	LVSD	Match	Match
Discharge instructions address activity	Match	Match	Contraindications to both ACEI and ARB at discharge	Match	<i>Mis-match</i>
Discharge instructions address diet	Match	Match	ACEI at discharge	Match	<i>Mis-match</i>
Discharge instructions address follow-up	Match	Match	ARB at discharge	Match	<i>Mis-match</i>
Discharge instructions address medications	Match	Match	Adult smoking history	Match	Match
Discharge instructions address symptoms worsening	Match	Match	Adult smoking counseling	Match	<i>Mis-match</i>

Table E-2: VBP Validation – Measure Level for Heart Failure

	Discharge Instructions Measure		LVF Assessment Measure		ACEI/ARB Measure		Smoking Cessation Measure	
	Hosp	CDAC	Hosp	CDAC	Hosp	CDAC	Hosp	CDAC
<i>Case 1</i>								
Numerator: delivered process of care	Yes	Yes	Yes	Yes	No	No	No	No
Denominator	Included	Included	Included	Included	Included	Included	Included	Included
Measure Match Status	Match		Match		Match		Match	
<i>Case 2</i>								
Numerator: delivered process of care	Yes	Yes	Yes	Yes	NA	No	Yes	No
Denominator	Included	Included	Included	Included	Excluded	Included	Included	Included
Measure Match Status	Match		Match		<i>Mismatch</i>		<i>Mismatch</i>	

Appendix F: VBP Program Monitoring, Evaluation, and Support for Quality Improvement

Appendix F presents a detailed discussion of the following important issues:

- VBP Program monitoring
- Strategies for monitoring the impact of Value-Based Purchasing on disparities
- Design elements of the VBP Program that seek to reduce disparities
- Quality improvement support for hospitals

VBP Program Monitoring

When implementing the Hospital VBP Program, CMS would foster an active learning system—within and across hospitals and at CMS—to promote breakthrough improvements in health care delivery. CMS’ ability to create an active learning system within a VBP Program would depend on having robust, ongoing program monitoring and evaluation functions. Building into the VBP Program, from its inception, appropriate ongoing program monitoring functions and sustained, systematic programmatic evaluation are critical program requirements. CMS would also require ongoing ready access to patient-level data to address many of the evaluation issues, such as unintended consequences among certain subpopulations.

Key areas to monitor would include:

- *Programmatic impact:* Has VBP (both the financial incentives and public reporting components) improved quality and efficiency within the Medicare program?
- *Distribution of payments:* Does VBP provide similar incentives to all hospitals to improve or maintain their performance, or do some hospitals consistently face challenges in improving or maintaining their performance that VBP may inadvertently exacerbate?
- *Implementation:* Are there aspects of the VBP Program infrastructure that could be strengthened to help hospitals participate more easily in the Program or to facilitate quality improvement?
- *Spillover effects:* Does the VBP Program incentivize broader systems-level changes within hospitals that result in improvements in other non-incentivized areas?

- *Best practices*: What best practices are being implemented by high-performing hospitals? How can these strategies be shared with other hospitals to improve care nationally?
- *Unintended consequences*: Does the VBP Program, as a function of its design, result in:
 - Hospitals “teaching to the test” (and potentially reducing quality in unmeasured areas)?
 - Hospitals dropping or avoiding caring for patients who are sicker or more difficult to manage, as observed through reduced access to care and increases in transfers? These actions could be monitored using administrative data to examine rates of transfers and changes in where vulnerable populations receive care, as well as patient experiences measured through HCAHPS.
 - Increased disparities in care (by region, race/ethnicity, etc.)?
 - Gaming of the data by hospitals to secure incentives? Data validation could be used to monitor for increase in use of exclusions.
 - Shifting of un-reimbursed costs to other payers?
- *Budget neutrality*: Does the VBP Program cause hospitals to increase their volume of services to offset potential losses in income because of being at risk for performance?
 - CMS recognizes that the pursuit of better outcomes could result in higher utilization rates and costs, which could increase Medicare spending. For example, striving to lower mortality rates could induce more frequent readmissions (Part A), more frequent ambulatory visits (Part B), or greater utilization of prescription drugs (Part D). Thus, performance would be monitored across the components of Medicare and all of the performance domains integrated within VBP.

More broadly, as it monitors the VBP Program, CMS could adjust design elements to minimize any negative effects. As an illustration:

- Consider that when the set of performance measures tied to incentives is narrow, there are more opportunities for providers to teach to the test. As the set of performance measures becomes more comprehensive, covering a wide array of the types of services that hospitals provide, hospitals would have to focus on a more expansive set of activities, as well as to make larger system improvements to demonstrate quality and secure the incentive.

- Large incentives may also distort behavior and encourage gaming, so it would be important for CMS to find a level of incentive that is important enough to encourage hospitals to invest resources and make organizational changes but not to game their data or engage aggressively in behaviors that reduce access for patient populations with more difficulties. Ongoing monitoring to identify such effects is an essential component of the overall VBP Program design and one that would need to be sustained throughout the life of the VBP Program.

Monitoring the Impact of Value-Based Purchasing on Disparities

CMS is committed to payment policies that promote care delivery that is safe, effective, timely, patient centered, efficient, and equitable—for all Medicare beneficiaries. As CMS proposes changes in its payment policies to incentivize improved care delivery, CMS will remain mindful of how these changes may affect the documented disparities in health care.²⁵ Disparities have been shown to exist across geographic regions, provider type, and patient characteristics such as age, sex, income, and race-ethnicity.

The impact of VBP on disparities in health care is currently unknown. Some have raised concerns that VBP may worsen disparities as providers attempt to avoid patient populations, such as minority or poor patients who may be perceived as more difficult to treat, in an effort to improve their quality scores.²⁶ Another concern raised in the IOM’s *Rewarding Provider Performance* report is that “[p]opulations most affected by disparities in health care are cared for disproportionately by undercapitalized providers who are likely to lack the resources necessary to invest in the infrastructure needed to facilitate participation in pay for performance.”²⁷

In some health care settings, quality improvement efforts have had positive effects on reducing disparities. For example, Medicare managed care plans demonstrated substantial improvements

²⁵ Institute of Medicine, 2002. *Insuring Health - Care without Coverage: Too Little, Too Late*. Committee on the Consequences of Uninsurance, eds. Washington, DC: National Academy Press. Institute of Medicine, 2005. *Quality through Collaboration: The Future of Rural Health*. Washington, DC: National Academy Press.

²⁶ Casalino LP, Elster A. Will pay-for-performance and quality reporting affect health care disparities? *Health Aff (Millwood)* 2007;26:w405-14.

²⁷ Institute of Medicine, 2006. *Rewarding Provider Performance: Aligning Incentives in Medicare*. Washington, DC: National Academy Press.

in quality from 1997 through 2003.²⁸ During that period, there was an accompanying improvement in decreasing disparities for clinical process-of-care measures. In contrast, after New York State published its Coronary Bypass Graft Surgery mortality report card, there was an increase in disparities in surgery utilization.²⁹ Unfortunately, little improvement occurred in disparities for outcome measures. Disparity reduction may be more straightforward for clinical process-of-care measures, which require simpler actions by hospitals and their providers once a patient is admitted (*e.g.*, provision of aspirin at arrival for patients having a heart attack), than it is for outcome measures (*e.g.*, 30-day mortality), which might also be affected by a host of other factors outside a hospital's control.

Studies of the hospital setting have generally focused on racial disparities, which have been shown to exist on two separate levels^{30,31}:

- *Between-hospital disparities*: Hospitals that care for a larger fraction of African Americans have lower quality of care than other hospitals do.
- *Within-hospital disparities*: The care for African American patients is worse than that for other patients within a specific hospital.

Between-hospital disparities are of particular concern, because the care for elderly African Americans is remarkably concentrated. Just 89 hospitals care for nearly one of every four elderly African Americans, and 21 percent of all hospitals care for 69 percent of elderly African Americans with a myocardial infarction.³² Within-hospital disparities are critical because of the implication that minority patients may not be able to secure equitable care solely by switching from poorer-quality to higher-quality hospitals. Both of these types of disparities are also likely to exist for patients because of income level and ethnicity.

There are several strategies for monitoring the VBP Program's impact on disparities, including:

²⁸ Trivedi AN, Zaslavsky AM, Schneider EC, Ayanian JZ. Trends in the quality of care and racial disparities in Medicare managed care. *N Engl J Med* 2005;353:692-700.

²⁹ Werner, RM, Asch, DA, Polsky, D. Racial profiling: the unintended consequences of coronary artery bypass graft report cards. *Circulation* 2005; 111:1257-63.

³⁰ Barnato AE, Lucas FL, Staiger D, Wennberg DE, Chandra A. Hospital-level racial disparities in acute myocardial infarction treatment and outcomes. *Med Care* 2005;43:308-19.

³¹ Jha AK, Li Z, Orav EJ, Epstein AM. Where do elderly blacks receive hospital care? The concentration and quality of hospitals that care for elderly black Americans. *Archives of Internal Medicine* 2007; Forthcoming June 2007.

³² Skinner J., Chandra A., Staiger D, Lee J, McClellan M. Mortality after acute myocardial infarction in hospitals that disproportionately treat black patients. *Circulation*. 2005; 112:2634-41.

- *Stratifying performance results by various sub-populations (e.g., race/ethnicity, income).*
 - Provide internal feedback reports to each hospital on the disparities of care for that hospital to promote actions to reduce disparities.
 - Adopt VBP measures stratified by sub-populations as measures of performance that would be incorporated as part of the incentive component of a VBP Program. This step would be taken in future years if performance differentials persist.
- *Monitoring for potential adverse impacts, such as increased difficulty accessing care.* One way to do this would be through targeted special studies.
- *Monitor whether hospitals that disproportionately care for vulnerable patients will be adversely impacted financially by analyzing payment rates and total annual payments.*

To examine whether the VBP Program is increasing or decreasing disparities in care provided to various sub-populations, CMS would need hospitals to capture standardized information on race/ethnicity or other patient characteristics that should be monitored. While the current Medicare beneficiary files contain limited information on race that could be applied to the Medicare portion of the discharges reported on a performance measure, the measures apply to patients with all categories of insurance coverage (commercial, Medicaid, uninsured). Therefore, CMS would need hospitals to submit additional data elements at the patient level (*i.e.*, race, ethnicity, language spoken, payer and patient zip code) when providing performance data for any VBP measures. Additionally, in the absence of income data, relevant Census data would be linked by patient zip code to provide a proxy for a patient's income that could be used to conduct disparities analyses.

For the mortality measures, which are based solely on Medicare patients, CMS could use the Medicare beneficiary file to obtain data on the race, ethnicity, and Census block of residence. Unfortunately, race and ethnicity data are currently poorly documented³³ or not collected using a standardized tool and CMS would need to work with the Social Security Administration to improve the quality and completeness of such data on the Medicare beneficiary file. Until the accuracy of race and ethnicity data is improved, current data could be augmented by using proxy data from Census data.

³³ Arday SL, Arday DR, Monroe S, Zhang J. HCFA's racial and ethnic data: current accuracy and recent improvements. Health Care Financ Rev 2000;21:107-16.

Design Elements of the VBP Program that Seek to Reduce Disparities

Several components of the VBP Program analyzed in this report are intended to address potential negative impacts on disparities:

- *Assistance with measuring disparities in care*
- *Payment structure*
 - Because minority populations often receive care at undercapitalized hospitals that also tend to be poorer performers, financial incentives based only on attainment would have a disproportionately negative impact on these hospitals financially and further exacerbate disparities. Basing financial incentives on both improvement and attainment would mitigate these concerns.
 - In future years, incentivized performance measures could include a focus on reductions in disparities.
- *Broader array of measures*
 - In the future, the array of measures could include those particularly applicable to vulnerable populations.
- *QIO support*
 - One key role for the QIOs in VBP could be to focus on providing technical assistance to and sharing best practices with hospitals that disproportionately care for vulnerable populations.

Quality Improvement Support for Hospitals

For many years, the Quality Improvement Organizations (QIOs) have provided expertise and hands-on support to hospitals in their quality improvement efforts. CMS could modify and expand the technical assistance provided to hospitals in improving quality of care and quality measurement through its 53 QIOs. An emphasis of the QIOs' role could be to provide technical assistance to small and rural hospitals that have more limited infrastructure to support quality improvement interventions, to hospitals with disparities in care among subgroups of patients, and to hospitals with poor performance scores. The type of support that the QIOs could provide may take a variety of forms, including:

- Creating tools to assist hospitals in improving their processes of care;
- Providing training about quality improvement techniques and methodologies;

- Holding forums to facilitate the exchange of best practices across hospitals;
- Conducting site visits to observe and consult with hospitals on current and best practices;
- Providing technical assistance to and sharing best practices with hospitals that disproportionately care for vulnerable populations.

Appendix G: Glossary of Terms and Acronyms

Acronym/Term	Definition
Absolute performance threshold	A pre-specified level of performance that would qualify a hospital for an incentive payment (<i>e.g.</i> , 90 percent of patients with AMI must have received aspirin at arrival)
APU	Annual Payment Update
Benchmark	A reference point or basis of comparison
CMS Abstraction and Reporting Tool (CART)	A software application for the collection and submission of data to the QIO Clinical Warehouse and for analysis of quality improvement data
Clinical Data Abstraction Center (CDAC)	The contractor used by CMS to carry out the process for validating data collected from medical records for the RHQDAPU Program
Composite measures	An aggregation of individual measures
CAH	Critical Access Hospital
CY	Calendar Year
DRA	Deficit Reduction Act of 2005
DRG	Diagnosis-Related Group
FY	Fiscal Year
HCAHPS	A standardized survey instrument and data collection methodology for measuring patients' perspectives on hospital care; called the HCAHPS, the Hospital CAHPS, and the CAHPS Hospital Survey
HQA	Hospital Quality Alliance
IOM	Institute of Medicine
IPPS	Inpatient Prospective Payment System
JCAHO	The Joint Commission
Medicare Hospital Compare	A tool on the CMS website that provides information on how well hospitals care for their adult patients with certain medical conditions (www.hospitalcompare.hhs.gov)
MedPAC	Medicare Payment Advisory Commission
MMA	Medicare Prescription Drug, Improvement, and Modernization Act of 2003
NQF	National Quality Forum
OPPS	Outpatient Prospective Payment System
QIO	Quality Improvement Organization
QIO Clinical Warehouse	Data repository maintained by the Iowa Foundation for Medicare Care that contains data uploaded from hospitals across the nation for various initiatives
QualityNet Exchange	A secure website approved by CMS for communications and data exchange that contains updates, tools, and applications useful for public reporting and data submission (www.qnetexchange.org)
PHQID	Premier Hospital Quality Incentive Demonstration, a pay-for-performance demonstration sponsored by CMS

P4P	Pay for performance
P4R	Pay for reporting
Relative performance threshold	A level of performance that would qualify a hospital for an incentive payment that is determined by comparing the performance of participating organizations (<i>e.g.</i> , 75 th percentile performance across all hospitals)
RHQDAPU	Medicare’s Reporting Hospital Quality Data for Annual Payment Update Program
Risk adjustment	A method to reduce effects on performance measures of characteristics of the patient population that affect results but are outside the control of providers and are not randomly distributed, such as level of illness in the population
Target	A performance goal that reflects a level of performance greater than the threshold or a desired level of improvement
Threshold	A minimum level of performance that would qualify a hospital for payment
VBP	Value-based purchasing

Appendix H: CMS Hospital VBP Workgroup Membership

Name	Group	Project Role
Benedicta Abel-Steinberg	Office of Beneficiary Information Services	Public Reporting Subgroup Co-Lead
Sheila Blackstock	Office of Clinical Standards and Quality	
Cheryl Bodden	Office of Clinical Standards and Quality	
Susan Bogasky	Assistant Secretary for Planning and Evaluation	ASPE Liaison
Erin Clapton	Office of Legislation	
Loretta Conyers	Center for Medicare Management	
Rachel Duguay	Office of Research, Development, and Information	
Shannon Flood	Office of Research, Development, and Information	
Laura Gange	Office of E-health Standards and Services	
Elizabeth Goldstein	Center for Beneficiary Choices	
Lisa Grabert	Special Program Office for Value-Based Purchasing	
Nilsa Gutierrez	CMS New York Regional Office	
Debbra Hattery	Office of Clinical Standards and Quality	
Valerie Hartz	Office of Information Systems	
Angelique Hebert	Office of Beneficiary Information Services	
Julianne Howell	Independent Technical Advisor, Special Program Office for Value-Based Purchasing	Project Coordinator
Terrence Kay	Center for Medicare Management	
Joseph Kelly	Center for Medicare Management	
Terris King	Office of Clinical Standards and Quality	VBP Forum Lead
Mark Koepke	Office of Clinical Standards and Quality	
Lisa Lang	Office of Clinical Standards and Quality	
Linda Magno	Office of Research, Development, and Information	
William Matos	Office of Clinical Standards and Quality	Data Infrastructure and Validation Subgroup Co-Lead
Renee Mentnech	Office of Research, Development, and Information	
Karen Milgate	Office of Policy	
David Miranda	Center for Beneficiary Choices	Public Reporting Subgroup Co-

Name	Group	Project Role
		Lead
Susan Nedza	CMS Chicago Regional Office	
Helen Nolt	Office of Clinical Standards and Quality	
James Poyer	Office of Clinical Standards and Quality	Data Infrastructure and Validation Subgroup Co- Lead
Sylvia Publ	CMS Chicago Regional Office	
Linda Radey	Office of Research, Development, and Information	
Michael Rapp	Office of Clinical Standards and Quality	Measures Subgroup Co- Lead
Thomas Reilly	Office of Research, Development, and Information	
Lesley Reis	Office of Legislation	
Liz Richter	Center for Medicare Management	
Sheila Roman	Office of Clinical Standards and Quality	Measures Subgroup Co- Lead
Steven Sheingold	Assistant Secretary for Planning and Evaluation	ASPE Liaison
Sarah Shirey- Losso	Center for Medicare Management	
Stewart Streimer	Center for Medicare Management	
Dennis Stricker	Office of Clinical Standards and Quality	
Donald Thompson	Center for Medicare Management	Incentive Structure Subgroup Lead
Karen Trudel	Office of E-health Standards and Services	
Thomas Valuck	Special Program Office for Value Based Purchasing	Project Lead
Timothy Walsh	Office of Beneficiary Information Services	
Mark Wynn	Office of Research, Development, and Information	

Appendix I: VBP Plan Support Contractors

RAND Corporation

Cheryl L. Damberg, PhD
Melony E. S. Sorbero, PhD
Ateev Mehrotra, MD
Marc N. Elliott, PhD
John L. Adams, PhD
Stephanie S. Teleki, PhD

Lily Bradley, BA
Susan L. Lovejoy, MS
Lee Hilborne, MD
Katrin Hambarsoomians, MS
Magdalen Paskell, BA

Brandeis University

Christopher P. Tompkins, PhD
Grant A. Ritter, PhD

Timothy C. Martin, PhD

Booz Allen Hamilton

Aparna Higgins, MA
Kristine Martin-Anderson, MBA

Mona Mahmoud, MPH
Fatima Riaz, MPH

Boston University

Gary Young, JD, PhD
James F. Burgess, PhD

Dan R. Berlowitz, MD

HSAG

Andrea B. Silvey, PhD

Mary Fermazin, MD

