



**California Physician Performance Initiative**

**Methodology for Physician Performance Scoring  
Cycle 4**

**July 31, 2009**

## TABLE OF CONTENTS

<b>I. MEASURES AND MEASURES SPECIFICATIONS</b> .....	<b>3</b>
<b>II. DATA USED TO CONSTRUCT PERFORMANCE MEASURES</b> .....	<b>8</b>
<b>III. ATTRIBUTION METHODOLOGY</b> .....	<b>8</b>
A. PRIMARY CARE PHYSICIAN ATTRIBUTION RULE .....	9
B. MEASURE-RELEVANT SUB-SPECIALIST ATTRIBUTION RULE.....	9
C. PRACTICE SITE ATTRIBUTION RULE .....	10
D. EVALUATION AND MANAGEMENT (E&M) CODES USED AS THE BASIS FOR ATTRIBUTION .....	11
<b>IV. CONSTRUCTING PHYSICIAN PERFORMANCE SCORES</b> .....	<b>12</b>
<b>V. CONSTRUCTING A MASTER PHYSICIAN LIST (MPL)</b> .....	<b>12</b>
<b>VI. ASSESSING THE RELIABILITY OF PERFORMANCE SCORES</b> .....	<b>13</b>
<b>VII. COMPOSITE MEASURES</b> .....	<b>14</b>
<b>VIII. DATA VALIDATION</b> .....	<b>16</b>
<b>IX. PHYSICIAN REPORTING</b> .....	<b>17</b>

## I. Measures and Measures Specifications

The California Physician Performance Initiative (CPPI), operated under the auspices of the California Cooperative Healthcare Reporting Initiative (CCHRI), generated 17 clinical performance measures for the Cycle 4 physician measurement work. The Cycle 4 measurement year (MY) spans a 12-month data period from October 1, 2007 to September 30, 2008.

All but two of the 17 measures—cervical cancer screening (CCS) and coronary artery disease: LDL drug therapy for patients with d (CAD7)—were run during Cycle 3 (MY 2007), as part of the CMS-Delmarva funded Better Quality Information (BQI) pilot project with CCHRI. Because of the involvement of the Centers for Medicare and Medicaid Services (CMS) in Cycles 1-3, the measures in those three cycles were generated using combined commercial PPO and Medicare FFS claims data. At the close of the BQI project in October 2008, Medicare no longer permitted pilot sites the use of Medicare data. Therefore, Cycle 4 measures were generated using combined commercial PPO and HMO claims/encounter data from three California health plans—Anthem Blue Cross, Blue Shield of California, and United Healthcare.

The 17 clinical measures contained in Cycle 4 are displayed in Table 1 on page 4, with a brief summary of their specifications and criteria for patient exclusions. The measures selected for Cycle 4:

- Are relevant to the Commercial-age population,
- Have been endorsed by the National Quality Forum (NQF) and/or the American Medical Association's Physician Consortium on Performance Improvement (PCPI),
- Could be calculated using the available administrative claims data, and,
- Were reviewed and approved by the CPPI Physician Advisory Group (PAG).

The Pharmacotherapy Management of COPD Exacerbation measure has two numerators (Systemic Corticosteroids and Bronchodilators) producing two separate performance rates.

Performance scores for each measure are calculated as a ratio, in which the denominator represents all patients who should have received a particular service (i.e., who meet all of the denominator qualification criteria) and the numerator represents the number of the denominator qualifying patients who received the service, based on the information found in the claims data maintained by each health plan.

Measures are first scored at the patient level (i.e., patients were categorized as either in or out of the numerator and denominator for each measure), and then measure events were attributed to physicians of the relevant specialty for each measure per the attribution rules (see Section III). Patient-level data (numerator and denominator events) were aggregated to generate physician-level scores, using aggregated commercial HMO and PPO data sources.

**Table 1: Cycle 4 CPPI Measures, Abbreviated Specifications, and Relevant Specialties**

	Measure Name	Measure Code	Measure Description	Relevant Specialties for the Measure	Measure Source	Exclusions
1	Breast Cancer Screening	BCS	Women, age 42-69 on 9/30/2008, who had mammogram during 10/01/2006-9/30/2008.	Family practice, internal medicine, and OB/GYN.	HEDIS	Bilateral mastectomy, Unilateral mastectomy (2 occurrences on 2 different dates)
2	Colorectal Cancer Screening*	COL	Patients, age 51-80, who had a FOBT during 10/01/2005 to 9/30/2008 or a sigmoidoscopy or DCBE during 10/01/2005 to 9/30/2008 or a colonoscopy during 10/01/2005 to 9/30/2008.	Family practice, internal medicine, and gastroenterology, colorectal surgeon.	HEDIS	Diagnosis of colorectal cancer or total colectomy.
3	Cervical Cancer Screening	CCS	Women, age 21-64 on 9/30/2008, who had a PAP test during 10/01/2005-9/30/2008.	Family practice, internal medicine, and OB/GYN.	HEDIS	Hysterectomy
4	Diabetes Care: LDL Screening	LDL	Diabetics, age 18-75, who had an LDL-C screening test during 10/01/2007- 9/30/2008	Internal medicine, family practice, endocrinology, cardiology.	HEDIS	Polycystic ovaries, gestational or steroid induced diabetes.
5	Diabetes Care: HbA1c Screening	HBA	Diabetics, age 18-75, who had an HbA1c screening test during 10/01/2007- 9/30/2008.	Internal medicine, family practice, endocrinology.	HEDIS	Polycystic ovaries, gestational or steroid induced diabetes.
6	Diabetes Care: Nephropathy Screening	NPH	Diabetics, age 18-75, who had a nephropathy screening test or evidence of nephropathy 10/01/2006- 9/30/2008	Internal medicine, family practice, endocrinology, and nephrology.	HEDIS	Polycystic ovaries, gestational or steroid induced diabetes.
7	Cardiovascular Care: LDL Testing	CMC	Patients, age 18-75, who were hospitalized during 10/01/2006 to 9/30/2008 for an AMI, CABG, or PTCA, or were diagnosed with IVD during 10/01/2006 to 9/30/2008, and who had an LDL test during 10/01/2007 to 9/30/2008.	Family practice, internal medicine, and cardiology.	HEDIS	None
8	Cardiovascular : Beta Blocker at 6 Months After a Heart Attack	PBH	Patients, age 35+, who were hospitalized during 4/01/2007 to 3/31/2008 for an AMI and received beta-blocker therapy for the 6 months after discharge.	Family practice, internal medicine, and cardiology.	PCPI	Contraindication to beta blocker therapy.
9	Coronary Artery Disease: LDL Drug Therapy	CAD2	Coronary artery disease patients, age 18+ on 10/1/2007, who were prescribed a lipid-lowering therapy	Family practice, internal medicine, and cardiology, endocrinology.	PCPI	None
10	Coronary Artery Disease: LDL Drug Therapy for Patients with Diabetes	CAD7	Coronary artery disease patients, age 18+ on 10/1/2007, who also have diabetes, who were prescribed ACE inhibitor or ARB therapy	Family practice, internal medicine, and cardiology, and endocrinology.	HEDIS	Contraindication to ACE/ARB therapy
11	Heart Failure: Left Ventricular Ejection Fraction (LVF) Testing	HF2	Heart failure patients aged 18+ who were hospitalized 10/01/2007- 9/30/2008 and had a LVEF test.	Family practice, internal medicine, and cardiology.	PCPI	None

	Measure Name	Measure Code	Measure Description	Relevant Specialties for the Measure	Measure Source	Exclusions
12	Heart Failure: Warfarin Therapy for Patients with Atrial Fibrillation	HF8	Heart failure patients, age 18+, who were hospitalized with paroxysmal or chronic atrial fibrillation during 10/01 2007-9/30/2008 and were prescribed warfarin therapy.	Family practice, internal medicine, and cardiology.	PCPI	Contraindication to Warfarin therapy
13	Monitoring Patients on Persistent Medications	MPM	Combined rate for patients, age 18+, who were prescribed at least a 180-days supply of ambulatory medication therapy for (1) ACE inhibitors or ARBs; (2) Digoxin; or (3) Diuretics during 10/1/2007-9/30/2008	Family practice, internal medicine, and cardiology.	HEDIS	None
14	COPD Care: Pharmacotherapy Management of COPD Exacerbation Bronchodilator	PCEb	Patients age 40+, with a COPD exacerbation, with an inpatient discharge or ED encounter between 10/1/2007 – 8/30/2008, who received a bronchodilator within 30 days.	Family practice, internal medicine, allergy/immunology, and pulmonology.	HEDIS	None
15	COPD Care: Pharmacotherapy Management of COPD Exacerbation Corticosteroid	PCEc	Patients age 40+, with a COPD exacerbation, with an acute inpatient discharge or ED encounter between 10/1/2007 – 8/30/2008, who received a systemic corticosteroid within 14 days of the event.	Family practice, internal medicine, allergy/immunology, and pulmonology.	HEDIS	None
16	COPD Care: Use of Spirometry Testing	SPR	Patients age 42+, with a new or newly active COPD diagnosis between 4/1/2007 – 3/31/2008, who received spirometry testing two years prior to diagnosis or within 6 months of diagnosis.	Family practice, internal medicine, allergy/immunology, and pulmonology.	HEDIS	None
17	Disease Modifying Anti-Rheumatic Drug	ART	Patients, age 18+, diagnosed with rheumatoid arthritis who received at least one ambulatory prescription for a disease modifying anti-rheumatic drug during 10/01/2007-9/30/2008.	Family practice, internal medicine, internal medicine, rheumatology, allergy/immunology.	HEDIS	HIV, pregnancy

\* The Colorectal Cancer Screening measure has a ten year look back period, but only three years of data are available for this report.

Measures were computed using claims and eligibility information from 2005, 2006, 2007 and 2008 to support measure calculations.

For 12 of the 17 measures, CPPI used the 2009 National Committee for Quality Assurance (NCQA) Healthcare Effectiveness Data and Information Set (HEDIS) specifications ([www.ncqa.org](http://www.ncqa.org)). HEDIS specifications are copyrighted by NCQA and permission to reproduce these is required by NCQA. The remaining four measures were generated using the Medicare Care Management Performance (MCMP) Demonstration and Quality Measurement specifications that were created by the Physician Consortium for Performance Improvement (PCPI) ([www.ama-assn.org](http://www.ama-assn.org)).<sup>1</sup> More detailed specifications can be found in Appendix A of this document, while the full specifications need to be obtained from the source organization.

The denominator, numerator and attribution periods for each measure are displayed in Table 2, page 6.

- **Denominator period:** the time period during which the member had to be enrolled and/or had a clinical condition/event to qualify for the measure.
- **Numerator period:** the time period during which the member was to have received the numerator qualifying service (e.g., mammogram).
- **Attribution period:** the time period during which the E&M visit that is used to attribute a denominator qualifying member to a physician must have occurred.

The CPPI measurement work aligns the numerator and attribution periods so that a physician must have seen the member for an E&M visit during the time period in which they were to have received the numerator service. Though the Colorectal Cancer Screening numerator timeframe is 10 years, the attribution period is set at two years and the numerator timeframe is set at three years given the available data.

**Table 2. Measure Denominator, Numerator and Attribution Periods, Cycle 4**

	Measure Name	Denominator Period	Numerator Period	Attribution Period
<b>HEDIS Measures</b>				
BCS	Breast Cancer Screening	10/01/2006 -- 09/30/2008	10/01/2006 -- 09/30/2008	10/01/2006 -- 09/30/2008
COL	Colorectal Cancer Screening	10/01/2006 -- 09/30/2008	10/01/1998 -- 09/30/2008*	10/01/2006 -- 09/30/2008
HBA	Diabetes: HbA1c Test	10/01/2006 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
LDL	Diabetes: LDL Screening	10/01/2006 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
CMC	Cardiovascular: LDL Screening	10/01/2006 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
NPH	Medical Attention for Nephropathy	10/01/2006 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
PBH	Persistence of Beta Blocker Treatment After a Heart Attack	4/1/07 – 3/31/08	180 day period after discharge date	180 day period after discharge date
MPM	Annual Monitoring for Patients on Persistent Medications	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
SPR	Use of Spirometry Testing in the Assessment and Diagnosis of COPD	4/1/07 – 3/31/08 (for index start date)	2 years before to 180 days after Index Start Date	180 day period after Index Episode Start Date
PCE	Pharmacotherapy Management of COPD Exacerbation	10/1/07 – 08/31/08	14 days and 30 days after discharge/ER visit	10/01/07 to 14 and 30 days after discharge/ER visit

<sup>1</sup> “Medicare Care Management Performance Demonstration and Quality Measurement Specifications,” a report prepared by Jody Blatt of CMS (Dated May 7, 2007; Updated August 3, 2007).

	Measure Name	Denominator Period	Numerator Period	Attribution Period
ART	Disease Modifying Anti-Rheumatic Drug Therapy for Rheumatoid Arthritis	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
CCS	Cervical Cancer Screening	10/01/2005 -- 09/30/2008	10/01/2005 -- 09/30/2008	10/01/2005 -- 09/30/2008
<b>Medicare Care Management Performance (MCMP) Demonstration and Quality Measurement Measures</b>				
CAD2	Coronary Artery Disease	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
HF2†	Heart Failure: Left Ventricular Ejection Fraction Testing	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
HF8	Heart Failure: Warfarin Therapy for Patients with Atrial Fibrillation	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008
CAD7	ACE inhibitor/angiotensin receptor blocker (ARB) Therapy	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008	10/01/2007 -- 09/30/2008

† All dates are inclusive. \*While the COL measure specification in HEDIS has a 10-year look-back period, for CPPI, the lookback period was truncated to three years.

There were a small number of measure issues and or modifications to specifications in Cycle 4.

- For the monitoring of patients on persistent medications (MPM) measure, the combined medication rate does not include the count of numerator positive events for anti-convulsants, as called for in the specifications. The CPPI PAG determined that the clinical evidence supporting annual monitoring of this drug category was relatively weak and recommended removal of the anticonvulsant category from the measure. The MPM rate was calculated by adding the numerators and denominators for the three remaining drug categories: digoxin, diuretics, and ACE/ARB.
- The numerator timeframe for colorectal cancer screening (COL) measure spans a 10 year period, which is longer than the time period for the claims and eligibility data available for Cycles 4. The Cycle 4 results are limited to a three-year look back period, based on available longitudinal claims data.
- The coronary artery disease measure (CAD7) is specified as “Percentage of CAD patients who also have diabetes and/or Left Ventricular Systolic Dysfunction (LVSD) who were prescribed ACE Inhibitor or ARB Therapy”. The identification of LVSD cannot be accurately done through administrative claims data because it requires test results which typically are not found in claims data. Therefore, the measure was modified to exclude LVSD as a denominator qualifying criteria. The description of the measure was changed accordingly to “Percentage of CAD patients who also have diabetes who were prescribed ACE Inhibitor or ARB Therapy”.

**HEDIS Measure Specification Updates:** There were no major specifications changes to the CPPI Cycle 4 measures derived from the NCQA HEDIS set, except for coding changes which occur each year as the NCQA reviews specifications and new codes are added or deleted (e.g., CPT, ICD-9, HCPCS, etc.).

**PCPI Measures Specification Updates:** The MCMP specifications were exactly the same as used in the BQI Cycle 3 (Dated May 7, 2007; Updated August 3, 2007) because the PCPI measures do not undergo the same annual cycle of review/update as the HEDIS measures. Three of the Cycle 4 measures (CAD-2, HF-8 and CAD-7) use NDC code lists, which have not been updated since the specifications were released. It is possible that performance rates may be affected by the exclusion of newer drugs from the specifications. Codes may also have been added or retired from other coding sets that would not be reflected in the 2007 specifications (e.g., in CPT, ICD-9 coding sets).

## II. Data Used to Construct Performance Measures

For Cycle 4, CCHRI combined commercial HMO and PPO claims data from five sources into an aggregated data base. The five sources were:

- Blue Shield of California: HMO
- Blue Shield of California: PPO
- Anthem Blue Cross: HMO
- Anthem Blue Cross: PPO
- United Healthcare: PPO

Complete data was available and used for all data sources for the time period 01/01/2006 to 9/30/2008 (for entire timeframe). We also used data for Blue Shield of California and Anthem Blue Cross (HMO and PPO) from 2005 (01/01/2005). There was partial data available for United Healthcare PPO in 2005 and this partial data was used.

We requested the following types of data files from each of the five plans, which were then used to generate the measures:

- Eligibility – member benefit enrollment information and member demographics
- Professional Claims – medical services (non-facility claims)
- Facility Claims – inpatient and outpatient facility services claims
- Pharmacy Claims – outpatient pharmaceutical claims
- Provider – physician and/or provider identifiers, demographics, and address.

Health plan members had to meet health plan continuous enrollment requirements within each data source (i.e., plan-product such as Blue Shield HMO) to qualify for a particular clinical measure. For example, if a member switched from a PPO plan to an HMO plan (within the same carrier or between carriers) during the continuous enrollment period, s/he would not meet continuous enrollment requirements. Members also had to meet any pharmacy coverage, per HEDIS specification (if measure requires pharmacy, then there has to be a pharmacy “flag” denoted in the member file) and anchor date requirements.

The health plan continuous enrollment requirement means that some patients who a physician is currently seeing will not be included in their measure rate calculation because the patients fail the continuous enrollment criteria. For example, the continuous enrollment requirement for Breast Cancer Screening is two years (10/1/2006 through 9/30/2008 for Cycle 4). If a member joined one of the participating health plans in January 2008, the member would not be included in the measure or in a physician’s score because they were only enrolled for part of the required two-year period, even if they saw the physician between January and September 2008 and had the recommended mammogram.

## III. Attribution Methodology

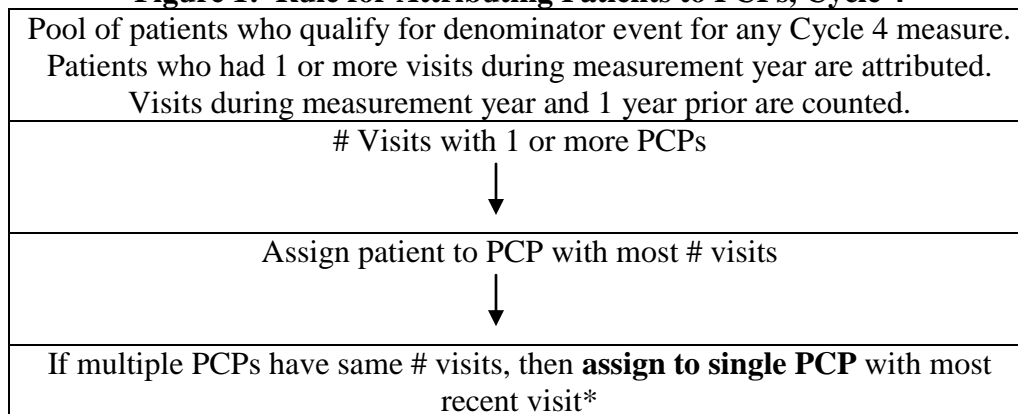
Under the CPPI attribution rules applied in Cycle 4, a patient event is attributed to one or more physicians of a relevant specialty for each measure (as noted in Table 1). To be assigned, the patient must have had  $\geq 1$  Evaluation & Management (E&M) visits with the designated physician(s) during the measurement period. Because many physicians have multiple specialty designations across the commercial health plans, we assigned a single, primary specialty to each physician and this specialty was used in the patient attribution step. Physicians are provided the opportunity during the “Review and Comment Period” to see their specialty designation and to make corrections.

For Cycle 4, we applied different rules to attribute Primary Care Physicians (PCPs) and sub-specialists. Additionally, as discussed in Section VII, we will group physicians of the same specialty into practice sites and attribute patients to these practice sites. The physician and practice-site attribution rules were reviewed and approved by the CPPI PAG and CCHRI Executive Committee.

### A. Primary Care Physician Attribution Rule

For all 17 measures in Cycle 4, a PCP is deemed a relevant specialist. In contrast to the approach used in Cycles 1-3, which assigned patient denominator events to all PCPs who had E&M visits with the patient during the measurement period (i.e., multiple PCP assignment), the Cycle 4 attribution rule assigns patients to a single PCP with whom the patient had the most ambulatory/outpatient E&M visits during the measurement year and one year prior (as shown in Figure 1). The rule imputes the patient’s PCP based on the patient’s visit activity over a two-year interval, to create a sense of greater ownership and responsibility on the part of the PCP who would be attributed responsibility for delivering the patient’s care.

**Figure 1: Rule for Attributing Patients to PCPs, Cycle 4**



\*Assignment to an “imputed PCP” can only occur if patient has seen a PCP within the measurement year (October 2007-Sept 2008). This implies that a patient who hasn’t been seen in most recent 12 months, but who may be eligible for a preventive screen (pap, mammography, colorectal), will go unassigned (and not measured).

The rule attributes a patient to single PCP who has the greatest number of E&M visits during the measurement year and year prior (i.e., 10/01/06 to 09/30/08) and with whom the patient had an E&M visit within the measurement year (i.e., 10/01/07 to 09/30/08). If the patient visit count is equivalent for two or more PCPs, then the patient is attributed to the PCP with whom the patient had the most recent visit. If the patient visits counts are equal and the most recent visit date is the same for multiple PCPs, the patient will be attributed to each of the PCPs.

### B. Measure-Relevant Sub-Specialist Attribution Rule

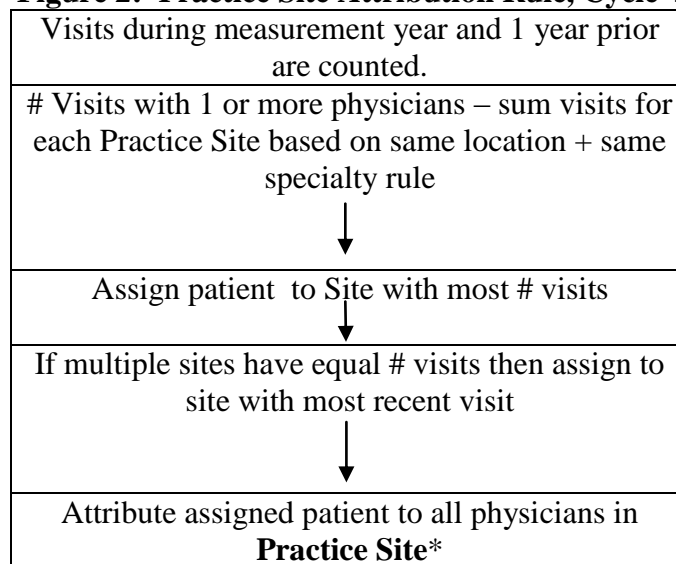
For Cycle 4, the rule for attributing patients to non-PCP measure relevant specialists remains unchanged from Cycles 1-3. Patients were attributed to any measure-relevant sub-specialist who had at least one E&M encounter with the patient during the measure’s attribution period (see Table 2). Patients can be attributed to multiple relevant sub-specialists for a given measure. For example, patients who qualify for the LDL screening measure for diabetes could be assigned to an endocrinologist and cardiologist if both types of physicians had an E&M visit with the patient during the attribution period. A physician’s specialty is determined through their primary specialty only, per the CPPI Master Physician List.

### C. Practice Site Attribution Rule

In addition to individual physician attribution for the 17 measures, CPPI is assigning physicians to practice sites and computing practice-site-level performance scores during Cycle 4. While narrow accountabilities may be more consistent with physician perceptions of the care they provide to patients, broader accountabilities (i.e., multi-provider) emphasize joint responsibility for ensuring the proper management of the patient during each and every care encounter.

CPPI defines a practice site as physicians of the same specialty who share a common address (note: the address must match at the office/suite level to be considered a common site). Practice sites are specialty specific; for example PCPs are grouped only with PCPs, cardiologist only with cardiologists etc. Physicians of different specialties who share the same location are assigned different practice site identifiers. The patient must have had an E&M visit with at least one physician in that same-specialty site to be attributed to that practice. The attribution logic for the practice site assignment rule is shown in Figure 2.

**Figure 2: Practice Site Attribution Rule, Cycle 4**



\*Assigns patient to all physicians in the multi-physician practice regardless of which physicians in practice have seen the patient.

The practice site results will be shared with physicians online via the same CCHRI pages that the physician uses to request their patient list. (**Note: this work is scheduled to occur after the close of the physician review and corrections period in September, to incorporate any changes to specialty or the numerators and denominators for measures, prior to constructing specialty specific practice sites and construction of site-level scores**)

## D. Evaluation and Management (E&M) Codes Used As the Basis for Attribution

The eligible E&M visit codes used in Cycle 4 are listed in Table 3. E&M visits are defined as professional encounters/claims with CPT codes identifying E&M services. The E&M visit did not have to have a diagnosis (Dx) code related to the measure for which the member qualified. Only office-based visits, hospital outpatient visits, and independent clinic visits were counted. These visits were identified by the Place of Service (POS) code on the claim/encounter.

- **Included POS Codes:** 11 (Office Visit), 22 (Outpatient Hospital), and 49 (Independent Clinic).
- **Excluded POS Codes:** Hospital Inpatient, Emergency Room, and Ambulatory Surgery Centers.

**Table 3: E&M Visit Codes Used to Attribute Patients to Physicians, Cycle 4**

E&M Visit	Definition
99201-99205	Office or Other Outpatient Services: New Patient
99211	Office or Other Outpatient Services: Established Patient: Presenting problems minimal (5 mins generally)
99212-99215	Office or Other Outpatient Services: Established Patient
99241-99245	Consultations: Office or Other Consultations: New or Established Patient
99271-99275	Consultations: Confirmatory Consultations: New or Established Patient
99341-99345	Home Services: New Patient
99347-99350	Home Services: Established Patient
99354-99357	Prolonged Services: Prolonged Physician Service with Direct (Face-to-Face) Patient Contact
99361	Case Management Services: Team Conferences
99374-99380	Care Plan Oversight Services
99381-99384	Preventive Medicine Services: New Patient (less than 18 years)
99385-99387	Preventive Medicine Services: New Patient (over 18 years)
99391-99394	Preventive Medicine Services: Established Patient (less than 18)
99395-99397	Preventive Medicine Services: Established Patient (over 18)
99401-99404	Preventive Medicine Services: Preventive Medicine, Individual Counseling
99411-99412	Preventive Medicine Services: Group Counseling
99420-99429	Preventive Medicine Services: Other Preventive Medicine Services
90801-90802	Psychiatric Diagnostic or Evaluative Interview Procedures
90804-90809	Office or Outpatient Facility: Insight Oriented, Behavior Modifying and/or Supportive Psychotherapy
90810-90815	Office or Outpatient Facility: Interactive Psychotherapy
90845-90857	Other Psychotherapy
90862-90899	Other Psychiatric Services or Procedures

There were cases in which denominator qualifying patients could not be attributed to a physician; these events are excluded from the performance results. The key reasons that patient events could not be attributed to a physician were:

1. member had an E&M visit, but not with a physician of a relevant specialty for the given measure,
2. member had an E&M visit, but the rendering physician could not be identified (i.e., rendering provider field on claim was blank or did not identify an individual physician), or
3. member did not have an E&M visit during the measurement period.

## IV. Constructing Physician Performance Scores

After patients are attributed to physicians, physician-level scores are calculated by aggregating attributed members across all data sources. This produces a set of physician-level scores. For each measure the numerator patient response is either 0 (did not have service) or 1 (had service) for each patient in the denominator. Each physician is attributed certain patients and the physician's performance is the aggregate of those numerator patient responses. Physician performance rates, or scores were calculated for each of the measures. Physician  $p$ 's true performance rate,  $\pi_k$ , is unknown. It was estimated from the data as:

$$r_p = \frac{1}{n_p} \sum_{i=1}^{n_p} X_{ip}$$

where physician  $p$  has a sample of  $n_p$  (denominator) patients and  $X_i = 1$  if patient  $i$  received the service and  $X_i = 0$  if patient  $i$  did not receive the service. That is, the performance rate was calculated as the numerator for the measure (i.e., the number of eligible members who had the service, such as breast cancer screening) divided by the denominator for the measure (i.e., the number of eligible members for the service).

## V. Constructing a Master Physician List (MPL)

Part of the work of aggregating data from different health plans involved consolidating the five data source provider files to identify a single set of unique physicians. Physicians were identified in numerous different ways across the commercial health plan data contributors. Data contributors have physicians stored in their systems with varying representations of the physicians' name, address, specialty, and identifier numbers. For instance, some data contributors use different health plan identifier numbers for the same physician within their own system. Across data contributors, there is even further variation in the way in which the same physician is represented. The most common identifier, tax identification number (Tax ID), is of limited value for identifying unique physicians because multiple physicians can use the same Tax ID and physicians can bill under more than one Tax ID. Further, those identifiers that are typically unique, such as the Unique Physician Identification Number (UPIN) and Drug Enforcement Administration (DEA) number, are often either entirely or intermittently absent from a given data contributors' files.

Given the absence of a common identifier across the sources of data, probabilistic linking of the claims records (professional, facility, and pharmacy) was used in this project to identify the universe of unique physicians practicing in California. Probabilistic linking does not pre-determine a point-value (or weight) for the various identifying variables, but instead allows the data itself to determine what those weights should be, based on the relative importance of an agreement on that variable. Probabilistic linking also allows for a variable's weight to be adjusted (or scaled) up or down to reflect the relative frequency of a particular value for the variable. For instance, in data where there are many last names of "Smith," then the weight for the last name variable would be scaled downward to reflect that agreement on this variable is less important; similarly, in data where there are very few last names of "Abercombe," then the weight for last name would be scaled upward to reflect that agreement on this variable is more important. Finally, probabilistic linking considers not only agreements but also disagreements in determining whether two records are linked (while most deterministic methods only include agreement measures). For instance, agreement on gender does not provide much information about whether two records are linked, but disagreement on gender does provide a lot of information that two records are not linked.

This process resulted in the creation of a CPPI Master Physician List (MPL) which contains the names, addresses, and specialties for all physicians that could be successfully identified within the data. A total of 80,616 unique physicians are contained in the MPL; 21,211 of whom (26%) are adult primary care (generalists, family physicians, internal medicine)". In addition, external reference source data used to validate and correct

physician specialty and complete missing address and identifier information including data from the California Medical Examiner Board, the NPI Registry and medical group physician directories. CCHRI has also created a web site through which physicians can validate their demographic information. On an ongoing basis, physicians will be encouraged to go to the web site and make any corrections to their information regarding specialty, address, and practice site designation.

## VI. Assessing the Reliability of Performance Scores

CPPI adopted a minimum reliability threshold score of 0.70 to indicate that there was sufficient signal in the estimate of performance to discriminate performance across physicians. Reliability is a measure that assesses the ratio of signal-to-noise. The signal in this case is the proportion of the variability that can be explained by real differences in performance. A reliability of zero implies that all of the variability in a measure is attributable to measurement error. A reliability of one implies that all of the variability is attributable to real differences in performance. High reliability does not mean that performance on a given measure is good, but rather that one can confidently distinguish the performance of one physician from another.

Measures of physician clinical quality, patient experience, peer review, medical errors, and utilization have been evaluated for their reliability (Hofer et al. 1999; Hayward and Hofer 2001; NCQA 2006; Hays and Revicki 2005). The primary motivation for using reliability as a metric is reliability's relationship to misclassification. If a provider is flagged as higher quality than his/her peers this classification is more likely to be correct the higher the reliability. The reliability of a HEDIS rate based on  $n$  denominator patients can be calculated from the Spearman-Brown prophecy formula:

$$R_i = \frac{\sigma_{MD}^2}{\sigma_{MD}^2 + \frac{\sigma_i^2}{n_i}}$$

In this formula,  $\sigma_{MD}^2$  is the between-physician variance, the variance among physician rates,  $\sigma_i^2$  is the binomial variance associated with the rate for physician  $i$ , and  $n_i$  is the rate denominator (sample size) for physician  $i$ . Consequently,  $R_i$  is the proportion of the total variance that is attributable to the variance among physicians. As is evident from the formula for  $R_i$ , the reliability of each physician's HEDIS rate increases with 1) the variation in rates among physicians, and 2) the physician's sample size (denominator).

We estimated the physician-to-physician variance,  $\sigma_{MD}^2$ , using a beta-binomial model, which assumed that the underlying true physician rates were distributed according to a beta distribution and that the observed rate for physician  $i$  was distributed according to a binomial distribution with sample size  $n_i$ . To estimate  $\sigma_{MD}^2$ , we used a SAS macro developed for this purpose (Wakeling, 2004). To estimate the binomial sampling variance,  $\sigma_i^2$ , for each physician, we used the property of the binomial distribution,  $\sigma_i^2 = p_i(1-p_i)$ , where  $p_i$  is the physician's true rate. For moderate to large sample sizes this variance can be estimated by substituting the physician's observed rate  $\hat{p}_i = (x_i / n_i)$  for  $p_i$  into the formula for  $\sigma_i^2$ , where  $x_i$  is the number of successes for the sample of  $n_i$  patients treated by physician  $i$ . However, many physicians have small denominators, resulting in unstable estimates of their true rates. Therefore, for the purpose of estimating the reliability,  $R_i$ , for physician  $i$ , we estimated each physician's rate,  $p_i$ , using a shrinkage estimator (explained in the next section) to compensate for the instability of estimated rates for small-denominator physicians. For example, without this adjustment, physicians with a denominator sample of size  $n = 1$  would always have  $p_i = 0$  or  $p_i = 1$ , resulting in an estimated reliability of 100 percent, which is clearly inappropriate.

We sought to establish a denominator sample size threshold that would achieve a certain level of reliability. However, for a given sample size,  $n_i$ , the reliability is different for different values of  $\sigma_i^2$ , which depends on the physician's HEDIS rate. Therefore, we established the sample size threshold as the sample size for which 90

percent of physicians had at least 70 percent reliability; we computed the observed 10<sup>th</sup> percentile of estimated reliability for each sample size n, and identified the value of n for which the 10<sup>th</sup> percentile of reliability was 0.70. This value of n was employed as a sample size threshold to categorize physicians as reliable.<sup>2</sup>

A physician's performance rate for a given measure was deemed to be reliable if his or her denominator sample size was at least as large as the sample size for which 90 percent of physicians had at least 70 percent estimated reliability. Note that some physicians will have slightly less than 70 percent estimated reliability using this sample size threshold, but the vast majority of physicians exceeding this sample size threshold will exceed or nearly exceed an estimated 70 percent reliability. It is also possible for a physician with a sample size below the threshold to exceed 70 percent estimated reliability. We addressed this by designating a physician's rate as reliable if his or her sample size fell below the sample size threshold, but his or her estimated reliability was at least 70 percent.

The patient denominator counts per physician to achieve the threshold of 0.70 or higher varies measure-to-measure. For measures that focused on less common conditions (e.g., heart failure), achieving reliable denominator sizes for individual measures at the physician-level is difficult given small patient samples per physician. Reliabilities can be improved through construction of composite measures at the individual physician level and by aggregating results to the practice site level.

## VII. Composite Measures

The Cycle 4 work includes construction of composite measures of performance at the individual physician and practice site-levels. **(Note: This work will commence at the close of the physician review and corrections period, once corrections have been made. This work will occur in October 2009, and as such, results are not contained in this document).**

A composite measure is a single, new score that is created by combining information across multiple individual measures. Among the reasons that composite measures are of interest:

1. Composite measures may increase the reliability of performance measures. To the extent that multiple measures yield similar results—measure the same underlying attribute—the effective sample size can be increased by pooling information over the related measures thereby increasing the “signal” (true performance rate) relative to the “noise” (sampling error).
2. Composite measures are more parsimonious, as they reduce the dimensionality of performance measurement information. This is particularly important when there are many measures. By distilling the information down to a few key performance dimensions, the use of composites makes it easier to assimilate information across many measures.

Our approach to constructing a composite measure seeks to satisfy the following properties:

1. The composite calculation should be understandable and reproducible by persons with average quantitative skills. This reflects our desire for “transparency” over a “black box.”
2. Each physician should be evaluated according to what he or she does, and not what he or she does not do. Not surprisingly, several papers have found that physicians tend to do well at the things they do all the time.

---

<sup>2</sup> For some measures, the 10<sup>th</sup> percentile of observed reliability never reached 70 percent. In those cases, we calculated the sample size threshold by solving the Spearman-Brown prophesy formula for n, and substituted the overall observed mean rate to estimate the binomial variance.

3. More weight should be put on measures that discriminate well than on measures that discriminate poorly. For example, if every ophthalmologist does an eye examination on every one of their diabetic patients it is difficult to discriminate among the ophthalmologists on the eye examination rate because there is no variance among them on that measure.
4. The composite score model should not mix physician specialties. It is likely that specialists face a much different case-mix than generalists. Case-mix could affect the probabilities of various actions.

A four-step process is used to create composite measures.

1. First, organize the individual measures into four clinical domains: preventive, diabetes, cardiac, and respiratory care. These domains are determined largely by the measures available for the study. The concentration on clinical domains makes the composites readily interpretable.
2. Second, determine the extent to which the individual measures are correlated (i.e., are related) with one another within each clinical domain. Examine correlations at the patient and physician level. Ensure that the correlations at the physician level are positive, suggesting an association with a single underlying measure of domain-specific quality.
3. Third, apply item-response theory (IRT) models to create physician-level composite quality measures based on the physician's observed rates for individual measures. These models, which measure the relative difficulty and discriminatory power among the individual measures, were recently suggested by Teixeira-Pinto and Normand (2008) for hospital composite quality measurement. IRT models are commonly used in education to estimate a latent trait like a student's math ability, based on responses to a set of math questions. For physician performance, the underlying physician trait is quality of care and the individual measures take on the role of test questions. Our approach took into account the relative difficulty of the measure and its discriminatory strength; a case mix adjustment. At least half the measures will be required to be non-missing within a composite to score a physician. To address missing data, for a given composite, if a physician has a missing value the unweighted average deviation from the all-physician mean for the remaining non-missing measures will be substituted for the missing value.
4. Fourth, to make the composite IRT score more transparent and interpretable, regress the estimated IRT score on reliability-adjusted individual measure scores. Therefore, the final composite score will represent a weighted average of reliability-adjusted individual scores (shrinkage estimates of individual scores), which can be easily calculated program. The weights are empirically derived.

Fourteen( 14) composite scores will be constructed: 7 of which will be physician-level and 7 of which will be practice site-level composite scores as shown below.

1. Preventive Care – PCP only: physician-level
2. Preventive Care – PCP only: practice-site-level
3. Cardiovascular – PCP: physician-level
4. Cardiovascular – cardiology: physician-level
5. Cardiovascular – PCP: practice-site-level
6. Cardiovascular – cardiology: practice-site-level
7. Diabetes – PCP: physician-level
8. Diabetes – endocrinology: physician-level
9. Diabetes – PCP: practice-site-level
10. Diabetes – endocrinology: practice-site-level
11. Respiratory – PCP: physician-level
12. Respiratory – Pulmonology/Immunology: physician-level
13. Respiratory – PCP: practice-site-level
14. Respiratory – Pulmology/Immunology: practice-site-level

The composite scoring will commence after the close of the physician review and comment period, so that we base our calculations on corrected data. The composite scoring work will occur in October 2009.

## VIII. Data Validation

During the Cycle 4 work, CPPI conducted several validation steps to inform its measurement and scoring work. Using 2007-2008 data, we tested the impact of two alternative approaches to attributing patients to PCPs—the method used in Cycles 1-3 which assigned to all PCPs that touched the patient vs. the method selected for use in Cycle 4, which assigned to the single PCP with the most E&M visits during the measurement year and one year prior. The validation work comprised both empirical analyses of the data and review of the findings by the CPPI PAG, the health plan medical directors, and the CCHRI Executive Committee. Additionally, per Table 4 below HMO and PPO performance rates were compared to national benchmarks to assess the degree of consistency. These benchmarks are drawn from administrative data only as a comparable reference point for the CPPI administrative-only data results.

**Table 4. Comparison of CPPI Cycle 4 Commercial Population-level Rates to External Benchmarks, (Rates Based on Administrative Data)**

Measure	Commercial		
	CPPI Rate	National Benchmark	Difference to Benchmark
BCS	62.0%	66.9%(1)	-4.9%
CCS	71.1%	73.5%(1)	-2.4%
COL	37.6%	42.5%(1)	-4.9%
CMC	61.5%	74.4% (1)	-9.9%
LDL	61.0%	72.7%(1)	-11.7%
HBA	64.0%	75.6%(1)	-11.6%
NPH	60.8%	64.16%(1)	-3.36%
PBH	70.4%	62.69(1)	-7.71%
MPM	66.5%	74.86%(1)	-8.36%
CAD-2	54.0%	*	*
CAD-7	73.1%	*	*
HF-2	43.2%	*	*
HF-8	69.3%	*	*
SPR	28.3%	33.66% (1)	-5.36%
PCE - Systemic Corticosteroid	48.9%	**	**
PCE - Bronchodilator	60.9%	**	**
ART	83.2%	78.93%(1)	4.27%

Notes: (1) Commercial HEDIS 2008 audit means for PPOs (NCQA). \*\* No benchmarks available: new HEDIS measures starting with measurement year 2008.

The CPPI scores compared to benchmark are an important element of the claims-based information approach but more work is needed to understand the extent to which incomplete data affects a physician score. The Cycle 4 physician review and corrections period is an opportunity to further understand the validity of claims-based measures as physicians self-validate their scores.

## **IX. Physician Reporting**

Physicians who have at least one measure with a denominator count of  $\geq 10$  received a report that displays the Cycle 4 results and shows the physician how she/he compares against her/his peers on a relative rank basis. The report displays performance results for measures that met the minimum reliability threshold of 0.70 as well as those that did not. Physician scores are displayed graphically only for measures with a denominator count of  $\geq 10$  patients. For measures in which the physician could be scored reliably, the bar is shaded in dark grey. If a measure did not achieve a 0.70 reliability or higher, then that particular measure is displayed with a light grey fill bar with an alert that the physician's true ranking likely differs from this result. All results, regardless of the level of reliability or size of patient count, are listed in a summary table, with scores displayed separately for HMO and PPO patients, and the combined results.

**Pilot Testing of Physician Reports:** CCHRI worked closely with physicians to engage them in the review and modification of the performance report during Cycle 3. Additionally, during Cycle 4, the PAG re-reviewed the tables and graphics contained in the physician reports.

To support action on the part of physicians for quality improvement and to support requests from physicians about the project, CCHRI worked with its Physician Advisory Group and the CCHRI Executive Committee to develop the Physician Review and Corrections Period. This process, accessed via [www.cchri.org/cppi](http://www.cchri.org/cppi) provides physicians the opportunity to do the following:

- 1- Update or change their demographic information, specialty designation, and practice site designation
- 2- Request a duplicate report
- 3- Request a patient list for their patients included in the measures
- 4- Submit additional data, if needed
- 5- Ask questions
- 6- Provide feedback

CPPI will track requests for information and modifications supplied by physicians; these findings will be reported back to the PAG and the CCHRI Executive Committee. To comply with HIPAA and the California Civil Code, The Physician Review and Corrections process includes a consent form that a physician is required to sign to authorize the release of the patient list derived from the aggregated commercial data.