Understanding Aetna Smart Compare™ designation methodology, 2023

1. Introduction

Aetna Smart Compare™ analyzes physician practices, so we can provide members actionable information about the most effective, high-quality physicians in their plan networks. For 2023, we enable designations for effectiveness and clinical quality in our authenticated provider search tools for members. Also, we actively recommend designated physicians to members through mail, email, and our authenticated portals.

We indicate designations in our authenticated portals using the phrases “effective care” for effectiveness designations, and “quality care” for clinical quality designations.

In California and Texas, Aetna Smart Compare designations for commercial customers are only for self-insured members and plans. Members excluded from the program will not see designations in their member portals.

This methodology describes common elements of our designations. You can also find a reference guide for some of our most common variables by visiting **Aetna.com/smartcompare**. In 2023, these specialties, procedures, and business lines are designated by Aetna Smart Compare under the following programs.

Programs available in 2023, Table 1

Program	Specialties	Procedure focus	Business line	Version
Cardiology	General Cardiology (Non-surgical)	Preventive care or disease management	Commercial	2022
Endocrinology[1]	Endocrinologists	General practice	Commercial	2023
Ob/Gyn	Gynecologists and Obstetricians, as well as associated midwives, physician assistants and nurse practitioners	General practice	Commercial	2023
Orthopedic hip & knee	Orthopedists	Treatment of hip or knee degeneration	Commercial	2023
PCP, Commercial	Family Medicine, General Pediatrics and Internal Medicine physicians, as well as associated physician assistants and nurse practitioners		Commercial	2023
PCP, Medicare	Family Medicine and Internal Medicine physicians, as well as associated physician assistants and nurse practitioners		Medicare	2023
Pulmonary medicine[1]	Pulmonologists	General practice	Commercial	2023
Spine Surgery	Neurosurgeons, and Orthopedists	Surgeries treating spine degeneration	Commercial	2022

[1] These programs are new in 2023

2. Two designations, two categories of measure

Each program in Aetna Smart Compare looks at two categories of measure — effectiveness and clinical quality. Each is designated independently. Either category (designation) may have one or more subcategories, and those may have many more individual measures. For instance, all effectiveness designations use episodes of care, of which there are hundreds in the *PCP, Commercial* program. All measurements are made at the physician practice or organization level, using tax identification numbers.

Subcategories and categories are evaluated for six possible outcomes:

1. **Criteria not met** — The performance of your physician practice(s) is worse than our metrics in a way that is statistically significant.
    a. We use this outcome at the subcategory and category level.
    b. We do not reach this outcome when we use external data; see **criteria not met – external data**.
    c. At the category level, this outcome indicates your practice is not designated: members will not see an Aetna Smart Compare designation for your practice.
2. **Criteria not met – external data** — The performance of your physician practice(s) does not meet thresholds or other performance criteria for external datasets.
    a. We only use this outcome at the subcategory level.
    b. This outcome replaces **criteria not met** for subcategories where we rely on external data.
3. **Designation earned** — After summarizing subcategories, the performance of your physician practice(s) is better than our metrics in a way that is statistically significant.
    a. We only use this outcome at the category level. It indicates your practice is designated in Aetna Smart Compare. Members will see designations for your practice for effectiveness and/or clinical quality.
4. **Performance not statistically significant** — The performance of your physician practice(s) is not significantly different from our metrics in a category or subcategory.
    a. We use this outcome at the subcategory and category level
    b. We do not reach this outcome when we use external data; see **criteria not met – external data**.
    c. At the category level, this outcome indicates your practice is not designated: members will not see an Aetna Smart Compare designation for your practice.
5. **Subcategory earned** — The performance of your physician practice(s) is better than our metrics in a way that is statistically significant.
    a. We only use this outcome at the subcategory level.
6. **Volume insufficient** — Your physician practice(s) does not have sufficient data for scoring.
    a. We use this outcome at the subcategory and category level.
    b. At the category level, this outcome indicates your practice is not designated: members will not see an Aetna Smart Compare designation for your practice.

We summarize subcategories at the category level. Summaries of **designation earned** indicate a designation in the relevant category, effectiveness or clinical quality. Each subcategory outcome has a value assigned to it (1, 0, -1). See Table 2 below to review values for outcomes.

Designation hierarchy, Table 2

Effectiveness or clinical quality designations in Aetna Smart Compare

Designated ≥ 1 > Not designated

Summary of subcategories

Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Criteria not met	**<0**

Subcategories

Subcategory X		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Subcategory Y		
	Subcategory earned	1
	Volume insufficient	0
	Criteria not met - external data	0
Subcategory Z		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

3. Volume thresholds

Program thresholds

Each program in Aetna Smart Compare has volume criteria to ensure statistical significance. The subcategories in each category often have volume criteria too. Where a physician practice does not meet volume thresholds, it is not eligible for the program's categories or subcategories. For some programs, we also exclude high-cost claimants. Refer to the program guides below for specific details (see sections 6-13).

COVID claim exclusions

For 2023 programs (see versions in Table 1), we exclude all claims containing coronavirus diagnoses, treatments or complications, using the COVID flag identifier in Optum Symmetry® Episode Treatment Groups® (ETG®) software version 10.1. (Previous versions of Aetna Smart Compare did not measure claim years after 2019; identifiable medical care for COVID in the United States begins with the initial outbreaks in early 2020.)

Pediatric practices

Except for the PCP programs, we exclude pediatric practices from programs using machine learning and statistical techniques to identify practices that primarily treat children and teenagers.

To identify pediatric groups, we split providers into two groups – providers with strong pediatric indicators (Group A) and all other providers (Group B). Strong pediatric indicators are defined as follows (Group A):

- Providers in the practice have a primary specialty code with a pediatric keyword (e.g. "pediatric," "adolescent," etc.), and
- 100% of their claims are related to pediatric specialty codes

For each physician (or practitioner) in each practice type (Groups A and B), we build a k-means clustering or grouping model using these features:

- Average member age, and
- The percentage of members with age 19 or less, and
- The attributed member count is equal to or greater than the 25th percentile of the total members in this practice

For physicians (or practitioners) in Group A, all providers clustered using these conditions are considered pediatricians. To account for providers with very small member volumes, we apply the threshold member count greater than or equal to the 25th percentile. Providers with member counts less than the 25th percentile are not considered pediatricians to avoid low volume misidentifications.

Using the clustering model, we identify physicians (or practitioners) in Group B as those who have:

- An average member age equal to 19 or less, and
- Where the percentage of these members is 90 percent or greater, and
- The attributed member count is equal to or greater than the 25th percentile of the total members in this practice.

The table below represents the results of this analysis.

Cluster analysis to identify pediatric practices, Table 3

Cluster sets	Ortho hip & knee	Ob/Gyn	Endocrinology	Pulmonary Medicine
Total PIN count	29,064	41,365	7,387	11,597
Pediatrics PIN count	938	108	904	699
Pediatrics percentage	3.2%	0.3%	12.2%	6.0%

See the Cardiology (section 6) and Spine surgery (section 13) programs to understand how we remove pediatric groups from these programs. When these programs are updated from the 2022 version of the methodology to the 2023 version, we will include this machine learning technique for these programs as well.

3.1 Common attribution

We use the attribution logic from Optum Symmetry® for all programs, except the PCP programs (section 3.2), where we use our own attribution logic for risk-adjusted utilization (section 4.2) and clinical quality measures (sections 5-5.2). Below is an outline of the logic we use from Symmetry.

- If an episode contains a surgical procedure as defined by Symmetry and it is also a major surgery as defined by Aetna, the episode is attributed to the physician who performed that procedure.
    - If there are two major surgeries, then the episode is attributed to the physician with the greatest allowed amounts.
    - If there is not a major surgery in the episode, but a minor surgery is found, the episode is attributed to the physician who performed the minor surgery.
    - If there are two minor surgeries, the episode is attributed to the physician with the greatest allowed amounts.

- If there is no surgery present, the episode is attributed to the physician with the highest number of visits based on management records as defined by Symmetry.
    - For physicians tied for highest number of visits, the episode is attributed to the physician with the most direct treatment provided.
    - If there is still more than one physician with the same number of visits and treatments, the episode is attributed to the physician with the highest allowed amount.

In addition to this attribution logic, we often use *involved episodes* in many of our programs. See sections 6-13 to learn more about involved episodes.

3.2 PCP attribution

- In PCP, Commercial and Medicare (sections 10 and 11), we use the following attribution logic for risk-adjusted utilization (section 4.2) and clinical quality measures (sections 5-5.2). Episodes of care are attributed using the common attribution method outlined in section 3.1.

- Member attribution is set each calendar quarter using the most recent 24 months of data. The most recent data is available in our data warehouse records one month prior to the start of the calendar quarter. For example, in the calendar quarter starting January 1, members are attributed based on the most recent 24 months of data available in our warehouse as of December. This data will include all claims received and paid through November 30.
- If a member has selected a PCP and there is a cap payment made to that PCP, then we attribute to the selected PCP (this will only occur in markets with PCP capitation); or
- We look for claims that must contain an outpatient place of service and one of the evaluation and management codes listed below:
    - Office or other outpatient visit for evaluation and management: 99201-05; 99211-15
    - Home visit for evaluation and management of a new patient: 99341-45; 99347-50
    - Prolonged physician service in the office or other outpatient setting requiring direct (face-to-face) patient contact beyond the usual service's first hour: 99354-55
    - Prolonged evaluation and management service before and/or after direct (face-to-face) patient care: 99358-59
    - Initial comprehensive preventive medicine evaluation and management: 99381-87
    - Periodic comprehensive preventive medicine reevaluation and management: 99391-97
    - Counseling and/or risk factor reduction intervention: 99401-04; or G-codes 0344, 0402, 0438, 0439
- And the rendering physician specialty is equal to Family Practice, Internal Medicine, Pediatrics; or
- If no Primary Care Physician (PCP) visits are found, then we use any claims for physician assistants or nurse practitioners, and
- We use the most recent 12 months of claims for physician assistants and nurse practitioners:
    - If the member has only one visit, then the member will be attributed to the rendering physician's group as defined by tax identification number.
    - If the member has more than one visit and the treating physicians are all with the same group, the member will be attributed to that group.
    - If the member has two or more visits in the current year, and the visits are treated by physicians of two different groups, the member will be attributed to the group with the most recent visit (if the member at least two visits with that group). Otherwise, we attribute to the group with the greatest number of visits, and if there is a tie, we attribute to the group with the most recent visit; or
    - If a Member has no PCP claims in the most recent 12 months, we include an additional 12 months of claims prior to the current year and we retry the previous three steps.

3.3 Claims years, and external program years

Throughout this document we identify various claim years for our internal measures, and program years for external data. See the table below for a summary of the claims data used in this guide. Here is an explanation of the table's abbreviations:
Our **Commercial claims years (CCY)**: wherever we use our Commercial claims data to calculate measures using internal or external specifications.
Medicare claims years (MCY): where we use internal or external Medicare claims data. When we use external claims data, this data is multi-payor.
External program year (EPY): where we use external programs identified by their calendar year conventions, or the year we verify external ratings or accreditations.

Claims years and external program years in Aetna Smart Compare, Table 4

	Effectiveness Category		Clinical quality category		
Programs	**Internal measures**	**External measures**	**HEDIS®**	**Internal measures**	**External measures**
Cardiology	2018-19 CCY	2018 MCY	2021 CCY & MCY		
Endocrinology	2020-21 CCY		2021 CCY		
Ob/Gyn	2020-21 CCY		2021 CCY	2021 CCY	
PCP, Commercial	2020-21 CCY		2021 CCY		2022 EPY
PCP, Medicare	2020-21 CCY		2021 CCY		2021 or 2022 EPY[2]
Pulmonary medicine	2020-21 CCY		2021 CCY		2020 EPY
Orthopedic hip & knee	2020-21 CCY	2019 MCY			2017-19 or 2022 EPY[3]
Spine surgery	2018-19 CCY	2018 MCY		2018-19 CY	

4. Effectiveness category: episodes of care subcategory

Introduction to episodes of care

Programs in Aetna Smart Compare contain various subcategories. For unique subcategories, please see our designation guides below in sections 6-13. This section provides an overview of the episodes of care subcategory, common to all programs.

"Episode of care" methodology analyzes medical cost and utilization. An episode of care for a member represents diagnostic and treatment services over time for a specific health condition. All relevant costs and utilization of services for the specific condition are part of a single grouping for analysis. An episode of care spans services from the onset of symptoms until treatment is complete. For chronic conditions, the episode lasts one year.

[2] CMS drug safety measures (2021), and NCQA Recognition (2022)
[3] CMS Care Compare (2017-19), and either accreditation from the Joint Commission (2022) or DNV (2022)

Episode treatment groups (ETG)

We use Optum Symmetry® Episode Treatment Groups® (ETG®) software version 10.1, an illness classification system, to build episodes of care. The ETG technology is a kind of "grouper" software. The software accepts health care claims (service line detail) and returns the ETG value, along with other patient details. The software combines all relevant doctor, hospital, pharmacy and ancillary testing claims data together.

An ETG ends when there is no treatment of the condition for a specific number of days. This interval is the “clean period.” For example, ETG 438300 (acute bronchitis) has a 30-day clean period. We consider any claims for this diagnosis within a 30-day period a recurrence of the same condition. When an episode starts for this ETG, all clinically consistent claims activity for the acute bronchitis group is added to this episode until the point where 30 days pass without any corresponding clinically consistent treatment. If we receive a claim for this condition after 30 days, a new episode starts.

We only use complete episodes in the evaluations of Aetna Smart Compare. Complete episodes are those that meet the clean periods before and after the measure, or episodes where a chronic condition lasts 365 days. (We look at multi-year conditions in episodes no greater than one year).

Risk adjustment: case-mix-adjusted expected allowed amounts, and episode severity

There are many variables that impact the use of health care resources such as the severity of illness, patient age, and comorbid conditions. To account for these variations in resource use, a case-mix adjusted “expected allowed amount” is created for each member assigned an ETG that is attributed to a physician group. This case-mix-adjusted, expected, allowed amount is compared to the actual, allowed amount for that episode of care. There can be variation in efficiency indexes over time for practices with low attribution numbers. To avoid attributing random variation to a practice pattern, the statistical significance of a provider group’s efficiency score is evaluated at $P \leq 0.10$.

Episode severity describes the intensity of a member’s condition. When determining an episode’s severity, we look at the relevant complication and comorbidity factors, indicating a sicker member who may require more extensive treatment for a related condition. The result is a severity score and severity level for episodes. Typically, the higher the severity score, the more costly the condition.

Severity-adjusting episodes provides a powerful unit of analysis for comparing provider performance when different providers care for members with the same condition but different severity levels.

When comparing the allowed cost to the expected cost of an episode, the expected cost reflects the case-mix-adjusted value for episodes. We also apply outlier logic at the case-mix category level. For case-mix categories, episodes of care are outliers if that episode is outside of two median absolute deviations away from the median in that case-mix category. We exclude these outliers from analysis.

Below we refer to all case-mix adjusted allowed amounts as “expected” allowed amounts.

Machine learning and market adjustment

We use machine learning models featuring the variables in Table 5 to risk-adjust case mix. These models allow us to compare similar episodes, outputting risk adjustments expressed in our expected allowed amounts.

Machine learning variables for expected allowed amounts, Table 5

Expected variables	All specialties (exceptions adjacent)	Exceptions
Age	18 year of age or older	All ages (*PCP, Commercial* program only)
Business line	Commercial	Medicare (*PCP, Medicare* program only)
ETG code	ETG Codes	
Concurrent episodes	Two or more episodes with different medical conditions that overlap on one or more days	Cardiology and Spine surgery programs only
Episode days	Length of an episode in days	
Episode volume minimum	200 or more episodes per ETG to generate expected values	
ERG risk score	Optum Symmetry® Episode Risk Groups® (ERG®)	
Gender	Female or male	
Hospital referral region (HRR)	HRR number level (See Appendix)	
Health profile database conditions	All health profile database (HPD) conditions (See Appendix)	Cardiology and Spine surgery programs only
ICD10 group codes	Frequency of ICD10 groups codes in an episode	
Pharmacy usage	With or without an Aetna pharmacy plan	
Severity level	As indicated on the episode based on Optum Symmetry® (severity increases from 0 to 4)	
Social Determinants of Health (SDoH)	Variables that impact health care cost in a census tract area: income, poverty, diversity, disability, education, physical inactivity, marital status, public transportation (CDC and US Census)	
Practice type	Gynecologists or Obstetricians Neurosurgeons or Orthopedic surgeons	Ob/Gyn and Spine surgery programs only
Timing of entry into an episode	Percentage of allowed amount that occurs in the episode before the practice's first claim	
Year of episodes (2023 program)	2 years: January 1, 2020 – December 31, 2021	January 1, 2018 – December 31, 2019 for Cardiology and Spine surgery programs only (2022 programs)

We make an additional adjustment to actual allowed amounts to ensure we are capturing economic differences across hospital referral regions. We refer to this adjustment as a market-adjusted allowed amount. In Table 4 below, we describe the variables for market adjustment by program.

Market adjustment variables, Table 6

Programs	Episodes or Population	Practice risk tier	Practice size	HRR
Cardiology		Low if practice average risk score is less than or equal to the 80th percentile All others are high risk	Small if episode count is less than or equal to the 80th percentile All others are not small	✓
Endocrinology		Low if practice average risk score is less than or equal to the 50th percentile All others are high risk	Small if episode count is less than or equal to the 50th percentile All others are not small	✓
Ob/Gyn	Gynecology	Low if practice average risk score is less than or equal to the 50th percentile All others are high risk	Small if episode count is less than or equal to the 80th percentile All others are not small	✓
	Obstetrics			✓
Ortho hip & knee	Hip			✓
	Knee	Low if practice average risk score is less than or equal to the 80th percentile All others are high risk	Small if episode count is less than or equal to the 80th percentile All others are not small	✓
PCP, Commercial	Adult, Chronic & Non-Chronic	Low if practice average risk score is less than or equal to the 40th percentile All others are high risk	Small if episode count is less than or equal to the 80th percentile All others are not small	✓
	Pediatrics	Low if practice average risk score is less than or equal to the 50th percentile All others are high risk	Small if episode count is less than or equal to the 50th percentile All others are not small	✓
PCP, Medicare		Low if practice average risk score is less than or equal to the 50th percentile All others are high risk	Small if episode count is less than or equal to the 80th percentile All others are not small	✓
Pulmonary medicine				✓
Spine surgery	Back	Low if practice average risk score is less than or equal to the 50th percentile All others are high risk	Small if episode count is less than or equal to the 50th percentile All others are not small	✓
	Neck	Low if practice average risk score is less than or equal to the 80th percentile All others are high risk	Small if episode count is less than or equal to the 80th percentile All others are not small	✓

Below is an example of the data you can find in our reporting packages. We use the market adjustment variables to create the market adjusted allowed (C; Table 7). We use this amount to create the practice's performance index per ETG ($\frac{C}{D} = E$; Table 7).

Practice performance example for episodes of care, Table 7

A	B	C	D	E
ETG Description	**Actual Allowed Per Episode**	**Market Adjusted Allowed Per Episode**	**Expected Allowed Per Episode**	**Performance Index**
Diabetes, with complication, with comorbidity, w/o surgery	$9,165.97	$9,461.07	$12,481.41	0.76
Rhinitis, allergic & non-allergic, w/o surgery	$131.88	$129.98	$127.14	1.02
Tonsillitis, adenoiditis or pharyngitis, w/o complication, w/o comorbidity, w/o surgery	$139.73	$137.72	$111.76	1.23

The practice's overall performance index (all ETGs) is calculated as the weighted average of E, where the expected allowed amounts in D weight E ($\frac{(E_2 \times D_2)+(E_3 \times D_3)\ldots}{D_2 + D_3 + \ldots}$).

To evaluate performance, we conduct a two-sample weighted t-test. We weight using mean allowed amounts from the relevant machine learning nodes. For the first sample, we weight the practice's market-adjusted performance by the mean of the practice node's allowed amounts. For the second sample, we weight the market's average market-adjusted performance index (from the relevant decision tree node) with the mean of the market node's allowed amounts. Both weighted samples are compared in the t-test:

- If the test has a p-value greater than 0.10, the practice results are not statistically significant, and the practice outcome will be "performance not statistically significant."
- If the practice's market-adjusted performance index is statistically significant and less than the average market-adjusted performance index, the practice will receive the outcome "subcategory earned."
- If the practice market-adjusted performance index is statistically significant and greater than the average market-adjusted performance index, the practice outcome will be "criteria not met."

4.1. Effectiveness: other subcategories

We often use multiple subcategories under the effectiveness designation. There are sometimes multiple *episodes of care* subcategories by population or specialty as well.

For instance, in the PCP, Commercial program there are:

- Two population subcategories for episodes of care
    - Chronic
    - Non-chronic
- Two pairs of effectiveness and clinical quality categories
    - One for members younger than 18 (pediatrics)
    - One for members 18 or older (adults).

To learn more about less common subcategories, see program guides below (sections 6-13). Below is summary of all effectiveness subcategories we use in our programs.

Effectiveness subcategories in Aetna Smart Compare (2023), Table 8

Programs	Cotiviti Risk-Readiness®	Population effectiveness	Population episodes of care	Procedure episodes of care	Specialty episodes of care	Utilization (Risk Adjusted)
Cardiology	✓					
Endocrinology						
Ob/Gyn					Gynecology, and Obstetrics	
Ortho hip & knee	✓			Hip, and knee treatments		
PCP, Commercial		<18 years of age, and ≥18	Chronic, and Non-Chronic			✓
PCP, Medicare			Chronic, and Non-Chronic			✓
Pulmonary medicine						
Spine surgery	✓			Back, and neck surgery		

4.2. Effectiveness: risk-adjusted utilization

We include risk-adjusted utilization measures in our two Primary care physician (PCP) programs, Commercial and Medicare. These measures assess either the use of specific resources or the outcomes associated with population management. Primary care practices play an important role in improving the health of the populations they manage. For instance, inpatient utilization and emergency room trends decline as the health of a population improves through the better management of chronic conditions. Also, as primary care practices coordinate care for members and provide education to members on treatment guidelines, the overuse or misuse of high-tech radiology declines. Our three utilization measures focus on these observations:

- Inpatient admission (IP) per 1,000 members
- Emergency room (ER) visits per 1,000 members
- MRI & CT scan utilization per 1,000 members

We risk-adjust our utilization measures by creating expected rates per 1,000. But we do not identify high-cost claimants for removal as outlined in PCP, Commercial (section 10) or PCP, Medicare (section 11). Expected rates assume the following variables in Table 9.

Variables for calculating expected rates for risk-adjusted utilization, Table 9

Expected variables	PCP designations (commercial and Medicare)
Age	≥ 18 for Commercial Adults and Medicare; <18 for Commercial Pediatrics
Business line	Commercial or Medicare
Practice risk tier	High or low
Practice size	Large or small
Geography	Urban and suburban compared to rural

Physician practices evaluated by our risk-adjusted utilization measures may request reports like the following.

Practice performance example for risk-adjusted utilization, Table 10

A	B	C	D	E	F
Measure	**Average member risk score**	**Actual rate per 1,000**	**Adjusted rate per 1,000**	**Expected rate per 1,000**	**Performance index**
Emergency room visits per 1,000	1.503	101.89	101.90	113.92	0.89
Inpatient admits per 1,000	1.503	21.39	21.41	24.42	0.88
MRI/CT scan utilization per 1,000	1.503	57.62	57.65	68.19	0.85

The three utilization measures are listed in the first column (A). The average member risk score (B) is derived from the Optum Symmetry® Episode Risk Groups® (ERG®): $\frac{\sum(ERG\ scores\ per\ member\ per\ month)}{all\ member\ months}$. The actual rate per 1,000 (C) is $\frac{\frac{event\ count}{(\frac{average\ member\ count}{1000})}}{average\ member\ risk\ score}$. The adjusted rate per 1,000 (D) is a Bayesian transformation (see below for an explanation of our most common transformations in section 5.1.) The expected value is calculated as the average adjusted rate per 1,000 within the peer groups defined by the variables in table 9. The performance index is calculated as $\frac{D}{E}$. We weight each measure according to its overall contribution to medical spend, and aggregate performance indices as a weighted average (see the weights in Table 11).

We run a t-test to identify practices that are statistically significant to award designation. The t-test indicates whether there is a statistically significant difference between the practice's index and the expected performance index.

- If the test result is P>.05, the practice results are not statistically significant, so the practice will receive the subcategory result "performance not statistically significant."
- If the test result is P<.05, and the practice's weighted performance index is <1.0, the practice will receive the subcategory result "subcategory earned."
- If the test result is P<.05, and the practice's weighted performance index is >1.0, the practice will receive the subcategory result "criteria not met."

Specifications for risk-adjusted utilization measures, Table 11

Eligibility specifications	IP admits per 1,000[4]	ER visits per 1,000[5]	MRI & CT scans per 1,000[6]
Business lines	Commercial or Medicare	Commercial or Medicare	Commercial or Medicare
Age	≥ 18 for Commercial adults and Medicare; <18 for Commercial Pediatrics	≥ 18 for Commercial adults and Medicare; <18 for Commercial Pediatrics	≥ 18 for Commercial adults and Medicare; <18 for Commercial Pediatrics
Continuous enrollment	Not applicable	Not applicable	Not applicable
Measurement period	12 months (2021)	12 months (2021)	12 months (2021)
Benefit	Full medical	Full medical	Full medical
Measure specifications	**IP admits per 1,000[4]**	**ER visits per 1,000[5]**	**MRI & CT scans per 1,000[6]**
Numerator	Count of acute IP hospitalizations	Count of ER visits to an acute care facility	Count of outpatient MRI and CT scan studies
Denominator	All attributed members	All attributed members	All attributed members
Performance index weights	3	2	1

4.3. Effectiveness: Cotiviti Risk-Readiness®

Introduction to Cotiviti Risk-Readiness

We use Cotiviti's Risk-Readiness tool to provide an objective measure in some of our Aetna Smart Compare programs. Risk-Readiness is an efficiency measurement tool that provides insights to identify and quantify sources of low-value care in Medicare claims. Cotiviti defines low-value care as health care services that increase medical spend without improving population health. Cotiviti Risk-Readiness benchmarks are based on provider geography and peer types, using leading industry research and guidelines. This includes the Dartmouth Atlas for Unwarranted Variation, the American Board of Internal Medicine (ABIM)'s Choosing Wisely initiative, and other proprietary clinical analyses. Above is a table of the programs that include Cotiviti Risk-Readiness® (see section 4.1; Table 8.)

Risk-Readiness methodology

Cotiviti applies its clinical methodology to Medicare data to create actionable benchmarks. The benchmarks identify outliers in practice patterns among a cohort of physician peers. These can indicate where a provider group may deliver more or less of the care influenced by physician preference. Below is an outline of the tool's parameters:

- Cotiviti uses six years of Medicare parts A, B, and D claims data to develop a large, multi-year, normalized dataset on more than 90 percent of all physicians. At the time of measurement, Cotiviti uses one calendar year of

4 This measure calculates the number of acute inpatient admissions per 1,000 members per year.
5 This measure calculates the rate of emergency room visits per 1,000 members per year.
6 This measure calculates the rate of MRI and CT-scan utilization per 1,000 members per year.

Medicare claims informed by their longitudinal study. 2023 programs in Aetna Smart Compare use 2019 claims data, and 2022 programs use 2018 data.

- Proprietary algorithms identify provider sub-specialties and create peer cohorts in each geography. Hospital referral regions (HRRs) are used to define peer cohort geography.
- Evidence from academic research — Choosing Wisely, etc — helps Cotiviti assess clinical quality and medical economics.
- Measures are risk-adjusted based on the provider's patient panel using the Medicare Hierarchical Condition Categories (HCC) risk-adjustment methodology.
- Provider scoring in each measure reflects their relative performance within their peer cohorts. They are placed into quintiles based on the relative intensity of visits, services, prescriptions and referrals they provide.
- Measure sets are specialized, using the ETG codes we identify for the specialties in each Aetna Smart Compare designation.
- Cotiviti's Risk-Readiness® tool evaluates provider groups on 4 domains and a score is created for each domain.
    - Visit score: measures how quickly provider visits escalate compared to peers
    - Procedure score: measures how intensely a provider practices medicine compared to peers
    - Pharmacy score: measures how a provider prescribes medications compared to peers
    - Referral score: measures performance and appropriateness in a provider's value chain compared to peers

Cotiviti aggregates the four domain scores into a composite score, known as the "overall value score."

Provider practices we score in the top 20 percent have practice patterns that align with high-value care delivery, receiving a result of "subcategory earned." To other practices, we assign an outcome of "criteria not met – external data," unless they have no episodes to measure, then we assign "volume insufficient."

5. Clinical quality category: HEDIS® measures

To score many of our programs on clinical quality, we use the most recent measures from Healthcare Effectiveness Data and Information Set (HEDIS). The National Committee for Quality Assurance (NCQA) creates the measure specifications[7]. We use HEDIS measures to evaluate the quality of preventive care, the outcomes of curative care, and the management of chronic conditions. We code our HEDIS measures using NCQA's HEDIS specifications to align with industry accepted standards.

We use subsets of the available HEDIS measures to customize our programs for the specialties and/or procedures they contain. Below is table of the HEDIS measures we currently use.

HEDIS measures in Aetna Smart Compare (2023), Table 12

ID	HEDIS measures	Program	Specialty[8]	PCP populations	Business line[9]
CWP	Appropriate testing for pharyngitis	PCP	FP, I, P	Pediatrics	Commercial
URI	Appropriate treatment for upper respiratory infection	PCP	FP, I, P	Adult (non-chronic), pediatrics	Commercial
BCS	Breast cancer screening	PCP, Ob/Gyn	FP, Gyn, I,	Adult (non-chronic)	Commercial, Medicare

7 They are available for purchase at store.ncqa.org/index.php

8 C: Cardiologist; E: Endocrinologist; FP: Family practice physician; Gyn: Gynecologist; I: Internal medicine physician; Ob: Obstetrician; P: Pediatrician; Plm: Pulmonologist

9 Currently only the PCP program is available for Commercial or Medicare.

ID	HEDIS measures	Program	Specialty[8]	PCP populations	Business line[9]
CRE	Cardiac rehabilitation	Cardiology	C		Commercial
COA	Care for older adults	PCP	FP, I	Adult (chronic)	Commercial
CCS	Cervical cancer screening	PCP, Ob/Gyn	FP, Gyn, I	Adult (non-chronic)	Commercial
CDC[10]	BP control (<140/90)	Endocrinology, Cardiology	E, C		Commercial
CDC[10]	HbA1c poor control (>9.0%)	Endocrinology, PCP	E, FP, I	Adult (chronic)	Commercial, Medicare
CDC[10]	HbA1c control (<8.0%)	Endocrinology	E		Commercial
CDC[10]	Hemoglobin a1c testing	PCP	FP, I	Adult (chronic)	Commercial
CDC[10]	Retinal eye exam	Endocrinology, PCP	E, FP, I	Adult (chronic)	Commercial, Medicare
WCV	Child and adolescent well-care visits	PCP	FP, I, P	Pediatrics	Commercial
CHL	Chlamydia screening in women	Ob/Gyn	Gyn		Commercial
COL	Colorectal cancer screening	PCP, Ob/Gyn	FP, Gyn, I	Adult (non-chronic)	Commercial, Medicare
CBP	Controlling high blood pressure	Cardiology	C		Commercial
CBP	Controlling high blood pressure for sleep apnea	Pulmonary medicine	Plm		Commercial
FMC	Follow-up after emergency department visit for people with multiple high-risk chronic conditions	PCP	FP, I		Medicare
KED	Kidney health evaluation for patients with diabetes	Endocrinology, PCP	E, FP, I, P	Adult (chronic)	Commercial
IMA	Immunizations for adolescents	PCP	FP, I, P	Pediatrics	Commercial
OMW	Osteoporosis management in women who had a fracture	PCP	FP, I		Medicare
PBH	Persistence of beta-blocker treatment after a heart attack	Cardiology	C		Commercial

10 CDC: comprehensive diabetes control

ID	HEDIS measures	Program	Specialty[8]	PCP populations	Business line[9]
PCE	Pharmacotherapy management of COPD exacerbation	Pulmonary medicine	Plm		Commercial
PPC	Postpartum care	Ob/Gyn	Ob		Commercial
SPC	Statin therapy for patients with cardiovascular disease (we measure both the dispensing rate and adherence)	Cardiology	C		Commercial
SPD	Statin therapy for patients with diabetes (we measure both the dispensing rate and adherence)	Endocrinology, PCP, Cardiology	E, C, FP, I	Adult (chronic)	Commercial
PPC	Timeliness of prenatal care	PCP	Ob		Commercial
DAE	Use of high-risk medications in older adults	PCP	FP, I	Adult (chronic)	Commercial
SPR	Use of spirometry testing in the assessment and diagnosis of COPD	Pulmonary medicine	Plm		Commercial
WCC	Weight assessment and counseling for nutrition and physical activity for children and adolescents	PCP	FP, I, P	Pediatrics	Commercial
W30	Well-child visits in the first 30 months of life	PCP	FP, I, P	Pediatrics	Commercial

5.1 Clinical quality: Bayesian transformations and performance indices

We calculate our HEDIS measures with an empirical Bayesian transformation. It assumes actual practice performance as a prior condition while seeking to transform the results of the smallest, mostly highly variable practices towards observations of national performance and variance. These adjustments prevent us from overstating or understating performance below a confidence interval of 90 percent. This holds even where we have few of the practice's patients to review for a given measure. Practices with few Aetna members should expect larger adjustments per measure. Practices with many Aetna members may see little to no adjustment for most — if not all — measures.

We transform our HEDIS measures by calculating the weighted national average for the measure, as well as the weighted national average variance for each measure. We also calculate variance within the practice we are measuring. By using national variance and practice variance as additional observations for actual practice performance, we can adjust this performance toward the national average, according to the degree of practice variance as compared to the national average variance for each measure. In short, our calculations regress actual practice performance toward the national average, whether that practice performance is greater or lower than national average, especially where practice variance is high.

Practice performance example for HEDIS® measures, Table 13

A	B	C	D	E
HEDIS® description	**Actual rate**	**Adjusted rate**	**Expected rate**	**Performance index**
Diabetes: Statin adherence 80%	0.67	0.69	0.78	0.88
Eye exam (retinal) performed	0.54	0.53	0.38	1.38
Functional status assessment	0.39	0.38	0.13	2.96

Practices that request reporting packages from us will receive reports that resemble Table 13 above. Actual rate (B) is the actual practice's performance for the respective HEDIS® measure: $\frac{HEDIS\ numerator}{HEDIS\ denominator}$. The adjusted rate (C) is the Bayesian transformation of the actual rate (B). The expected rated (D) is the result of peer group analysis, calculated as $\frac{(lower\ bound\ of\ the\ peergroup)+(upper\ bound\ of\ the\ peer\ group)}{2}$. We calculated the performance index (E) of each measure as $\frac{C}{D}$.

Once we determine the performance index per measure, we evaluate that performance by normalizing the measure, where the normalized performance index per measure is equal to $\frac{(peformance\ index)-(peer\ group\ mean)}{peer\ group\ standard\ deviation}$. This notation (performance index, peer group mean, and peer group standard deviation) is calculated for each measure (A). To determine the aggregate performance of measures under a clinical quality subcategory, we average all normalized performance indexes. This average is compared to 0, and a statistical test is run to identify how the practice compares to its peers.

- If the test result is P>.10, the practice results are not statistically significant, so the practice will receive the subcategory outcome "performance not statistically significant."
- If the test result is P<.10, and the practice's weighted performance index is >0, the practice will receive the subcategory outcome "subcategory earned."
- If the test result is P<.10, and the practice's weighted performance index is <0, the practice will receive the subcategory outcome "criteria not met."

5.2 NCQA Recognition

In our PCP, Commercial and Medicare programs, we value NCQA recognition as a subcategory, where we observe 75 percent of members are attributed to physicians and practitioners who are recognized by NCQA's Patient-Centered Medical Home (PCMH) Recognition program[11].

The program addresses several PCMH concepts:

- Team-based care and practice organization helps structure a practice's leadership, care team responsibilities and how the practice partners with patients, families and caregivers.
- Knowing and managing your patients sets standards for data collection, medication reconciliation, evidence-based clinical decision support and other activities.
- Patient-centered access and continuity: guides practices to provide patients with convenient access to clinical advice and helps ensure continuity of care.
- Care management and support help clinicians set up care management protocols to identify patients who need more closely-managed care.

11 ncqa.org/programs/health-care-providers-practices/patient-centered-medical-home-pcmh/

- Care coordination and care transitions ensures that primary and specialty care clinicians are effectively sharing information and managing patient referrals to minimize cost, confusion and inappropriate care.
- Performance measurement and quality improvement: improvement helps practices develop ways to measure performance, set goals and develop activities that will improve performance.

Under these concepts the program contains 40 required criteria, and an additional 25 credits of elective criteria in each of the concepts.

When we observe practices meeting the 75 percent threshold, we assign "subcategory earned." We assign all other practices "criteria not met – external data," unless they have no physicians with NCQA recognition, then we assign "volume insufficient."

6. Our Cardiology program

The commercial Cardiology program measures episodes of general cardiology. The program includes all subspecialties of cardiology. We use episodes of care and Cotiviti Risk-Readiness® in the effectiveness category. We use HEDIS measures in the clinical quality category.

This program is for Commercial members, but we do use both Commercial and Medicare claims in the clinical quality category. We use only Commercial members in the effectiveness category. You can find more detailed information above on the various measurements (see sections 3, 3.1, 3.3, 4, 4.3, 5, and 5.1).

Thresholds

To be part of the program, cardiology practices must see enough adult members (18 years of age and older) for us to perform a statistically valid evaluation. For measures in the episodes of care category, we evaluate practices with at least five inlier episodes of care in 2018-19. For measures in Cotiviti Risk-Readiness, we look at practices with at least one inlier episode of care in 2018. For measures in the clinical quality category, we only use quality measures that have at least four members in the denominator, using 2019 claims data.

Involved Episodes

For clinical quality measures we use involved episodes in addition to standard episodes. Involved episodes are defined this way:

- The episode's responsible provider is not a cardiologist (episode belongs to another specialist).
- At least one of the episode's claims includes care from at least one of the practice's cardiologists.
- The service location of the cardiology claim is an office or outpatient setting.

High-cost claimants

We remove high-cost claimants (HCC) from any category. If a member has claim costs (total allowed amount) per year in the 99th percentile in either 2018 or 2019, we excluded the member from all analysis. Below is the HCC threshold by claim year.

High-cost claimants in 2018-19, Table 14

Commercial Population	Year	HCC threshold (99th percentile)
Adults	2018	$80,125
Adults	2019	$82,895

Designation hierarchies

Below is a summary of the designations we administer in this program.

Commercial clinical quality for the Cardiology program, Table 15[12]

Clinical quality for Cardiology

Designated = 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategory		
HEDIS Measures		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

Commercial effectiveness for the Cardiology program, Table 16[12]

Effectiveness for Cardiology

Designated ≥ 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories		
Episodes of care		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Cotiviti Risk-Readiness®		
	Subcategory earned	1
	Volume insufficient	0
	Criteria not met - external data	0

7. Our Endocrinology program

Thresholds

In the commercial Endocrinology program, we measure episodes of care in the effectiveness category. We use HEDIS measures in the clinical quality category.

You can find more detailed information above on the various measurements by reading sections 3, 3.1, 3.3, 4, 5, and 5.1.

To be part of the program, endocrinology practices must see enough adult members for us to perform a statistically valid evaluation. For measures in the episodes of care category, we evaluate practices with at least five inlier episodes of care in 2020-21. For measures in the clinical quality category, we only use quality measures that have at least three members in the denominator. To be measured, practices must have five measures that meet the denominator threshold. We use claims from the 2021 calendar year for all quality measures.

[12] In 2022 programs, we communicated letters with "insufficient information" for either "volume insufficient" or "performance not statistically significant." Additionally, we used "designation earned" instead of "subcategory earned." For more detail on the 2022 programs, please review the cardiology methodology guide on **Aetna.com/smartcompare**.

Involved episodes

For effectiveness and clinical quality measures we use involved episodes in addition to standard episodes. Involved episodes are defined this way:

- The episode's responsible provider is not an Endocrinologist, and
- An Endocrinologist enters the episode before ten percent of the total allowed is spent, and
- This Endocrinologist owns 30 percent or more of the episode's physician office visits.

High-cost claimants

We remove high-cost claimants (HCC) from any category. If a member has claim costs (total allowed amount) per year in the 99th percentile in either 2020 or 2021, we excluded the member from all analysis. Below is the HCC threshold by claim year and population:

High-cost claimants in 2020-21, Table 17

Commercial Population	Year	HCC threshold (99th percentile)
Adults	2020	$81,092
Adults	2021	$85,337

Designation hierarchies

Below is a summary of the designations we administer in the Endocrinology program.

Commercial clinical quality for the Endocrinology program, Table 18

Clinical quality for Endocrinology

Designated = 1 > Not designated

Summary of subcategories

Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategory

HEDIS Measures

Subcategory earned	1
Volume insufficient	0
Performance not statistically significant	0
Criteria not met	-1

Commercial effectiveness for the Endocrinology program, Table 19

Effectiveness for Endocrinology

Designated = 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategory		
Episodes of care		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

8. Our Ob/Gyn program

The Commercial Ob/Gyn program measures practices that focus on obstetrics and/or gynecology. Measures are split into either specialty, which serve as the subcategories in this program, where we include all physicians practicing as obstetricians and/or gynecologists. In the effectiveness category, we measure obstetric episodes of care and gynecologic episodes of care. In the clinical quality category, we use HEDIS measures in either specialty as subcategories, except in the obstetric category, where we also include a low-risk C-section measure endorsed by the Society for Maternal-Fetal Medicine (SMFM) (see *additional measures* below).

You can find more detailed information on the various measurements by reading sections 3, 3.1, 3.3, 4, 5, and 5.1.

Thresholds

To be part of the program, physician practices must see enough members for a statistically valid evaluation. For obstetric effectiveness, practices must have 6 or more inlier pregnancies. For gynecologic effectiveness, practices must have 11 or more inlier episodes. We calculate all effectiveness measures with 2020-21 claims data.

For clinical quality measures, we only evaluate measures with 3 or more members in the denominator, using 2021 claims data. An Ob/Gyn practice must have at least 3 valid clinical quality measures to receive an overall quality designation.

High-cost claimants

We remove high-cost claimants (HCC) from analysis. If a member has claim costs (total allowed amount) per year greater than the 99th percentile in either 2020 or 2021, the member is excluded from analysis (the entirety of the member's claims; not only the claims that meet the threshold). Below is the HCC threshold by claim year.

High-cost claimants in 2020-21, Table 20

Commercial population	Year	HCC threshold (99th percentile)
Adults	2020	$81,092
Adults	2021	$85,337

General Ob/Gyn practices

Only general Ob/Gyn practices are eligible for this designation. Physicians classified as general Ob/Gyns are identified using a machine learning technique called clustering. This technique involves inputting variables into clustering or grouping algorithms to find similarities in large datasets. Using 2021 claims data, we apply this technique to all Ob/Gyn claims.

Before we run clustering algorithms, we extract diagnosis and procedure codes and transform them into data arrays with an index score based on global co-occurrence or the likelihood of their appearing together. The arrays and associated scores are then aggregated at a provider level using a weighted average of the associated claim count.

After we aggregate and weight provider data, we input this data into a density-based, hierarchical clustering algorithm. This algorithm is used to identify groupings of providers who have similar diagnosis and procedure group clusters based on similar scores (i.e., low distance from cluster averages). The following table shows the clusters identified by this algorithm. All providers appearing in "general Ob/Gyn" (cluster 5) were included in this Ob/Gyn program.

Clusters from machine learning analysis of Ob/Gyn claims, Table 21

Specialty	**Reproductive endocrinology and infertility**	**Female pelvic medicine and reconstructive surgery**	**Maternal-fetal medicine**	**Gynecologic oncology**	**General Ob/Gyn**	**All others**
Cluster sets	**Cluster 1**	**Cluster 2**	**Cluster 3**	**Cluster 4**	**Cluster 5**	**Outliers**
Total provider count	1,485 (4%)	976 (3%)	4,849 (13%)	516 (1%)	29,217 (78%)	302 (1%)
Claim count (K)	1,285 (15%)	114 (1%)	908 (8%)	37 (1%)	6,191 (72%)	80 (1%)
Allowed Amount ($M)	256 (14%)	23 (2%)	179 (12%)	15 (1%)	1,295 (73%)	8 (1%)

Additional measures

Our Ob/Gyn program includes a measure to evaluate low-risk C-sections, based on a similar measure supported by the Society for Maternal-Fetal Medicine. Below are the specifications for the measure. We calculate this measure according to these specifications, which otherwise conform to our calculations of HEDIS measures, including Bayesian transformation (see sections 5 and 5.1).

Low-risk C-section measure from the clinical quality designation for obstetrics, Table 22

Cesarean section rate in singleton low-risk deliveries					
Description	This measure calculates the percentage of members who do not have a high-risk medical indication requiring a Cesarean section, and who deliver a term, singleton live birth in a vertex position by Cesarean				
Denominator	Members that deliver a term, singleton live birth in a vertex position, without a medical indication of high risk (see Appendix I) by an Ob/Gyn provider and assigned a delivery (DRGs: 765, 766, 767, 768, 774, 775, 783, 784, 785, 786, 787, 788, 796, 797, 798, 805, 806, 807)				
Numerator	Members in the denominator that deliver by cesarean section (DRGs: 765, 766, 783, 784, 785, 786, 787, 788)				
Product Lines	Commercial	**Benefit**	Full medical	**Risk Adjustment**	None, but see 5.1 for transformation
Continuous Enrollment	Member must be active as of the hospitalization admission date through the hospitalization discharge date within the 12-month assessment period				
Attribution	Ob/Gyn	**Population**	Adults	**Measure type**	Obstetric
Data	Aetna claims data from measurement year plus three months of claims runout				

Designation hierarchies

Below is a summary of the designations we administer in this program.

Commercial clinical quality for the Ob/Gyn program, Table 23

Clinical quality for Gynecology and/or Obstetrics

Designated ≥ 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories		
Gynecologic HEDIS Measures	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Obstetric HEDIS Measures, and Low-risk C-section rate	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

Commercial effectiveness for the Ob/Gyn program, Table 24

Effectiveness for Gynecology and/or Obstetrics

Designated ≥ 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories		
Gynecologic episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Obstetric episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

9. Orthopedic hip & knee program

This Commercial program for orthopedists evaluates knee episodes of care, hip episodes of care, and Cotiviti Risk-Readiness® in its effectiveness category. The program also looks at surgical facility measures in the clinical quality category. In this program, we only measure orthopedists who treat hip and/or knee degenerations. You can find more detailed information above on the various measurements (see sections 3, 3.1, 3.3, 4, 4.1, 4.3, and 9.1).

Thresholds

For measurement purposes, physician practices must see enough members for us to perform a statistically valid evaluation of a physician's practice patterns. We only measure orthopedic practices with 11 or more inlier episodes in either hip or knee treatments for the effectiveness measure. That means 11 in each – 22 or more – for

both measures. In effectiveness, we measure using 2020-21 commercial claims. For more information about quality, see section 9.1 below.

To measure Cotiviti Risk-Readiness®, we require at least one episode in 2019

To measure clinical quality, physician practices must have three or more total episodes to be evaluated (2020-21).

Pediatric practices

This program does not measure pediatric orthopedics, so we exclude practices where 51% or more of hip and knee conditions are managed for members younger than 18 years of age. We also exclude episodes for all members younger than 18.

Designation hierarchies

Below is a summary of the designations we administer in this program.

Commercial clinical quality for the Orthopedic hip & knee program, Table 25

Clinical quality for Orthopedic hip & knee

Designated = 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategory		
Facility measures (CMS Care Compare, and either Joint Commission or DNV)	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

Commercial effectiveness for the Orthopedic hip & knee program, Table 26

Effectiveness for Orthopedic hip & knee

Designated ≥ 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories		
Hip episodes of care		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Knee episodes of care		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Cotiviti-Risk Readiness®		
	Subcategory earned	1
	Volume insufficient	0
	Criteria not met - external data	0

9.1 Orthopedic hip and knee program: clinical quality

In the orthopedic hip and knee program, the clinical quality category looks at facilities where orthopedic specialists perform hip or knee replacement procedures. Complication rates and compliance with clinical guidelines affect this measure.

We only measure practices where the majority of surgical spend is within the respective member's network plan (in-network; not out-of-network). We also exclude practices that use facilitates with confirmed fraudulent activity. Practices that fail to meet one of these criteria will receive the "volume insufficient" designation.

CMS measures

The clinical quality designation is based on publicly available external data on facilities where the orthopedic practice performs hip or knee procedures. We include groups with three episodes in Commercial and Medicare populations. We use two measures from Centers for Medicare & Medicaid Services (CMS) Care Compare program for comparing hospitals:

- 30-day readmission rate following elective hip or knee arthroplasty
- Hospital risk adjusted complication rate following elective hip or knee arthroplasty

We evaluate each of the CMS measures by calculating a performance index as the ratio between actual and expected facility performance for each measure. We link orthopedic practices to facilities based on 2021 episodes. If an orthopedic practice performs cases at more than one facility, we use a weighted average score based on Aetna case count to determine practice-level performance. Each measure is scored separately, reaching one of the following outcomes:

- We conclude "positive performance" if practice performance is lower than the index average less 0.5 standard deviations.
- We conclude "negative performance" if practice performance is higher than index average plus 0.5 standard deviations.
- We conclude "neutral performance" otherwise.

We combine the performance of the CMS measures according to the following diagram.

CMS Care Compare (hospital) measures, combined performance, Table 27

		Hospital risk-adjusted complication rate following elective hip or knee arthroplasty		
		Positive performance	**Neutral performance**	**Negative performance**
30-day Readmission following elective hip or knee arthroplasty	**Positive performance**	Positive performance	Positive performance	Positive performance
	Neutral performance	Positive performance	Neutral performance	Neutral performance
	Negative performance	Positive performance	Neutral performance	Negative performance

External accreditations

In addition to the CMS measures, we consider two accreditations:

1. The Joint Commission Advanced Total Hip and Knee Replacement Certification
2. Det Norske Veritas (DNV) Healthcare Total Hip and Knee Replacement Certification

Either accreditation is available to hospitals and ambulatory surgical centers. We link the accreditation to practices that use an accredited facility for 25 percent or more episodes in 2019.

The Joint Commission Certification is based on several outcomes. These include surgical site infections, fall rates, improvement in patient education prior to discharge, length of stay, early ambulation and pain management. To be eligible for the certification, a facility must have served a minimum of 10 patients.

The DNV Healthcare Hip & Knee Replacement Program Certification (HKRPC) integrates requirements related to the CMS Conditions of Participation for hospitals. HKRPCs recognize excellence in orthopedic surgery within the scope of hip and knee replacement, and any related procedures. The certification means that a hospital has shown it complies with the DNV Healthcare standard and adherence to guidelines of the American Academy of Orthopedic Surgeons (AAOS).

Integrating CMS measures and external accreditations

We integrate the CMS Measures and one of the accreditations (from the Joint Commission or DNV) to create the results for the quality category. See the table below to understand how designations are awarded based on the measures and accreditations outlined above.

Clinical quality measure for the Orthopedic hip & knee program, combined performance, Table 28

		CMS Care Compare		
		Positive performance	**Neutral performance**	**Negative performance**
Joint Commission Certification or DNV Healthcare	**Positive performance**	Designation earned	Designation earned	Designation earned
	Neutral performance	Designation earned	Performance not statistically significant	Performance criteria not met

10. Primary care physician (PCP), Commercial program

The Commercial program for PCPs evaluates episodes of care and risk-adjusted utilization measures in its effectiveness category. In the clinical quality category, we evaluate HEDIS measures and NCQA Recognition. The current program measures our claims from 2020-21. We measure family practice, internal medicine, and pediatric practices, including physicians' assistants and nurse practitioners managed by these physicians. You can find more detailed information above on the various measurements (see sections 3, 3.1, 3.2, 3.3, 4, 4.2, 5, and 5.1).

Populations and thresholds

To be included in the program, physician practices must see enough members for us to perform a statistically valid evaluation. We evaluate adult and pediatric populations separately. (Pediatric populations are less than 18 years of age.) We measure family practice and internal medicine physicians on each population, where the practice mees the minimum member criteria.

To measure episodes of care, we require no less than 20 valid inlier episodes for each population in 2020-21 claims data. To measure risk-adjusted utilization, we require 25 attributed, valid members in 2021 claims data. To measure clinical quality, we must have 5 valid members in the denominator of each measure to evaluate practices on those measures, using 2021 claims data. To be evaluated for each clinical quality population (adult or pediatric), we must be able to evaluate practices on 2 HEDIS measures.

Here are the rules for attribution and member inclusion:

- Valid members are members not attributed to any other practices during calendar year 2021. They also have at least six months of medical benefit eligibility in calendar year 2021. For risk-adjusted utilization and quality measures, we include only members attributed to a primary care practice at the end of 2021.
- Attribution methodology uses our standard attribution logic, which considers both volume and recency of claims. We do measure members receiving care from nurse practitioners and physician assistants managed by primary care practices (see sections 3.1 and 3.2).

Involved episodes

- In addition to standard episodes, we add involved episodes for episodes of care measures in effectiveness. Here are the rules we use to identify involved episodes:
    - The episode's responsible provider is not a PCP (or nurse practitioner, or physician assistant managed by a PCP), and
    - At least one of the episode's claims are serviced by PCP (or nurse practitioner, or physician assistant managed by a PCP), and
    - The PCP (or nurse practitioner or physician assistant) enters the episode before 10 percent of the episode's allowed amounts are spent.

High-cost claimants

- High-cost claimants (HCC) are removed from the analysis. If a member has claim costs (total allowed amount) per year greater than the 99th percentile in either 2020 or 2021, we exclude them from analysis (the entirety of the member's claims, not only the claims that meet the threshold). Below is the HCC threshold by claim year and population.

High-cost claimants in 2020-21, Table 29

Commercial population	Year	HCC threshold (99th percentile)
Adults	2020	$81,092
Pediatrics	2020	$29,247
Adults	2021	$85,337
Pediatrics	2021	$34,656

Designation hierarchies

Below is a summary of the designations we administer in this program.

Commercial adult clinical quality for PCPs, Table 30

Clinical quality for PCP, Commercial, adults

Designated ≥ 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories		
Chronic HEDIS Measures	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Non-Chronic HEDIS Measures	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
NCQA PCMH Recognition	*Subcategory earned*	1
	Volume insufficient	0
	Criteria not met - external data	0

Commercial adult effectiveness for PCPs, Table 31

Effectiveness for PCP, Commercial, adults

Designated ≥ 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories		
Chronic episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Non-chronic episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Risk-adjusted utilization	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

Commercial pediatric clinical quality for PCPs, Table 32

Clinical quality for PCP, Commercial, pediatrics

Designated ≥ 1 > Not designated

Summary of subcategories

Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories

HEDIS Measures		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
NCQA PCMH Recognition		
	Subcategory earned	1
	Volume insufficient	0
	Criteria not met - external data	0

Commercial pediatric effectiveness for PCPs, Table 33

Effectiveness for PCP, Commercial, pediatrics

Designated ≥ 1 > Not designated

Summary of subcategories

Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories

Episodes of care		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Risk-adjusted utilization measures		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

11. Primary care physician (PCP), Medicare program

The Medicare program for PCPs evaluates episodes of care and risk-adjusted utilization measures in its effectiveness category. In the clinical quality category, we evaluate HEDIS measures and NCQA Recognition. We measure family practice and internal medicine practices, including physician assistants and nurse practitioners managed by these physicians. You can find more detailed information above on the various measurements (see sections 3, 3.1, 3.3, 4, 4.2, 5, and 5.1).

Thresholds

To be included in the program, physician practices must see enough members for us to perform a statistically valid evaluation. To measure episodes of care, we require no less than 20 valid inlier episodes for each population in 2020-21 claims data. To measure risk-adjusted utilization, we require 25 attributed, valid members in 2021 claims data. To measure clinical quality, we must have 5 valid members in the denominator of each measure to evaluate practices on those measures, using 2021 claims data as well. To be evaluated for a clinical quality designation, we must be able to evaluate practices on 2 HEDIS measures.

Here are the rules for attribution and member inclusion:

- Valid members are members not attributed to any other practices during calendar year 2021. They also have at least six months of medical benefit eligibility in calendar year 2021. For risk-adjusted utilization and quality measures, we include only members attributed to a primary care practice at the end of 2021.
- Attribution methodology uses our standard attribution logic, which considers both volume and recency of claims. We do measure members receiving care from nurse practitioners and physician assistants managed by primary care practices (see sections 3.1 and 3.2).

Involved episodes

In addition to standard episodes, we only add involved episodes for episodes of care measures in effectiveness. Here are the rules we use to identify involved episodes:

- The episode's responsible provider is not a PCP (or nurse practitioner, or physician assistant managed by a PCP), and
- At least one of the episode's claims are serviced by PCP (or nurse practitioner, or physician assistant managed by a PCP), and
- The PCP (or nurse practitioner or physician assistant) enters the episode before 10 percent of the episode's allowed amounts are spent.

High-cost claimants

- High-cost claimants (HCC) are removed from the analysis. If a member has claim costs (total allowed amount) per year greater than the 99th percentile in either 2020 or 2021, we exclude them from analysis (the entirety of the member's claims, not only the claims that meet the threshold). Below is the HCC threshold by claim year.

High-cost claimants in 2020-21, Table 34

Medicare population	Year	HCC threshold (99th percentile)
Adults	2020	$93,235
Adults	2021	$95,091

Additional measures

In clinical quality for the PCP, Medicare program, we use drug safety measures alongside HEDIS measures. Below are the four measures we use from Centers for Medicare and Medicaid Services (CMS) 2022 Part C & D Star Ratings[13]. These measures were published in 2021, using dates of service from 2020. We benchmark and adjust these like we do HEDIS measures (see sections 5 and 5.1).

13 www.cms.gov/files/document/2022-star-ratings-technical-notes-oct-4-2022.pdf

2021 drug safety measures from CMS Part C & D Star Ratings, Table 35

Measure ID	Measure name
D08	Medication adherence for diabetes medications
D09	Medication adherence for hypertension (RAS antagonists)
D10	Medication adherence for cholesterol (statins)
D12	Statin use in persons with diabetes (SUPD)

Designation hierarchies

Below is a summary of the designations we administer in this program.

Medicare adult, clinical quality, Table 36

Clinical quality for PCP, Medicare

Designated ≥ 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories		
HEDIS measures, and CMS drug safety measures	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
NCQA PCMH Recognition	*Subcategory earned*	1
	Volume insufficient	0
	Criteria not met - external data	0

Medicare adult, effectiveness, Table 37

Effectiveness for PCP, Medicare

Designated ≥ 1 > Not designated

Summary of subcategories

Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories

Chronic episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Non-chronic episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Risk-adjusted utilization	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

12. Pulmonary medicine program

In the Commercial Pulmonary medicine program, we measure episodes of care in the effectiveness category. We use HEDIS measures in the clinical quality category, as well as quality measures from the Merit-based Incentive Payment System (MIPS) created by the Centers for Medicare and Medicaid Services (CMS)

Thresholds

To be part of the program, pulmonary medicine practices must see enough adult members for us to perform a statistically valid evaluation. For measures in the episodes of care category, we evaluate practices with at least six inlier episodes of care in 2020-21 claims data. For measures in the clinical quality category, we only use quality measures that have at least two members in the denominator as observed in 2021 claims data. To be measured, practices must have two measures that meet the denominator threshold.

Involved episodes

For effectiveness and clinical quality measures we use involved episodes in addition to standard episodes. Involved episodes are defined this way:

- The episode's responsible provider is not a Pulmonologist, and
- A Pulmonologist enters the episode before 10 percent of the total allowed is spent, and
- This Pulmonologist owns 30 percent or more of the episode's physician office visits.

High-cost claimants

We remove high-cost claimants (HCC) from any category. If a member has claim costs (total allowed amount) per year in the 99th percentile in either 2020 or 2021, we excluded the member from all analysis. Below is the HCC threshold by claim year.

High-cost claimants in 2020-21, Table 38

Commercial population	Year	HCC threshold (99th percentile)
Adults	2020	$81,092
Adults	2021	$85,337

Additional measures

In addition to the HEDIS measures above (section 5), we include several MIPS measures (2020) in the Pulmonary medicine program. You can find specifications for these at **qpp.cms.gov/mips/explore-measures**. We evaluate TINs where every Pulmonologist is evaluated by MIPS (100% threshold). Table 39 contains the measures we use. We aggregate all four measures and weight each provider's star rating by attributed membership. For all TINs with weighted averages 4 stars or greater, we evaluate "subcategory earned." For all others, we evaluate "criteria not met – external data," unless practices do not meet the threshold, which we indicate with "volume insufficient."

Clinical quality measures from the 2020 CMS Merit-based Incentive Payment System (MIPS), Table 39

NQF ID	CMS quality ID	Measure name
0102	052	Chronic obstructive pulmonary disease (COPD): long-acting inhaled bronchodilator therapy
-	111	Pneumococcal vaccination status for older adults
-	279	Sleep apnea: assessment of adherence to positive airway pressure therapy
-	277	Sleep apnea: severity assessment at initial diagnosis

Below is a summary of the designations we administer in this program. You can find more detailed information above on the various measurements (see sections 3, 3.2, 3.3, 4, 4.3, 5, and 5.1).

Designation hierarchies

Commercial clinical quality category for the pulmonary medicine program, Table 40

Clinical quality for Pulmonology

Designated ≥ 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories		
HEDIS measures		
	Subcategory earned	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
CMS MIPS measures		
	Subcategory earned	1
	Volume insufficient	0
	Criteria not met - external data	0

Commercial effectiveness category for the pulmonary medicine program, Table 41

Effectiveness for Pulmonology

Designated = 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategory		
Episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

13. Spine surgery program

In the Commercial Spine surgery program for orthopedists and neurosurgeons, we evaluate episodes of care and Cotiviti Risk-Readiness® in the effectiveness category. We measure reoperations, surgical site infections, and complications in the clinical quality category. In this program, we only measure orthopedists and neurosurgeons who perform spinal surgeries.

Thresholds

We measure physician practices who see enough members for us to perform a statistically valid evaluation of the practice's patterns. This program only considers adult episodes, so we exclude episodes for all members younger than 18. Only practices with six or more inlier episodes in either back or neck treatments are measured in effectiveness, using 2018-19 claims data. To measure Cotiviti Risk-Readiness, we require at least one episode in 2018. To measure clinical quality, physician practices must have five or more members in the denominator to be evaluated, using 2018-19 claims data.

Designation hierarchies

Below is a summary of the designations we administer in this program. You can find more detailed information above on the various measurements (see sections 3, 3.2, 3.3, 4, 4.3, 5, and 5.1).

Commercial clinical quality category for the Spine surgery program, Table 42[14]

Clinical quality for spine surgery

Designated = 1 > Not designated

Summary of subcategories	
Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategory		
Reoperations, complications, and surgical site of infection measures	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1

[14] In 2022 programs, we communicated letters with "insufficient information" for either "volume insufficient" or "performance not statistically significant." Additionally, we used "designation earned" instead of "subcategory earned." For more detail on the 2022 programs, please review the Spine surgery methodology guide on **Aetna.com/smartcompare**.

Commercial effectiveness category for the Spine surgery program, Table 43[15]

Effectiveness for Spine surgery

Designated ≥ 1 > Not designated

Summary of subcategories

Designation earned	**>0**
Volume insufficient	**0**
Performance not statistically significant	**0**
Performance criteria not met	**<0**

Subcategories

Back episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Neck episodes of care	*Subcategory earned*	1
	Volume insufficient	0
	Performance not statistically significant	0
	Criteria not met	-1
Cotiviti-Risk Readiness®	*Subcategory earned*	1
	Volume insufficient	0
	Criteria not met - external data	0

[15]In 2022 programs, we communicated letters with "insufficient information" for either "volume insufficient" or "performance not statistically significant." Additionally, we used "designation earned" instead of "subcategory earned." For more detail on the 2022 programs, please review the Spine surgery methodology guide on **Aetna.com/smartcompare**.

Additional measures

Our spine surgery program includes three measures that address surgical complications, reoperations, and surgical sites of infection, all based on similar measures supported by the American Academy of Orthopedic Surgeons (AAOS). Below are the specifications for the measures (tables 44 and 45). Because of their similarity, the two measures for surgical complications and surgical sites of infection are collapsed into the same table. We calculate this measure according to these specifications, which otherwise conform to our calculations of HEDIS measures, including Bayesian transformation (see sections 5 and 5.1).

Reoperations measure for clinical quality in the Spine surgery program, Table 44

Unplanned reoperations after spine surgery					
Description	The percentage of spine surgeries where an unplanned reoperation occurred within 90 days of the first surgery.				
Denominator	Members 18 years or older who had a spinal surgery during the first 21 months of the measurement period, as identified with these procedure codes: 0095T, 0098T, 0163T, 0164T, 0165T, 0200T, 0201T, 0202T, 0219T, 0220T, 0221T, 0222T, 0274T, 0275T, 0627T, 0628T, 0629T, 0630T, 0656T, 0657T, 20930, 20931, 20936, 20937, 20938, 20939, 22100, 22101, 22102, 22103, 22110, 22112, 22114, 22116, 22206, 22207, 22208, 22210, 22212, 22214, 22216, 22220, 22222, 22224, 22226, 22310, 22315, 22318, 22319, 22325, 22326, 22327, 22328, 22510, 22511, 22512, 22513, 22514, 22515, 22526, 22527, 22532, 22533, 22534, 22548, 22551, 22552, 22554, 22556, 22558, 22585, 22586, 22590, 22595, 22600, 22610, 22612, 22614, 22630, 22632, 22633, 22634, 22800, 22802, 22804, 22808, 22810, 22812, 22818, 22819, 22830, 22840, 22841, 22842, 22843, 22844, 22845, 22846, 22847, 22848, 22849, 22850, 22852, 22853, 22854, 22855, 22856, 22857, 22858, 22859, 22861, 22862, 22864, 22865, 22867, 22868, 22869, 22870, 22899, 27080, 27280, 61343, 61885, 62267, 62268, 62269, 62287, 62292, 63001, 63003, 63005, 63011, 63012, 63015, 63016, 63017, 63020, 63030, 63035, 63040, 63042, 63043, 63044, 63045, 63046, 63047, 63048, 63050, 63051, 63055, 63056, 63057, 63064, 63066, 63075, 63076, 63077, 63078, 63081, 63082, 63085, 63086, 63087, 63088, 63090, 63091, 63101, 63102, 63103, 63170, 63172, 63173, 63185, 63190, 63194, 63195, 63196, 63197, 63198, 63199, 63200, 63250, 63251, 63252, 63265, 63266, 63267, 63268, 63270, 63271, 63272, 63273, 63275, 63276, 63277, 63278, 63280, 63281, 63282, 63283, 63285, 63286, 63287, 63290, 63295, 63300, 63301, 63302, 63303, 63304, 63305, 63306, 63307, 63308, 63650, 63655, 63661, 63662, 63663, 63664, 63685, 63700, 63702, 63704, 63706, 63707, 63709, 63710, 63740, 63741, 63744, and 63746				
Numerator	Members from the denominator that had a second spinal surgery within 90 days and where that reoperation has either modifier 78 or ICD code Y83.9. This measure is reverse scored in external reporting packages for providers who request additional information about this measure.				
Product Lines	Commercial	**Benefit**	Full medical	**Risk Adjustment**	None, but see 5.1 for transformation
Continuous Enrollment	To be included in this measure, members must have continuous enrollment between January 1, 2018 and December 31, 2019, without gaps in enrollment.				
Population	Adults	**Measure type**	Spine surgery	**Attribution**	Orthopedist or Nurosurgeon
Data	Aetna claims data from measurement years 2018-19				

Surgical site infection and complications measures for clinical quality in the Spine surgery program, Table 45

Surgical site infection and complications after spine surgery (two measures)					
Description	The percentage of spine surgeries where a surgical site infection or complication occurred within 30 days of the spine surgery				
Denominator for either numerator	Members 18 years of age or older who had a spine surgery between the 2nd and 23rd month of the measurement period. We create two denominators for each of the two sub-measures: surgical site infection and complication. The date of the spine surgery will be considered the anchor major procedure. We identify spine surgery with the following CPT codes: 0095T, 0098T, 0163T, 0164T, 0165T, 0200T, 0201T, 0202T, 0219T, 0220T, 0221T, 0222T, 0274T, 0275T, 0627T, 0628T, 0629T, 0630T, 0656T, 0657T, 20930, 20931, 20936, 20937, 20938, 20939, 22100, 22101, 22102, 22103, 22110, 22112, 22114, 22116, 22206, 22207, 22208, 22210, 22212, 22214, 22216, 22220, 22222, 22224, 22226, 22310, 22315, 22318, 22319, 22325, 22326, 22327, 22328, 22510, 22511, 22512, 22513, 22514, 22515, 22526, 22527, 22532, 22533, 22534, 22548, 22551, 22552, 22554, 22556, 22558, 22585, 22586, 22590, 22595, 22600, 22610, 22612, 22614, 22630, 22632, 22633, 22634, 22800, 22802, 22804, 22808, 22810, 22812, 22818, 22819, 22830, 22840, 22841, 22842, 22843, 22844, 22845, 22846, 22847, 22848, 22849, 22850, 22852, 22853, 22854, 22855, 22856, 22857, 22858, 22859, 22861, 22862, 22864, 22865, 22867, 22868, 22869, 22870, 22899, 27080, 27280, 61343, 61885, 62267, 62268, 62269, 62287, 62292, 63001, 63003, 63005, 63011, 63012, 63015, 63016, 63017, 63020, 63030, 63035, 63040, 63042, 63043, 63044, 63045, 63046, 63047, 63048, 63050, 63051, 63055, 63056, 63057, 63064, 63066, 63075, 63076, 63077, 63078, 63081, 63082, 63085, 63086, 63087, 63088, 63090, 63091, 63101, 63102, 63103, 63170, 63172, 63173, 63185, 63190, 63194, 63195, 63196, 63197, 63198, 63199, 63200, 63250, 63251, 63252, 63265, 63266, 63267, 63268, 63270, 63271, 63272, 63273, 63275, 63276, 63277, 63278, 63280, 63281, 63282, 63283, 63285, 63286, 63287, 63290, 63295, 63300, 63301, 63302, 63303, 63304, 63305, 63306, 63307, 63308, 63650, 63655, 63661, 63662, 63663, 63664, 63685, 63700, 63702, 63704, 63706, 63707, 63709, 63710, 63740, 63741, 63744, and 63746				
Denominator exclusion	Exclude anchor procedure where another major procedure occurred 30 days prior to the anchor date, or 30 days after anchor date, except where they are in the same procedure group as the anchor procedure, or where they occurred on the same day as the anchor procedure.				
Numerator for sugical site infection	Members from the denominator who had a surgical site infection within 30 days of the anchor surgery date				
Numerator for complication	Members from the denominator who had a complication within 30 days of the anchor surgery date				
Diagnosis codes used to identify numerator	J95.00, J95.03, J95.04, J95.09, O75.4 , H95.51, H95.52, H95.53, H95.54, T81.83XA, D78.11, D78.12, D78.31, D78.32, D78.33, D78.34, D78.81, D78.89, E36.11, E36.12, E36.8 , E89.820, E89.821, E89.822, E89.823, E89.89, G97.41, G97.82, H59.211 , H59.212 , H59.213 , H59.219 , H59.221 , H59.222 , H59.223 , H59.229 , H59.331, H59.332, H59.333, H59.339, H59.341, H59.342, H59.343, H59.349, H59.351, H59.352, H59.353, H59.359, H59.361, H59.362, H59.363, H59.369, H59.88, H59.89, H95.31, H95.32, H95.88, H95.89, J95.5 , J95.71, J95.72, K91.71, K91.72, L76.11, L76.12, L76.31, L76.32, L76.33, L76.34, L76.81, L76.82, M96.820 , M96.821 , M96.840, M96.841, M96.842, M96.843, M96.89, M96.89, T81.500A, T81.504A, T81.508A, T81.510A, T81.518A, T81.519A, T81.520A, T81.528A, T81.529A, T81.530A, T81.534A, T81.535A, T81.536A, T81.538A, T81.539A, T81.590A, T81.594A, T81.595A, T81.596A, T81.598A, T81.599A, T81.60XA, T81.61XA, T81.69XA, T81.710A, T81.711A, T81.718A, T81.719A, T81.72XA, T81.89XA, T81.9XXA, T86.8481, T86.8482, T86.8483, T86.8489, T86.8491, T86.8492, T86.8493, T86.8499, T81.33XA, T81.40XA, T81.41XA, T81.42XA, T81.43XA, A41.01, A41.02, M00.08, M00.18, M00.28, M00.88, M01.X8, M46.20, M46.30 , T81.4XXA, T84.50XA, T84.60XA, T84.7XXA , T85.79XA, T85.79XA, M46.32, M46.33, M46.34, M46.35, M46.36, M46.22, M46.23, M46.24, M46.25, and M46.26				
Product Lines	Commercial	**Benefit**	Full medical	**Risk Adjustment**	None, but see 5.1 for transformation
Continuous Enrollment	To be included in this measure, members must have continuous enrollment between January 1, 2018 and December 31, 2019, without gaps in enrollment.				
Population	Adults	**Measure type**	Spine surgery	**Attribution**	Orthopedist or Nurosurgeon
Data	Aetna claims data from the measurement years 2018-19				