

(ii) Can demonstrate that the Part D sponsors that it has accredited meet or exceed applicable Medicare requirements; and

(iii) Resubmits the application in its entirety.

(2) An accreditation organization that has requested reconsideration of CMS' denial of its request for approval may not submit a new request until the reconsideration is administratively final.

§ 423.180 Basis and scope of the Part D Prescription Drug Plan Quality Rating System.

(a) *Basis.* This subpart is based on sections 1851(d), 1852(e), 1853(o) and 1854(b)(3)(iii), (v), and (vi) of the Act and the general authority under section 1856(b) of the Act requiring the establishment of standards consistent with and to carry out Part D.

(b) *Purpose.* Ratings calculated and assigned under this subpart will be used by CMS for the following purposes:

(1) To provide comparative information on plan quality and performance to beneficiaries for their use in making knowledgeable enrollment and coverage decisions in the Medicare program.

(2) To provide quality ratings on a 5-star rating system.

(3) To provide a means to evaluate and oversee overall and specific compliance with certain regulatory and contract requirements by Part D plans, where appropriate and possible to use data of the type described in § 423.182(c).

(c) *Applicability.* Except for § 423.182(b)(3), the regulations in this subpart will be applicable beginning with the 2019 measurement period and the associated 2021 Star Ratings that are released prior to the annual coordinated election period for the 2021 contract year.

[83 FR 16743, Apr. 16, 2018]

§ 423.182 Part D Prescription Drug Plan Quality Rating System.

(a) *Definitions.* In this subpart the following terms have the meanings:

Absolute percentage cap is a cap applied to non-CAHPS measures that are on a 0 to 100 scale that restricts move-

ment of the current year's measure-threshold-specific cut point to no more than the stated percentage as compared to the prior year's cut point.

CAHPS refers to a comprehensive and evolving family of surveys that ask consumers and patients to evaluate the interpersonal aspects of health care. CAHPS surveys probe those aspects of care for which consumers and patients are the best or only source of information, as well as those that consumers and patients have identified as being important. CAHPS initially stood for the Consumer Assessment of Health Plans Study, but as the products have evolved beyond health plans the acronym now stands for Consumer Assessment of Healthcare Providers and Systems.

Case-mix adjustment means an adjustment to the measure score made prior to the score being converted into a Star Rating to take into account certain enrollee characteristics that are not under the control of the plan. For example age, education, chronic medical conditions, and functional health status that may be related to the enrollee's survey responses.

Categorical Adjustment Index (CAI) means the factor that is added to or subtracted from an overall or summary Star Rating (or both) to adjust for the average within-contract (or within-plan as applicable) disparity in performance associated with the percentages of beneficiaries who are dually eligible for Medicare and enrolled in Medicaid, beneficiaries who receive a Low Income Subsidy, or have disability status in that contract (or plan as applicable).

Clustering refers to a variety of techniques used to partition data into distinct groups such that the observations within a group are as similar as possible to each other, and as dissimilar as possible to observations in any other group. Clustering of the measure-specific scores means that gaps that exist within the distribution of the scores are identified to create groups (clusters) that are then used to identify the four cut points resulting in the creation of five levels (one for each Star Rating), such that the scores in the same Star Rating level are as similar as possible and the scores in different

Star Rating levels are as different as possible. Technically, the variance in measure scores is separated into within-cluster and between-cluster sum of squares components. The clusters reflect the groupings of numeric value scores that minimize the variance of scores within the clusters. The Star Ratings levels are assigned to the clusters that minimize the within-cluster sum of squares. The cut points for star assignments are derived from the range of measure scores per cluster, and the star levels associated with each cluster are determined by ordering the means of the clusters.

Consolidation means when an MA organization that has at least two contracts for health and/or drug services of the same plan type under the same parent organization in a year combines multiple contracts into a single contract for the start of the subsequent contract year.

Consumed contract means a contract that will no longer exist after a contract year's end as a result of a consolidation.

Cut point cap is a restriction on the change in the amount of movement a measure-threshold-specific cut point can make as compared to the prior year's measure-threshold-specific cut point. A cut point cap can restrict upward movement, downward movement, or both.

Display page means the CMS website on which certain measures and scores are publicly available for informational purposes; the measures that are presented on the display page are not used in assigning Part C and D Star Ratings.

Domain rating means the rating that groups measures together by dimensions of care.

Dual-eligible (DE) means a beneficiary who is enrolled in both Medicare and Medicaid.

Guardrail is a bidirectional cap that restricts both upward and downward movement of a measure-threshold-specific cut point for the current year's measure-level Star Ratings as compared to the prior year's measure-threshold-specific cut point.

Health equity index means an index that summarizes contract performance among those with specified social risk

factors (SRFs) across multiple measures into a single score.

Highest rating means the overall rating for MA-PDs, the Part C summary rating for MA-only contracts, and the Part D summary rating for PDPs.

Highly-rated contract means a contract that has 4 or more stars for its highest rating when calculated without the improvement measures and with all applicable adjustments in § 423.186(f).

Low-income subsidy (LIS) means the subsidy that a beneficiary receives to help pay for prescription drug coverage (see § 423.34 for definition of a low-income subsidy eligible individual).

Mean resampling refers to a technique where measure-specific scores for the current year's Star Ratings are randomly separated into 10 equal-sized groups. The hierarchical clustering algorithm is done 10 times, each time leaving one of the 10 groups out. By leaving out one of the 10 groups for each run, 9 of the 10 groups, which is 90 percent of the applicable measure scores, are used for each run of the clustering algorithm. The method results in 10 sets of measure-specific cut points. The mean cut point for each threshold per measure is calculated using the 10 values.

Measurement period means the period for which data are collected for a measure or the performance period that a measure covers.

Measure score means the numeric value of the measure or an assigned 'missing data' message.

Measure star means the measure's numeric value is converted to a Star Rating. It is displayed to the nearest whole star, using a 1–5 star scale.

Overall rating means a global rating that summarizes the quality and performance for the types of services offered across all unique Part C and Part D measures.

Part C summary rating means a global rating that summarizes the health plan quality and performance on Part C measures.

Part D summary rating means a global rating that summarizes prescription drug plan quality and performance on Part D measures.

Plan benefit package (PBP) means a set of benefits for a defined MA or PDP service area. The PBP is submitted by

Part D plan sponsors and MA organizations to CMS for benefit analysis, bidding, marketing, and beneficiary communication purposes.

Reliability means a measure of the fraction of the variation among the observed measure values that is due to real differences in quality (“signal”) rather than random variation (“noise”); it is reflected on a scale from 0 (all differences in plan performance measure scores are due to measurement error) to 1 (the difference in plan performance scores is attributable to real differences in performance).

Restricted range is the difference between the maximum and minimum measure score values using the prior year measure scores excluding outer fence outliers (first quartile $-3 \times$ Interquartile Range (IQR) and third quartile $+3 \times$ IQR).

Restricted range cap is a cap applied to non-CAHPS measures that restricts movement of the current year’s measure-threshold-specific cut point to no more than the stated percentage of the restricted range of a measure calculated using the prior year’s measure score distribution.

Reward factor means a rating-specific factor added to the contract’s summary or overall ratings (or both) if a contract has both high and stable relative performance.

Statistical significance assesses how likely differences observed in performance are due to random chance alone under the assumption that plans are actually performing the same.

Surviving contract means the contract that will still exist under a consolidation, and all of the beneficiaries enrolled in the consumed contract(s) are moved to the surviving contracts.

Traditional rounding rules mean that the last digit in a value will be rounded. If rounding to a whole number, look at the digit in the first decimal place. If the digit in the first decimal place is 0, 1, 2, 3 or 4, then the value should be rounded down by deleting the digit in the first decimal place. If the digit in the first decimal place is 5 or greater, then the value should be rounded up by 1 and the digit in the first decimal place deleted.

Tukey outer fence outliers are measure scores that are below a certain point

(first quartile $-3.0 \times$ (third quartile $-$ first quartile)) or above a certain point (third quartile $+3.0 \times$ (third quartile $-$ first quartile)).

(b) *Contract ratings*—(1) *General*. CMS calculates an overall Star Rating, Part C summary rating, and Part D summary rating for each MA–PD contract and a Part D summary rating for each PDP contract using the 5-star rating system described in this subpart. For PDP contracts, the Part D summary rating is the highest rating. Measures are assigned stars at the contract level and weighted in accordance with § 423.186(a). Domain ratings are the unweighted mean of the individual measure ratings under the topic area in accordance with § 423.186(b). Summary ratings are the weighted mean of the individual measure ratings for Part C or Part D in accordance with § 423.186(c), with the applicable adjustments provided in paragraph (f) of this section. Overall Star Ratings are calculated by using the weighted mean of the individual measure ratings in accordance with § 423.186(d), with the applicable adjustments provided in paragraph (f) of this section. CMS includes the Star Ratings measures in the overall and summary ratings that are associated with the contract type for the Star Ratings year.

(2) *Plan benefit packages*. All plan benefit packages (PBPs) offered under an MA contract or PDP plan sponsor have the same overall and/or summary Star Ratings as the contract under which the PBP is offered by the MA organization or PDP plan sponsor. Data from all the PBPs offered under a contract are used to calculate the measure and domain ratings for the contract.

(3) *Contract consolidations*. (i) In the case of contract consolidations involving two or more contracts for health and/or drug services of the same plan type under the same parent organization, CMS assigns Star Ratings for the first and second years following the consolidation based on the enrollment-weighted mean of the measure scores of the surviving and consumed contract(s) as provided in paragraph (b)(3)(ii) of this section.

(ii) The Star Ratings posted on Medicare Plan Finder for contracts that consolidate are as follows:

(A)(1) For the first year after consolidation, CMS uses enrollment-weighted measure scores using the July enrollment of the measurement period of the consumed and surviving contracts for all measures, except survey-based measures, call center measures, and improvement measures. The survey-based measures will use enrollment of the surviving and consumed contracts at the time the sample is pulled for the rating year. The call center measures will use average enrollment during the study period. The Part C and D improvement measures are not calculated for first year consolidations.

(2) For contract consolidations approved on or after January 1, 2022, if a measure score for a consumed or surviving contract is missing due to a data integrity issue as described in § 423.184(g)(1)(i) and (ii), CMS assigns a score of zero for the missing measure score in the calculation of the enrollment-weighted measure score.

(B)(1) For the second year after consolidation, CMS uses the enrollment-weighted measure scores using the July enrollment of the measurement year of the consumed and surviving contracts for all measures except for CAHPS. CMS ensures that the CAHPS survey sample includes enrollees in the sample frame from both the surviving and consumed contracts.

(2) For contract consolidations approved on or after January 1, 2022, for all measures except CAHPS if a measure score for a consumed or surviving contract is missing due to a data integrity issue as described in § 423.184(g)(1)(i) and (ii), CMS assigns a score of zero for the missing measure score in the calculation of the enrollment-weighted measure score.

(iii) This provision governing the Star Ratings of surviving contracts is applicable to contract consolidations that are approved on or after January 1, 2019.

(c) *Data sources.* (1) Part D Star Ratings measures reflect structure, process, and outcome indices of quality. This includes information of the following types: Beneficiary experiences, benefit administration information, clinical data, and CMS administrative data. Data underlying Star Ratings measures may include survey data,

data separately collected and used in oversight of Part D plans' compliance with contract requirements, data submitted by plans, and CMS administrative data.

(2) Part D sponsors are required to collect, analyze, and report data that permit measurements of health outcomes and other indices of quality. Part D sponsors must provide unbiased, accurate, and complete quality data described in paragraph (c)(1) of this section to CMS on a timely basis as requested by CMS.

(3) For 2021 Star Ratings only, Part D sponsors are not required to submit CAHPS data that would otherwise be required for the calculation of the 2021 Star Ratings.

[83 FR 16743, Apr. 16, 2018; 84 FR 15841, Apr. 16, 2019, as amended at 85 FR 19290, Apr. 6, 2020; 85 FR 33911, June 2, 2020; 86 FR 6118, Jan. 19, 2021; 88 FR 22337, Apr. 12, 2023]

§ 423.184 Adding, updating, and removing measures.

(a) *General.* CMS adds, updates, and removes measures used to calculate the Star Ratings as provided in this section. CMS lists the measures used for a particular Star Rating each year in the Technical Notes or similar guidance document with publication of the Star Ratings.

(b) *Review of data quality.* CMS reviews the quality of the data on which performance, scoring and rating of a measure is based before using the data to score and rate performance or in calculating a Star Rating. This includes review of variation in scores among MA organizations and Part D plan sponsors, and the accuracy, reliability, and validity of measures and performance data before making a final determination about inclusion of measures in each year's Star Ratings.

(c) *Adding measures.* (1) CMS will continue to review measures that are nationally endorsed and in alignment with the private sector, such as measures developed by National Committee for Quality Assurance (NCQA) and the Pharmacy Quality Alliance (PQA) or endorsed by the National Quality Forum for adoption and use in the Part D Quality Ratings System. CMS may develop its own measures as well when