42 C.F.R. § 423.186

Calculation of Star Ratings.

(a) *Measure Star Ratings*—(1) *Cut points.* CMS will determine cut points for the assignment of a Star Rating for each numeric measure score by applying either a clustering or a relative distribution and significance testing methodology. For the Part D measures, CMS will determine MA–PD and PDP cut points separately.

(2) Clustering algorithm for all measures except CAHPS measures.

(i) The method maximizes differences across the star categories and minimizes the differences within star categories using mean resampling with the hierarchal clustering of the current year's data. Effective for the Star Ratings issued in October 2023 and subsequent years, prior to applying mean resampling with hierarchal clustering, Tukey outer fence outliers are removed. Effective for the Star Ratings issued in October 2022 and subsequent years, CMS will add a guardrail so that the measure-threshold-specific cut points for non-CAHPS measures do not increase or decrease more than the value of the cap from 1 year to the next. The cap is equal to 5 percentage points for measures having a 0 to 100 scale (absolute percentage cap) or 5 percent of the restricted range for measures not having a 0 to 100 scale (restricted range cap). New measures that have been in the Part C and D Star Rating program for 3 years or less use the hierarchal clustering methodology with mean resampling with no guardrail for the first 3 years in the program.

(ii) In cases where multiple clusters have the same measure score value range, those clusters would be combined, leading to fewer than 5 clusters.

(iii) The clustering algorithm for the improvement measure scores is done in two steps to determine the cut points for the measure-level Star Ratings. Clustering is conducted separately for improvement measure scores greater than or equal to zero and those with improvement measure scores less than zero.

(A) Improvement scores of zero or greater would be assigned at least 3 stars for the improvement Star Rating.

(B) Improvement scores less than zero would be assigned either 1 or 2 stars for the improvement Star Rating.

(3) *Relative distribution and significance testing for CAHPS measures.* The method combines evaluating the relative percentile distribution with significance testing and accounts for the reliability of scores produced from survey data; no measure Star Rating is produced if the reliability of a CAHPS measure is less than 0.60. Low reliability scores are defined as those with at least 11 respondents, reliability greater than or equal to 0.60 but less than 0.75, and also in the lowest 12 percent of contracts ordered by reliability. The following rules apply:

(i) A contract is assigned 1 star if both of the criteria in paragraphs (a)(3)(i)(A) and (B) of this section are met plus at least one of the criteria in paragraphs (a)(3)(i)(C) or (D) of this section is met:

(A) Its average CAHPS measure score is lower than the 15th percentile; and

(B) Its average CAHPS measure score is statistically significantly lower than the national average CAHPS measure

score;

(C) The reliability is not low; or

(D) Its average CAHPS measure score is more than one standard error below the 15th percentile.

(ii) A contract is assigned 2 stars if it does not meet the 1-star criteria and meets at least one of these three criteria:

(A) Its average CAHPS measure score is lower than the 30th percentile and the measure does not have low reliability; or

(B) Its average CAHPS measure score is lower than the 15th percentile and the measure has low reliability; or

(C) Its average CAHPS measure score is statistically significantly lower than the national average CAHPS measure score and below the 60th percentile.

(iii) A contract is assigned 3 stars if it meets at least one of these three criteria:

(A) Its average CAHPS measure score is at or above the 30th percentile and lower than the 60th percentile, and it is not statistically significantly different from the national average CAHPS measure score; or

(B) Its average CAHPS measure score is at or above the 15th percentile and lower than the 30th percentile, the reliability is low, and the score is not statistically significantly lower than the national average CAHPS measure score; or

(C) Its average CAHPS measure score is at or above the 60th percentile and lower than the 80th percentile, the reliability is low, and the score is not statistically significantly higher than the national average CAHPS measure score.

(iv) A contract is assigned 4 stars if it does not meet the 5-star criteria and meets at least one of these three criteria:

(A) Its average CAHPS measure score is at or above the 60th percentile and the measure does not have low reliability; or

(B) Its average CAHPS measure score is at or above the 80th percentile and the measure has low reliability; or

(C) Its average CAHPS measure score is statistically significantly higher than the national average CAHPS measure score and above the 30th percentile.

(v) A contract is assigned 5 stars if both of the following criteria in paragraphs (a)(3)(v)(A) and (B) of this section are met plus at least one of the criteria in paragraphs (a)(3)(v)(C) or (D) of this section is met:

(A) Its average CAHPS measure score is at or above the 80th percentile; and

(B) Its average CAHPS measure score is statistically significantly higher than the national average CAHPS measure score;

(C) The reliability is not low; or

(D) Its average CAHPS measure score is more than one standard error above the 80th percentile.

(4) *5-Star Scale.* Measure scores are converted to a 5-star scale ranging from 1 (worst rating) to 5 (best rating),

with whole star increments for the cut points.

(b) *Domain Star Ratings.* (1)(i) CMS groups measures by domains solely for purposes of public reporting the data on Medicare Plan Finder. They are not used in the calculation of the summary or overall ratings. Domains are used to group measures by dimensions of care that together represent a unique and important aspect of quality and performance.

(ii) The 4 domains for the Part D Star Ratings are: Drug Plan Customer Service; Member Complaints and Changes in the Drug Plan's Performance; Member Experience with the Drug Plan; and Drug Safety and Accuracy of Drug Pricing.

(2) CMS calculates the domain ratings as the unweighted mean of the Star Ratings of the included measures.

(i) A contract must have scores for at least 50 percent of the measures required to be reported for that contract type for that domain to have a domain rating calculated.

(ii) The domain ratings are on a 1 to 5 star scale ranging from 1 (worst rating) to 5 (best rating) in whole star increments using traditional rounding rules.

(c) *Part D summary ratings.* (1) CMS will calculate the Part D summary ratings using the weighted mean of the measure-level Star Ratings for Part D, weighted in accordance with paragraph (e) of this section and with the applicable adjustments provided in paragraph (f) of this section.

(2)

(i) A contract must have scores for at least 50 percent of the measures required to be reported for the contract type to have a summary rating calculated.

(ii) The Part D improvement measure is not included in the count of the minimum number of rated measures.

(3) The summary ratings are on a 1 to 5 star scale ranging from 1 (worst rating) to 5 (best rating) in half-star increments using traditional rounding rules.

(d) *Overall MA-PD rating.* (1) The overall rating for a MA-PD contract will be calculated using a weighted mean of the Part C and Part D measure-level Star Ratings, weighted in accordance with paragraph (e) of this section and with the applicable adjustments provided in paragraph (f) of this section.

(2)

(i) An MA-PD must have both Part C and Part D summary ratings and scores for at least 50 percent of the measures required to be reported for the contract type to have the overall rating calculated.

(ii) The Part C and D improvement measures are not included in the count of measures needed for the overall rating.

(iii) Any measures that share the same data and are included in both the Part C and Part D summary ratings will be included only once in the calculation for the overall rating.

(iv) The overall rating is on a 1 to 5 star scale ranging from 1 (worst rating) to 5 (best rating) in half-increments using traditional rounding rules.

- 4 -