IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF TEXAS
FORT WORTH DIVISION

ELEVANCE HEALTH, INC.; COMMUNITY INSURANCE COMPANY; FREEDOM HEALTH, INC.; GROUP RETIREE HEALTH SOLUTIONS, INC.; WELLPOINT INSURANCE COMPANY; and WELLPOINT TEXAS, INC., Plaintiffs, v. ROBERT F. KENNEDY, JR., in his official capacity as Secretary of Health and Human Services; and DR. MEHMET OZ, in his official capacity as Administrator of the Centers for Medicare & Medicaid Services, Defendants.[1]	Civil Action No. 4:24-cv-01064-P

DEFENDANTS' REPLY BRIEF
IN SUPPORT OF THEIR CROSS MOTION FOR SUMMARY JUDGMENT

CHAD E. MEACHAM
ACTING UNITED STATES ATTORNEY

/s/ Andrea Hyatt
Andrea Hyatt
Assistant United States Attorney
Texas Bar No. 24007419
Burnett Plaza, Suite 1700
801 Cherry Street, Unit #4
Fort Worth, Texas 76102-6882
Telephone: 817.252.5230
Fax: 817.252.5458
Andrea.Hyatt@usdoj.gov

Attorneys for Defendants

[1] Xavier Becerra has been substituted with Robert F. Kennedy, Jr. as Secretary of the United States Department of Health and Human Services, and Chiquita Brooks-LaSure has been substituted with Dr. Mehmet Oz as Administrator of the Centers for Medicare & Medicaid Services, pursuant to Federal Rule of Civil Procedure 25(d).

Table of Contents

Table of Authorities

Cases

Statutes

Regulations

Regulatory

Rules

Other Authorities

I. Summary

Elevance's Response and Reply underscores that its arguments are predicated on a series of basic errors and misunderstandings about how the Centers for Medicare & Medicaid Services (CMS) calculates Star Ratings:

- First, to support its contention that the Secretary's regulations clearly authorize CMS to case-mix adjust Consumer Assessment of Healthcare Providers and Systems measures) (CAHPS measures), Elevance cites the part of the Secretary's regulations that prescribe rules with regard to calculation of measure Star Ratings—not numerical CAHPS measure scores—and contends that because this section does not mention case-mix adjustment, it must not be authorized. But measure Star Ratings are not case-mix adjusted, numerical measure scores are.

- Next, Elevance contends that CMS should have calculated a national average CAHPS contract measure score. But the regulations require that CMS calculate a "national average CAHPS measure score," not a national average *contract* measure score.

- Elevance further contends that CMS calculates and rounds overall scores to the sixth decimal place. CMS does not; it rounds to the first decimal, or tenths place, looking to the second decimal, or hundredths place and calculates with at least six digits of precision.

- Finally, Elevance contends that using the same "seed"—the initiating number sequence used to create the random groups for purposes of mean resampling—year over year makes mean resampling less random because CMS has failed to account for the remote possibility that the group of contracts does not change year to year. But the group of contracts always changes year to year.

Elevance's arguments are statistically dubious and internally contradictory. Elevance maintains its request to require CMS to double round. Elevance argues that case-mix adjustment is not supported by the Secretary's regulations. At the same time, it contends that CMS has inadequately addressed nonresponse bias, despite ample evidence that case-mix adjustments address the problem of nonresponse bias. If the Court grants Elevance the relief it seeks here, it opens its doors to these kinds of dubious, internally contradictory challenges motivated only by Medicare Advantage organizations' desire to nudge their scores up into the next half-star category. The Court should reject these efforts.

II. Argument and Authorities

A. CMS Case-Mix Adjusted CAHPS Measure Follows from the Applicable Regulations.

1. The applicable regulations authorize CMS to case-mix adjust CAHPS measures.

The Secretary's regulations repeatedly authorize CMS to case-mix adjust CAHPS measures:

- The regulations state that case-mix adjustments "to the measure score [are] 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 a plan." 42 C.F.R. § 422.162(a).
- The regulations confirm that some CAHPS measure scores are "*already* case-mix adjusted." 42 C.F.R. § 422.166(f)(2)[2] (emphasis added).
- Additionally, the regulations refer to "[m]easures that *are* case-mix adjusted in the Star Ratings." § 422.166(f)(3) (emphasis added).

Despite these explicit regulatory authorizations, Elevance maintains that the Secretary's regulations do not authorize CMS to case-mix adjust CAHPS-based measure scores. Pls.' Consol. Br. in Support of Resp. and Reply ("Pls.' Resp. Br.") 5-14, ECF No. 42. Elevance's proposed reading of these regulations is implausible because these regulations clearly authorize case-mix adjusting.

2. Elevance's arguments against case-mix adjusting are flawed.

a. Elevance's misunderstands section 422.166(a)(3) in two ways.

Elevance misapprehends 42 C.F.R. § 422.166(a)(3), which sets out the methodology for calculating CAHPS measure Star Ratings. Elevance contends that because section 422.166(a)(3) does not mention case-mix adjusting, CMS is not empowered to case-mix adjust CAHPS measures. Pls.' Resp. Br. 5-7. Elevance makes two errors. First, Elevance states that "42 C.F.R.

[2] For ease of reference, this brief omits reference to the parallel citations in 42 C.F.R. § 423.182 *et seq*.

§ 422.166(a)(3) dictates how CAHPS-based measure scores are calculated for any contract." *Id*. at 5. This is wrong. Section 422.166(a), entitled "Measure Star Ratings," is not about "how CAHPS-based measure scores are calculated"—it concerns how measure Star Ratings are calculated.

Second, Elevance states that the "regulation expressly provides that, to calculate measure scores for CAHPS-based individual measures, Defendants use the relative distribution and significance testing methodology . . ." *Id.* This is also wrong. CMS uses relative distribution and significance testing to calculate *measure Star Ratings*. 42 C.F.R. § 422.166(a)(3). Elevance is conflating numerical measure scores with measure Star Ratings. As Elevance correctly explains, section 422.166(a)(3) dictates the methodology for determining whether a measure receives 1, 2, 3, 4, or 5 stars. Pls.' Resp. Br. 5. The case-mix adjustment is part of the calculation of the numeric measure scores for CAHPS measures. *See* 42 C.F.R. § 422.162 ("Case-mix adjustment means an adjustment to the measure score *made prior to the score being converted into a Star Rating* . . ." (emphasis added)). Section 422.166(a)(3) explains how to convert CAHPS measure scores into measure Star Ratings. That regulatory provision does not explain how measure scores should be calculated. It would be improper (and make no sense) for the Secretary to authorize case-mix adjusting in the portion of its regulations pertaining to calculation of measure Star Ratings.

b. Elevance's reliance on the Categorial Adjustment Index and Health Equity Index is misplaced.

Elevance points to the Categorical Adjustment Index (CAI) and Health Equity Index (HEI), 42 C.F.R. §§ 422.166(f)(2), (f)(3), to assert that "Defendants know how to apply [case-mix adjustment] in the regulations." Pls.' Resp. Br. 7. But the CAI and HEI provisions support the opposite reading of the regulations Elevance attempts to advance. As CMS has already explained, Defs.' Summ. J. Brief, ECF 38 (Defs.' Br.), 19-20, the upshot of the CAI and HEI provisions is that the Secretary accounted for the fact that some measures, and specifically CAHPS measures,

will have already been case-mix adjusted. *See* Medicare Program; Contract Year 2019 Policy, 83 Fed. Reg. 16440, 16,581 (Apr. 16, 2018) ("Measures would be excluded as candidates for adjustment if the measures are already case-mix adjusted for [socioeconomic status] (for example, CAHPS and [Health Outcomes Survey] outcome measures.")). Elevance simply does not engage with the fact that the case-mix adjusted measures referred to in these provisions are the same CAHPS measures they contend the regulations prohibit CMS from case-mix adjusting. If case-mix adjusting CAHPS measures violates CMS's regulations as Elevance contends, the CAI regulations would not have to exclude them from the CAI calculation on the basis that they are already case-mix adjusted. Because Elevance's reading of the Secretary's regulations would render portions of § 422.166(f) "inoperative or superfluous, void or insignificant," the Court should reject it. *See Gulf Fishermens Ass'n v. Nat'l Marine Fisheries Serv.*, 968 F.3d 454, 464-65 (5th Cir. 2020) (describing anti-surplusage canon).

c. Elevance's theory asks the Court to overlook regulatory text.

Elevance contends that "the mere existence of a regulatory definition for a term does not confer authority on the agency to implement the defined concept." Pls.' Resp. Br. 10. That an authorization for CMS to take a discrete action appears in a section of a regulation labeled "Definitions" is no matter—this Court's analysis should begin and end with the text. *Kisor v. Wilkie*, 588 U.S. 558, 575 (2019) ("a court must carefully consider the text") (cleaned up). The regulation states: "a 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." § 422.162(a). The regulation unambiguously permits case-mix adjustments to be "made." "Made"—the past participle of "make"—means "to perform an action." *See Make*, Cambridge University Press & Assessment, https://dictionary.cambridge.org/ (last visited April 21, 2025). Here, that action was case-mix adjustment. The Secretary's use of the past participle of "make"

serves to codify the existing practice of case-mix adjusting CAHPS measures, which began with the initiation of CAHPS in 1998. Defs.' Br. 22.

In this litigation, Elevance has already taken the opposite position that it takes in its response and reply—it has argued that provisions of subsection 422.162(a) not only authorize the Agency to act in certain ways, but expressly require it to do so. Pls.' Mem. of Law in Supp. Pls.' Mot. for Summ. J. ("Pls.' Br.") 12, 28, 34, ECF No. 32. Elevance cites to and quotes the regulation's definition of "mean resampling" : "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." Pls.' Br. 28, 34 (citing at § 422.162(a)). Based on that definition alone, Elevance contends that "CMS regulations *expressly require* mean resampling to achieve a random separation of the applicable contracts into groups." *Id.* (emphasis added); *see also* Pls.' Resp. Br. 43 (asserting randomness is "required by the regulation"). Citing only to the definition subsection, § 422.162(a), Elevance describes its assertion that mean sampling must be random as a "regulatory requirement." *Id.* at 12-13, 35. But even though "case-mix adjustment" is defined in a similar manner as "mean resampling," in the same subsection of the regulation and conspicuously, Elevance contends that "the applicable regulation does not call for [the case-mix] adjustment," Pls.' Br. 18. In contrast, according to Elevance, not only does the "mean resampling" definition in the same regulatory section authorize mean resampling—it imposes the specific "regulatory requirement that means [sic] resampling must be random." *Id.* 12-13, 34, 35. And unlike "case-mix adjustment," the randomness "regulatory requirement" appears nowhere else in the Secretary's regulations pertaining to Star Ratings, only in the definition subsection, § 422.162(a). Elevance concedes that the definitions in subsection 422.162(a) authorize CMS to take the described, discrete actions.

d. Elevance's theory fails to recognize that CMS's case-mix adjustment definition is constraining and is supported by preamble text.

CMS could not, as Elevance suggests, "case-mix adjust seemingly any measure score within the Medicare Advantage universe," nor does CMS have "*carte blanche* to determine whether a measure is case-mix adjusted." Pls.' Resp. Br. 9-10. The definition limits when CMS can case-mix adjust—"prior to the score being converted into a Star Rating." § 422.162(a). Case-mix adjusting must only consider "enrollee characteristics that are not under the control of the plan," like "age, education, chronic medical condition." *Id.* Those characteristics may be related to "the enrollee's survey responses." *Id.* Only a handful of measures are premised on survey responses—nine measures from the CAHPS Survey and three measures from the Health Outcomes Survey, only some of which are currently case-mix adjusted. A.R. 38-113. In light of this definition, it would make no sense for the Secretary to impose these constraints on case-mix adjusting survey-based measures and also prohibit case-mix adjusting CAHPS measures.

The Secretary's preamble language "inform[s] the interpretation of a regulation." *Texas v. HHS*, No. 6:24-cv-348, 2025 WL 818155, at *9 (E.D. Tex. Mar. 13, 2025) (quoting *Peabody Twentymile Mining, LLC v. Sec'y of Lab.*, 931 F.3d 992, 998 (10th Cir. 2019)). Unlike in *Elevance Health, Inc. v. Becerra*, 736 F. Supp. 3d 1, 23-24 (D.D.C. 2024) and *Scan Health Plan v. HHS*, Case No. 23-cv-03910, 2024 WL 2815789, at *6–7 (D.D.C. June 3, 2024), where the courts found that the preamble and regulations text were in conflict, here, the preamble of the 2018 rule accords with the reasonable reading of the regulatory text as authorizing case-mix adjustments. Elevance's position is that because subsection 422.166(a)(3) does not explicitly authorize case-mix adjustments, there is a conflict between the preamble and the regulation text. This supposed regulatory silence does not create the preamble-regulation conflict that was at issue in *Elevance Health, Inc.* and *Scan Health Plan*. In those cases, for the 2024 Star Ratings, the courts concluded

that the regulatory text required application of a guardrail to *actual* cut points and preamble text required application of a guardrail to *hypothetical* cut points, creating a true conflict. *See Scan Health Plan*, 2024 WL 2815789, at *6; *Elevance Health, Inc.*, 736 F. Supp. 3d at 23–24. There is no similar conflict here. Nothing in subsection 422.166(a)(3) forecloses CMS from case-mix adjusting CAHPS measure scores. As already established, subsection 422.166(a)(3) makes no mention of case-mix adjustments because it relates to measure Star Ratings (which are not case-mix adjusted), not measure scores. And as also established, other regulatory provisions authorize case-mix adjustment—the definition of case-mix adjustment at subsection 422.162(a) and the CAI and HEI provisions at subsection 422.166(f). The Secretary explained in preamble text in the 2018 Final Rule that CAHPS measures are case-mix adjusted: "For CAHPS measures, contracts are first classified into base groups by comparisons to percentile cut points defined by the current-year distribution of *case-mix adjusted* contract means." 83 Fed. Reg. at 16,568 (emphasis added). As part of this rulemaking, CMS responded to comments related to case-mix adjustments. *Id.* at 16,527, 16,555. The extensive discussion of CMS's application of case-mix adjustments in preamble text informs the interpretation of CMS's regulations as permitting case-mix adjusting of CAHPS measures.

In each of the other cases that Elevance cites for the proposition that preamble text does not create law, the courts concluded that the regulatory text conflicted with the preamble. *See Tex. Children's Hosp. v. Burwell*, 76 F. Supp. 3d 224, 237 (D.D.C. 2014) ("To the extent that this definition is contradicted by the Rule's Preamble, the definition controls."); *AT&T Corp. v. FCC*, 970 F.3d 344, 351, 449 (D.C. Cir. 2020) (explaining that when "there is a discrepancy between the preamble and the Code, it is the codified provisions that control"); *Air Prods. & Chems., Inc. v.*

GSA, 700 F. Supp. 3d 487, 501 (N.D. Tex. 2023) (explaining that preamble words "cannot override unambiguous text"). Not so here.

e. Elevance's theory ignores the evidence in the record.

Finally, Elevance briefly asserts that there is no evidence in the record of consideration "of any implications on the weighting of any racial, age, educational bias or other demographic scores utilized by CMS to 'rationalize' these survey measure scores." Pls.' Resp. Br. 11. This is false. As an initial matter, contrary to Elevance's claim that there is a lack of basic information underscoring these case-mix adjustments, the case-mix coefficients are publicly available. A.R. 115-21. CMS adequately explained its determination to case-mix adjust survey-based measures. CMS explained that certain populations, for example, beneficiaries who are in the 75–79 age range may tend to respond more positively or negatively to certain survey questions. *Id.* at 114. Without applying a case-mix adjustment, contracts that serve enrollees who are more likely to give positive responses would be advantaged and contracts that serve enrollees who are more likely to give fewer positive responses would be disadvantaged. Elevance fails to explain why that outcome is preferable. It even contends that CMS has not adequately addressed one form of bias that case-mix adjustments sufficiently address—survey nonresponse bias. Pls.' Resp. Br. 34-38. Elevance has not offered any policy objection to the case-mix adjustment. It cites to *Motor Vehicle Mfrs. Ass'n, Inc. v. State Farm*, 463 U.S. 29 (1983), but fails to state what aspect of the problem it thinks CMS failed to consider.

The regulatory provision regarding the definition of case-mix adjustment, read together with the CAI and HEI provisions at subsection 422.166(f), as well as the preamble language in the 2018 Final Rule, provides ample support for case-mix adjusting certain CAHPS measures.[3]

B. CMS's Method for Calculating the "National Average" Is Consistent with the Regulations, and Nothing Compels Elevance's Preferred Interpretation.

Elevance argues that CMS must calculate the national average *contract* score and claims that CMS used a "national weighted average" instead of a "national average," in violation of the plain language of 42 C.F.R. § 422.166(a)(3). Pls.' Resp. Br. 17-20. This claim fails for two reasons.

1. Elevance's Amended Complaint does not assert a claim that CMS incorrectly calculated the national average.

In its Amended Complaint, Elevance claimed only that "CMS may double-penalize contracts by first applying the case-mix adjustment and then further reducing the CAHPS score for a measure when the score is significantly lower than the national mean for that measure" and that this "double-whammy is not contemplated by the applicable statute or regulations." Am. Compl. ¶¶ 34, 40. It is not sufficient for Elevance to simply reference the national average in relation with its case-mix adjustment claim. The central part of Elevance's new claim—that CMS allegedly incorrectly calculated the national average—is completely absent from the Amended Complaint. Granted, Elevance is not required to use "magic words" in its complaint. *Arredondo v. Schlumberger Ltd.*, 583 F. Supp. 3d 783, 800 (W.D. Tex. 2022) *aff'd*, 81 F.4th 419 (5th Cir. 2023). Even a "short and plain" statement that Elevance believed CMS incorrectly calculated the national average would have sufficed. *See* Fed. R. Civ. P. 8(a)(2) (explaining that a pleading that states a claim for relief must contain "a short and plain statement of the claim showing that the pleader is

[3] In its response, Elevance appears to claim that case-mix adjusting CAHPS measures conflicts with the statute. *See* Pls.' Resp. Br. 8. But Elevance did not include that claim in its Amended Complaint or in its summary-judgment motion. *See* Am. Compl. ¶ 67 ("CMS violated its regulations and acted arbitrary and capriciously when it calculated CAHPS measures by first adjusting for the case-mix index.").

entitled to relief"). But Elevance's Amended Complaint provided neither notice of what its new claim was nor the underlying grounds.

2. Even if Elevance had properly pled this claim, it would fail on the merits.

Elevance claims "[t]he applicable regulations require [CMS] to calculate measure-specific Star Ratings by comparing Plaintiffs' contract to the 'national average' *contract score*." Pls. Resp. Br. 2 (emphasis added) (citing 42 C.F.R. § 422.166(a)(3)(i)-(v)). This is not what the regulation says. The regulation requires CMS to find the "national average CAHPS measure score." 42 C.F.R. § 422.166(a)(3)(i)-(v). No part of subsection 422.166(a) requires CMS to find the national average *contract* score. Elevance accuses CMS of using "linguistic gymnastics" and "sleight of hand" to mischaracterize its claim that CMS "must use the national *contract*-level average as the national average," Pls. Resp. 19-20. But, as noted, Elevance affirmatively states in its brief that it is asking this Court to require that CMS use a contract-level average. Pls.' Resp. Br. 2.

Elevance's definitions of "average" are helpful in establishing that Elevance is requesting that CMS use a contract-level average. Plaintiffs state, "Cambridge Dictionary defines 'average' as 'the result you get by adding two or more amounts together and dividing by the total number of amounts.'" Pls.' Resp. Br. 18 (citing Cambridge Dictionary). Elevance contends that the numerator, the "two or more amounts" added together, should be the sum of the contract scores for a given measure. It thinks that the denominator, the "total number of amounts," should be the total number of contracts. Pls.' Resp. Br. 18-19 ("the national average without any weighting is the simple average of all contract values . . . where each contract is weighted equally"). Whether a contract has 2,000 or 500,000 enrollees, Elevance thinks that each contract should count for one in the denominator. Adding up all of the contracts' measure scores and dividing by the total number of contracts yields a *per-contract* average measure score.

But the Secretary's regulations nowhere expressly mandate CMS to use an average of the contracts. 42 C.F.R. § 422.166(a)(3)(i)-(v). Elevance's approach to the national average would add the word "contract" such that subsection 422.166(a)(3) would say: "statistically significantly higher [or lower] than the national average CAHPS measure **contract** score." (emboldened term added). Elevance offers no explanation as to why the regulations unambiguously compel its interpretation of "national average" as the contract-level average except that doing so might push two of its contracts over the line into the next half-star category.

Conversely, CMS's approach is to find the national average CAHPS measure score, consistent with the regulation. To calculate the national average of CAHPS measure scores, CMS takes the contract-level scores for each CAHPS measure, weights those scores by beneficiary enrollment, and then averages those scores. CMS, *Summary of Analyses for Reporting, MA & PDP CAHPS* 2 (Aug. 2024). By weighting for beneficiary enrollment, CMS disaggregates beneficiaries from their contracts to avoid capturing the contract average—the regulation does not explicitly mandate the contract-level average. For CMS's national average, the numerator is effectively the sum of the scores for all enrollees across the nation. The denominator is the total number of enrollees in the nation. This approach yields a national average, consistent with the regulation.

Elevance argues that CMS "do[es] not and cannot in good faith dispute that an 'average' and 'weighted average' are two different things." Pls.' Resp. Br. 18. Not so. A "weighted average" is simply a kind of average—they are not "two different things." If the steps CMS takes to disaggregate beneficiaries from their contracts to determine an average qualifies as weighting by beneficiary enrollment as Elevance suggests, Pls.' Resp. 17, Elevance's national average contract score theory fares no better. By advocating that CMS use the national average contract score, Elevance is arguing that CMS should calculate the national average by weighting by contract,

"where each contract is weighted equally." Pls.' Resp. Br. 18-19; Pls.' App. 11-12. This is simply a different kind of weighting—one that Elevance would prefer but is not unambiguously compelled by the regulations. The Court should reject Elevance's efforts to have the Court substitute its judgment for CMS and require CMS to calculate a national average contract score.

C. Elevance Maintains Its Argument that This Court Should Require CMS to Double Round, in Violation of Traditional Rounding Rules and the Secretary's Regulations.

1. Reliance on Dr. Goldstein's Declaration is appropriate.

Elevance labels the facts that Dr. Goldstein provides in her declaration as "*post hoc* rationalizations" and repeatedly asserts that she is unqualified and lacks expertise in the subjects relevant to this litigation. Pls.' Resp. Br. 21, 39, 44.

Elevance's attack on Dr. Goldstein misses the mark for two reasons. First, her declaration is a fact-witness declaration—not an expert-witness declaration, so there was no need for her declaration to describe all her expertise and experience. Instead, her declaration stated the basis for her personal knowledge of the facts in her declaration, as required by Federal Rule of Civil Procedure 56(c)(4). Specifically, her declaration stated that her role at CMS is the Director, Division of Consumer Assessment and Plan Performance, Medicare Drug Benefit and C & D Data Group. Defs.' App. 1, Goldstein Decl. ¶ 1. And her declaration clearly stated that it was "based on her personal knowledge, information contained in agency files, and information furnished to her in the course of her official duties." *Id*. Her declaration is limited to establishing that CMS does not calculate or round to the sixth decimal, that mean resampling is random because contracts differ from year to year, that CMS, with its contractor RAND, undertook its own simulation establishing that changing the seed used in the Star Ratings methodology would not change the 2025 overall Star Ratings for 92.4% of contracts on average across simulation seeds, and that

CMS's use of weighted, case-mix adjusted means for CAHPS measures results in unbiased assignment of contract scores to base groups. *See id.* 1-3 (Goldstein Decl. ¶¶ 2-6).

Reliance on her declaration is appropriate. "[A] court may obtain from the agency, either through affidavits or testimony, such additional explanations of the reasons for the agency decision as may prove necessary," provided that any new materials are "merely explanatory of the original record and should contain no new rationalizations." *See Kirwa v U.S. Dep't of Defense*, 285 F. Supp. 3d 257, 269 (D.D.C. 2018). Here, Dr. Goldstein's declaration does not contain new rationalizations; instead, it further explains the original records of CMS's decisions and rebuts Elevance's extra-record evidence, Pls.' App. 1-77. In other words, Dr. Goldstein's explanations support what the Administrative Record amply supports—most critically, that CMS does not calculate or round to the sixth decimal.

Second, Elevance's attacks on Dr. Goldstein's expertise are baseless. If CMS did wish to offer her as an expert declarant, she would be more than qualified.[4]

2. CMS rounded in accordance with traditional rounding rules.

CMS's application of rounding rules in this case asks a basic question: is Elevance's overall score of 3.749565 closer to 3.5 or 4? Elevance directs this Court to that simplified articulation in the Administrative Record. Pls.' Resp. Br. 24. For some reason, helpfully, Elevance even

[4] Dr. Elizabeth Goldstein has over 30 years of experience in survey research, quality measurement, long-term care research, and public policy. She has worked at CMS since 2000 and leads the Medicare Advantage Star Ratings program, oversees the national implementation of CAHPS, and directs the administration of all other division data collection efforts. (Supp. App. to Supp. Defs.' Summ.-J. Reply ("Defs.' Supp. App."), at 1, 3-10, ECF No. 45-1(Supp. Decl. of Elizabeth Goldstein ¶ 1 & Ex. 1).) Dr. Goldstein received her Ph.D. in Economics from the University of Wisconsin in 1993. (Defs.' Supp. App. at 3.) By virtue of her training and experience, she has significantly more expertise and experience relevant to the Part C and D Star Ratings and CAHPS than Elevance's experts. (These facts about Dr. Goldstein's experience and qualifications are contained in the Defendants' proposed Supplemental Appendix to Support Defendant's Summary-Judgment Reply. (Defs.' Supp. App. 1 ¶ 1, 3-10.) Defendants have sought leave to file this proposed supplemental appendix in a separate motion, filed contemporaneously with this reply brief.).

highlighted it. On page 24 of its response and reply brief, Elevance excerpted Table 22 of the 2025 Technical Notes, A.R. 31. Defendants also excerpted a version of this chart in its own brief. Defs.' Br. 29. This section of the 2025 Technical Notes is entitled, "Rounding Rules for Summary and Overall Ratings." *See* Pls.' Resp. Br. 24. The highlighted text states, "Table 22 summarizes the rounding rules for converting the Part C and D summary and overall ratings into the publicly reported Star Ratings." *Id.* Below the table, the highlighted text states in part: "For example, a summary or overall rating of 3.749999 rounds down to a rating of 3.5, and a rating of 3.750000 rounds up to rating of 4." *Id.* So, under this simplified articulation of CMS's rounding rules, if Elevance had achieved an overall score of 3.749999, its overall rating would *still* round down to 3.5. Elevance did not achieve an overall score that high—3.749565 is less than 3.749999. Returning to the question at issue, is Elevance's overall score of 3.749565 closer to 3.5 or 4? It is closer to 3.5. It should consequently be rounded to 3.5. Nothing Elevance has said in this litigation changes that basic mathematical truth.

Elevance's position is that the traditional rules of rounding require CMS to round the second digit after the decimal, the hundredth place, and to do so, to look to the third digit after the decimal, the thousandths place. Pls.' Resp. Br. 23. Thus, under Elevance's reading of the traditional rounding rules, 3.749 rounds to 3.75. *Id.* CMS's position is that the traditional rules of rounding require CMS to round to the nearest half-star by rounding the first digit from the decimal, the tenths place, and to do so, look to the second digit after the decimal, the hundredths place. Thus, under CMS's reading, 3.74 rounds to 3.5. Because CMS looks to the hundredth decimal place to round the tenth decimal place, all of the other digits that follow the hundredth decimal place are inconsequential. Put differently, nothing that happens at the sixth decimal point or later impacts CMS's rounding determinations.

Why is CMS's position consistent with the traditional rules of rounding and Elevance's not? Consider again the definition of traditional rounding rules at 42 C.F.R. § 422.162(a). It says that "[t]raditional rounding rules mean that the last digit in value will be rounded. If rounding to a whole number, look to the digit in the first decimal place." *Id*. § 422.162(a). The value being rounded is not being rounded to a whole number. It is a value with a digit in the tenths place, such as 3.0, 3.5, 4.0, etc. This is because in specifying that CMS determine overall quality ratings for plans on a 1- to 5-star scale ranging from 1 (worst rating) to 5 (best rating) in half-star increments, Congress permitted CMS to round to the nearest half-star, or 0.5. *See* 42 U.S.C. § 1395w-23(o)(4)(A) ("The quality rating for a plan shall be determined according to a 5-star rating system."); § 1395w-24(b)(1)(C)(v) (indicating a 5-star rating system on half-star increments). Consequently, the rule's instruction that when rounding to a whole number "look to the digit in the first decimal place," should be shifted one place over to account for the tenths place digit. CMS is rounding the first-place digit and is required to look to the digit in the second decimal place to see whether it is less than or equal to 4, or greater than or equal to 5. *See* § 422.162(a).

The highlighted table, Table 22, Pls.' Resp. Br. 24, A.R. 31, is simply a visual representation of this rounding methodology. Elevance maintains that this table demonstrates that CMS is "not applying traditional rounding rules at all, but rather 'converting' numbers rounded to the sixth decimal to an Overall Star Rating." Elevance does not explain how. The overall scores here are displayed with six decimal places for practical purposes only. Defs.' App. 2, ¶ 2 (Goldstein Decl. ¶ 2). Using six decimal places for display purposes does not indicate rounding in the calculation of the overall rating. Elevance fails to point to any part of the Administrative Record, any regulation, any guidance, that supports its contention that CMS is rounding the sixth decimal place of an overall Star Rating. Elevance cites the rounding rules applied to calculation of measure

scores and asserts that there is a "dichotomy between what Defendants do for Overall Star Ratings . . . versus the measure scores." Pls.' Resp. Br. 25. The critical difference is that measure Star Ratings are whole numbers—there are no half-star increments for measure scores (unlike for overall Star Ratings). *See* 42 C.F.R. § 422.166(a)(4) ("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."). Elevance repeatedly describes CMS's explanation of its rounding rules as "*post hoc* rationalizations." *See, e.g.*, Pls.' Resp. Br. 23. CMS explained how its rounding rules worked in the 2025 Technical Notes that predate this litigation. That guidance is consistent with the regulation defining traditional rounding rules.

In its response and reply, Elevance again asks this Court to require CMS to double round. It says, "the only possible conclusion is that 3.749565 rounds to 3.75, which is 4 Stars." Pls.' Resp. Br. 28. Under Elevance's preferred rounding methodology, the first rounding step would be to round 3.749565 to 3.75, and the second rounding step would be to round 3.75 to 4. At no point in its brief does Elevance ever deny that it is advocating double rounding. Elevance asserts that CMS's double rounding argument is "a red herring" and "predicated upon a flawed premise." *Id.* But it fails to explain why. Double rounding is at odds with the Secretary's requirement that CMS use traditional rounding rules. CMS's regulation mandating the use of traditional rounding rules, Elevance contends, "does not foreclose CMS from applying its own policy to determine whether the rounded number is above, below, or equal to the midpoint of .75 or .25." *Id.* But that is exactly what the traditional rounding rules definition does—it commits CMS to a policy of rounding values once, traditionally. The regulatory definitions are meant to constrain CMS. *See* 42 C.F.R. § 422.162(a). It cannot subsequently come up with its own *ad hoc* rounding rules. And it certainly cannot apply rounding rules that are arbitrary and statistically dubious.

Because it is so central to its claims, Elevance insists again and again that CMS rounds to the sixth decimal place. *See, e.g.*, Pls.' Resp. Br. 24, 26, 27, 28, 28, 30, 32, 34. There is no record evidence to support this assertion. Again, CMS rounds overall Star Ratings to the nearest half-star for overall and summary ratings. Elevance contends that CMS's use of the "FIXED" function in Microsoft Excel shows CMS does round to the sixth decimal place. Pls.' Resp. Br. 26-27. CMS, however, does not use Microsoft Excel to calculate the Star Ratings. CMS explained in its 2025 Technical Notes that "CMS and its contractors have always used software called SAS (an integrated system of software products provided by SAS Institute Inc.) to perform the calculations used in producing the Star Ratings," not Microsoft Excel. A.R. 30. CMS also explained that "[c]ontracts may request a contract-specific calculation spreadsheet which emulates the actual SAS calculations." *Id.* Elevance and its expert, Dr. Diver, are relying on Excel spreadsheets meant to help contracts understand the Star Ratings calculations. The Excel "FIXED" function and Dr. Diver's analysis, Pls.' Supp. App. 325-36, have no application to CMS's Star Ratings calculation methodology.

Finally, Elevance alleges s that instead of following traditional rounding rules, CMS is "converting" numbers rounded to the sixth decimal place to an Overall Star Rating. Pls.' Resp. Br. 24, 25, 32. Elevance does not explain what it means by this. As demonstrated, CMS rounds in accordance with traditional rounding rules. The Court should reject Elevance's contentions that it rounds to the sixth decimal place or is otherwise at odds with traditional rounding rules.

D. By Rounding to the Half-Star Increment, CMS Accounts for Imprecision.

Elevance continues to insist that CMS calculates and rounds to the sixth decimal place. Because CMS does not do so, this Court should reject Elevance's arguments pertaining to precision. Additionally, CMS's Star Ratings calculation methodology is in fact precise.

1. CMS does not calculate or round to the sixth decimal place.

Elevance's arguments pertaining to precision rest on its incorrect premise that CMS "calculated and rounded the Final Summary Score to the sixth decimal, and in doing so, completely ignore[d] the imprecision inherent in the statistical methodology." Pls.' Resp. Br. 30. Elevance calls this an "illusion of precision." *Id.* Elevance's (incorrect) premise that CMS calculates and rounds to the sixth decimal place is the entire basis for *all* of Elevance's arguments related to precision. But, as established, CMS does not round to the sixth decimal place. And as also established, CMS generally calculates numerical scores with at least six digits of precision and publicly displays them to the sixth decimal place. A.R. 30-31; Defs.' Br. 26. Elevance states that CMS is "mistaken [in its] belief that it truncates . . . at the sixth decimal place." Pls.' Resp. Br. 28. CMS neither calculates to nor truncates measure scores at the sixth decimal place. In calculating measure scores, CMS carries at least six decimal places of data. "CMS calculates each step out to the sixth decimal place because nothing that happens at the sixth decimal place ever impacts the second, hundredth decimal place. Practically, this is no different than if the calculations were completed with infinite decimal places but is computationally more efficient." Defs.' App. 2 ¶ 2 (Goldstein Decl. ¶ 2).

Elevance's position is that "calculating and rounding a plan's Final Summary Score to the second decimal place, as opposed to the sixth, is better reflective of the uncertainty which exists due to random chance in Defendants' methodology to determine an Overall Star Rating." Pls.' Resp. Br. 30. CMS does Elevance one better—it rounds overall Star Ratings to the first decimal place, to the half-star increment. By rounding to the nearest half star, or 0.5, CMS has sufficiently accounted for uncertainty due to random chance.

Elevance's contention is that a very small amount, 0.000435, could be determinative as to whether a contract receives 3.5 or 4 stars. Pls.' App. 26. But the Star Ratings are not so sensitive.

Elevance recognizes CMS's argument that "by allegedly rounding at the second decimal, [CMS] account[s] for that imprecision," Pls.' Resp. Br. 32, but it does not seriously engage with it. Elevance falls back on its refrain that CMS calculates overall scores to six decimal places. As demonstrated, CMS does not. Elevance takes issue with the fundamental truth that in *any* Star Ratings system, cutoffs are a requirement. Contracts will necessarily miss those cutoffs by small amounts. That Elevance missed the cutoff between 3.5 and 4 stars by a small amount is a fact that is unrelated to CMS's rounding methodology.

Because CMS does not round to the sixth decimal place, the remainder of Elevance's claims pertaining to precision fail. Elevance's argument is that rounding to the sixth decimal is improper because doing so does not account for imprecision and statistical variance in CMS's calculation—specifically, for CAHPS measures, the alleged failures to consider nonresponse bias and sampling and measurement error. *See* Pls.' Resp. Br. 34 (alleging CMS fails to account for inherent imprecision when calculating and rounding to the sixth decimal place and arguing that the source of that imprecision includes nonresponse bias, sampling error, and measurement error). Elevance did not bring freestanding arbitrary and capricious claims related to nonresponse bias and sampling error. Nonresponse bias, sampling error, and measurement error are not mentioned in Elevance's complaint; only Elevance's "illusion of precision" theory. *See* Am. Compl. ¶ 38. This lawsuit is Elevance's effort to nudge their scores over the line into the next half-star category, *see id.* 23 (requesting that this Court "[o]rder Defendants to recalculate the score for contract H3655 to 3.75 and corresponding Star Rating to 4.0 Stars"). It is unclear how a finding by this Court that, unrelated to rounding to the sixth digit, CMS failed to adequately consider nonresponse bias, sampling error, or measurement error would provide a basis for the Court to order CMS to increase Elevance's Star Ratings. Put differently, Elevance makes no attempt to explain how

nonresponse bias, sampling error, and measurement error are problematic in light of the fact that by rounding to the nearest half star, or 0.5, CMS has accounted for uncertainty from random chance. If the Court agrees that the Agency does not round to the sixth decimal place to determine the overall rating, it need not reach any of Elevance's remaining claims related to precision.

2. CMS's Star Ratings calculation methodology is sufficiently precise.

CMS's contractor, RAND, "conducted its own larger, more accurate simulation" as compared with Dr. Diver's simulation. Defs.' Br. at 35–37. Elevance contends that the Court should not consider this analysis because Dr. Goldstein is not being offered as an expert declarant. Pls.' Resp. Br. 31. Elevance additionally contends that Dr. Goldstein is describing CMS's contractor, RAND's statements and would-be testimony, so her declaration amounts to hearsay. A "declaration used to support or oppose a motion must be made on personal knowledge, set out facts that would be admissible in evidence, and show that the . . . declarant is competent to testify on the matters stated." *See* Fed. R. Civ. P. 56(c)(4). Again, Dr. Goldstein is being offered as a fact witness. Her knowledge of RAND's study is not based on RAND's statements, but her personal knowledge, work, and collaboration with CMS's contractor, RAND, in her role as a director at CMS. Defs.' App. 1, 3.

Elevance contends that RAND's simulation shows that CMS "calculate[s] and assign[s] Star Ratings based upon nothing more than the roll of the dice." Pls.' Resp. Br. 32. The RAND simulations show that 92.4 percent of contracts receive the same rating on average across seeds. Defs.' App. 3. This is not "a roll of the dice." Elevance admits that not only was Dr. Diver's number of simulations smaller, but critically, unlike RAND's simulations, Dr. Diver's simulations failed to consider the reward factor threshold as part of his simulations and the Parts C and D improvement measure scores. Pls.' Resp. Br. 33–34. These limitations prevent Elevance from drawing any meaningful conclusions regarding the impact of different seeds.

Elevance further contends CMS failed to sufficiently consider nonresponse bias. Pls.' Resp. Br. 34–39. The record demonstrates that CMS sufficiently considered the issue of nonresponse bias in CAHPS surveys and that the procedures CMS employs sufficiently address nonresponse bias. Elevance misconstrues CMS's obligations under the Administrative Procedure Act (APA). An agency must "examine the relevant data and articulate a satisfactory explanation for its action including a rational connection between the facts found and the choice made." *State Farm*, 463 U.S. at 43 (internal quotation marks omitted). An agency has acted arbitrarily and capriciously only when "it has entirely failed to consider an important aspect of the problem, offered an explanation for its decision that runs counter to the evidence before the agency, or is so implausible that it could not be ascribed to a difference in view or the product of agency expertise." *Id.* Even a decision that is not fully explained may be upheld "if the agency's path may reasonably be discerned." *Bowman Transp., Inc. v. Ark.-Best Freight Sys., Inc.*, 419 U.S. 281, 286 (1974).

CMS stated "meta-analyses of surveys that follow the rigorous probability sampling and survey approaches used by MA and PDP CAHPS find little relationship between response rates and nonresponse bias." 83 Fed. Reg. at 16,555. CMS continued, "research specific to patient experience, CAHPS, and [Medicare Advantage] and [Prescription Drug Plan] CAHPS surveys finds no evidence nonresponse bias affects comparison of case-mix adjusted scores between contracts or other similar reporting units." *Id.* at 16,555 (citing five studies related to nonresponse bias). Despite clear evidence that CMS considered nonresponse bias, Elevance states "[t]he Administrative Record is *entirely devoid of any evidence* that [CMS] considered or investigated nonresponse bias." Pls.' Resp. Br. 36 (emphasis added). Not so. Elevance implies that CMS only considered the topic of nonresponse bias in hospital CAHPS data and not CAHPS data for

Medicare Advantage Organization surveys. *Id.* But the 2018 Final Rule explicitly discusses Medicare Advantage-related CAHPS surveys, not hospital surveys. 83 Fed. Reg. at 16,555.

Additionally, CMS's 2020 Final Rule observed that "[the case-mix] adjustment ensures that contract-level scores fairly represent all contracts. Analyses of nonresponse in CAHPS data (Elliott et al. 2005; Elliott et al. 2009) have shown little or no evidence of nonresponse bias in the presence of CAHPS case-mix adjustment." Medicare Program; Contract Year 2021 Policy and Technical Changes to the Medicare Advantage Program, 85 Fed. Reg. 33,796, 33,840 (June 2, 2020). Although the two Elliot papers CMS cites involve Hospital CAHPS data, Elevance's expert declarant, Dr. Lavrakas, fails to establish that real differences between hospital surveys and the Medicare Advantage surveys at issue in this case prevent CMS from drawing conclusions about the efficacy of the case-mix adjustment in preventing nonresponse bias.

While it is true that the Bland analysis was not cited in the 2018 rule, as Elevance points out, Pls.' Resp. Br. 37–38, and CMS inadvertently omitted it from the Administrative Record it filed in this case,[5] it is a study upon which CMS relied in both evaluating nonresponse bias and making determinations concerning case-mix adjustments. Defs.' Supp. App. 2, ¶ 3 (Supp. Dec. of Elizabeth Goldstein ¶ 3). In that analysis, researchers found that "[a]fter the [case]-mix adjustment process is completed, the survey results no longer have nonresponse bias, which has been shown in some [non-Medicare Advantage and Prescription Drug Plans] CAHPS surveys." A.R. 1693 (Cynthia Rae Bland et al., *Challenges Facing CAHPS Surveys and Opportunities for Modernization* (RTI Press 2022)). The goal of case-mix adjustment is to permit unbiased comparison of contract performance. A.R. 114. It does so by adjusting for any differences in respondents across contracts that are relevant to patient experience, whether those differences arise

[5] Defendants are filing a supplement to the Administrative Record, contemporaneous with this reply, to correct this inadvertent omission by adding the study. Supp. to A.R., ECF No. 46.

from nonresponse bias differences in enrollment by enrollee characteristics, or other causes. *Id.* Across multiple rules, CMS has explained that case-mix adjusting measure scores sufficiently addresses multiple forms of bias, including nonresponse bias.

On the issue of sampling error, Dr. Lavrakas states that "the size of the margin of sampling error, and resulting Confidence Intervals, in all CAHPS surveys are such that they make invalid any comparisons of means of percentages smaller than several integers in the size of the difference." Pls.' Supp. App. 347. Consequently, according to Dr. Lavrakas, "it is statistically incorrect to draw conclusions about survey-based differences that are far less than one integer in size, let alone differences that are as small as CMS does for some of Elevance Health plans." *Id.* Dr. Lavrakas premises his expert opinions on the incorrect notion that CMS rounds at the sixth decimal; but as discussed above CMS does not premise its Star Ratings on small differences. Dr. Lavrakas asserts that "[i]n particular, no CAHPS survey data can validly be used to conclude that 3.749565 is reliably different from 3.750000." *Id.* As established, the difference for Elevance's H3655 contract between a 3.5 Star Rating and a 4.0 Star Rating is not 0.000435, as Dr. Lavrakas assumes—it is 0.249565. Defs.' Br. 33. Rounding error is the difference between an unrounded value and the value that it is rounded to. If 3.749565 were rounded to 4 as Dr. Lavrakas suggests, the rounding error is 4 minus 3.749565, or 0.250435. Clearly, the rounding error is greater when rounded to 4 using double rounding. Elevance tellingly does not contend with the fact all of Dr. Lavrakas's conclusions regarding sampling error are premised on its incorrect assertions regarding the calculation and rounding of Star Ratings.

As established, CMS's CAHPS scoring methodology accounts for uncertainty created by potential sampling error in a principled manner. Defs.' Br. 40. When measure scores are below the 15th percentile (in base group 1), significantly below average, and have low reliability, 1 star will

be assigned if and only if the measure score is at least 1 standard error below the unrounded cut point between base groups 1 and 2. 42 C.F.R. § 422.166(a)(3). Similarly, when the measure score is at or above the 80th percentile (in base group 5), significantly above average, and has low reliability, 5 stars would be assigned if and only if the measure score is at least 1 standard error above the unrounded cut point between base groups 4 and 5. *Id.* § 422.166(a)(3); A.R. 151-55.

With regard to measurement error in CAHPS survey data, Elevance's argument is again premised on incorrect notions about how CMS calculates and rounds to overall Star Ratings. Elevance states that "[e]ven though many CAHPS measures may have strong reliability, all are imperfect and therefore their usage to calculate Star Ratings result in Star Ratings that contain errors that are larger than one millionth of a decimal place." Pls.' Resp. Br. 41. As established, errors that are "larger than one millionth of a decimal place" are not problematic in light of the fact that CMS rounds to the nearest half-star, or 0.5. Dr. Lavrakas's opinions otherwise amount to an assertion that CAHPS measures "are imperfect." Pls.' Supp. App. 349. "The standard of review under the arbitrary and capricious test is only reasonableness, not perfection." *Kennecot Greens Creek Min. Co. v. Mine Safety and Health Admin.*, 476 F.3d 946, 954 (D.C. Cir. 2007). The regulations conceived of the possibility that measures may benefit from improvements: they require CMS to solicit feedback for new measures before proposing and finalizing them, § 422.164(c)(2), and to follow notice-and-comment rulemaking for substantive updates, § 422.164(d). The regulations confer significant discretion upon CMS to engage in an iterative process to develop and update measures such that they may be "nationally endorsed," "align[ed] with the private sector," "appropriate to measure," and "reflect[ive of] performance specific to Medicare." § 422.164(c)(1). CAHPS measures are not so imperfect or "statistically incorrect" that they do not sufficiently account for nonresponse bias, sampling error, or measurement error.

E. Because Contracts Vary from Year to Year, CMS's Mean Resampling Is Random.

In its Summary Judgment Brief, CMS established that its mean resampling is random because its list of considered contracts does not remain the same from year to year. Defs.' Br. 43–44. Even though CMS uses the same seed every year, the 8-6-7-5-3-0-9 seed has different effects in different years because it is applied to different data from different contracts. Elevance's response is to say that the fact that contracts change from year to year is a "fortuitous alignment of outside factors" over which CMS has no control. Pls.' Resp. Br. 42. In essence, Elevance argues that CMS's methodology would not ensure randomness if hypothetically contracts did not change from year to year. But the number of contracts from year to year never stays the same. Defs.' Supp. App. 2 ¶ 2 (Supp. Decl. of Elizabeth Goldstein ¶ 2). Consider, for example, the contract changes for the last five years with regard to the Breast Cancer Screening measure. *Id.* In 2025, there were 478 contracts that were the same as the previous year. *Id.* There were 63 new contracts as compared with the prior year. *Id.* Forty-three contracts no longer existed as compared with 2024. *Id.* These kinds of changes are reflected for the other years. *Id*. CMS's methodology has actually resulted in randomness because contracts have in fact changed year to year.

More fundamentally, Elevance makes no showing that CMS did not sufficiently ensure randomness in the mean resampling CMS performed with regard to the specific contracts at issue in this case. Elevance, for instance, does not assert and cannot show that the number of contracts for 2025 are not different than for 2024.

III. Conclusion

For the reasons described above and in their earlier Summary Judgment Brief, Defendants respectfully request that the Court grant Defendants' cross-motion for summary judgment and deny Plaintiffs' motion for summary judgment.

Respectfully submitted,

CHAD E. MEACHAM
ACTING UNITED STATES ATTORNEY

/s/ Andrea Hyatt
Andrea Hyatt
Assistant United States Attorney
Texas Bar No. 24007419
Burnett Plaza, Suite 1700
801 Cherry Street, Unit #4
Fort Worth, Texas 76102-6882
Telephone: 817.252.5230
Fax: 817.252.5458
Andrea.Hyatt@usdoj.gov

Attorneys for Defendants

Certificate of Service

On April 21, 2025, I electronically submitted the foregoing document with the clerk of court for the U.S. District Court, Northern District of Texas, using the electronic case filing system of the court. I hereby certify that I have served all parties electronically or by another manner authorized by Federal Rule of Civil Procedure 5(b)(2).

/s/ Andrea Hyatt
Andrea Hyatt
Assistant United States Attorney