



AMERICAN ACADEMY of ACTUARIES

Objective. Independent. Effective.™

November 24, 2021

Allison Yadsko

Deputy Director of the Division of Policy and Analysis, Financial Management Group
Center for Consumer Information and Insurance Oversight (CCIIO)
Centers for Medicare & Medicaid Services (CMS)
Department of Health and Human Services (HHS)
200 Independence Avenue SW
Washington, DC 20201

Re: Comments on the 2021 Risk Adjustment Model Update Technical Paper

Dear Ms. Yadsko,

On behalf of the American Academy of Actuaries (Academy)¹ Risk Sharing Subcommittee of the Individual and Small Group Markets Committee, we appreciate the opportunity to provide comments on the *HHS-Operated Risk Adjustment Technical Paper on Possible Model Changes*, [exposed for comment](#) on Oct. 26, 2021.

Comments on expanded explanation of refinements to the risk adjustment model specification

The bulk of this technical paper is focused on improving the risk adjustment model initially proposed in October 2020 as part of the proposed HHS Notice of Benefit and Payment Parameters for 2022 (Payment Notice). We provided [comments](#) on these proposals and thank CMS for publishing additional supporting detail for these potentially meaningful proposed changes to the risk adjustment model.

Importance of transfer results and the choice of data used to evaluate model performance

As noted in our previous [comment letter](#), the best measure of performance of the risk adjuster is its effect on transfer payments and issuer financial outcomes. We recognize that CMS leveraged certain data to produce these results, but CMS has not published results. It is extremely challenging to provide robust commentary on the proposed changes in the absence of this data, and we request an extension or reopening of the comment period once modeled transfer results are published, which CMS has indicated will be in late 2021.

¹ The American Academy of Actuaries is a 19,500-member professional association whose mission is to serve the public and the U.S. actuarial profession. For more than 50 years, the Academy has assisted public policymakers on all levels by providing leadership, objective expertise, and actuarial advice on risk and financial security issues. The Academy also sets qualification, practice, and professionalism standards for actuaries in the United States.

We note that model performance was typically evaluated on the data set used for calibration of the coefficients. This limits the usefulness of evaluation of the risk adjuster as a predictive model—at best, the analysis performed can tell us how well the model fits the data used to create it. It has limited use with regard to informing stakeholders of how the model may perform. To better evaluate the risk adjuster, we would have liked to see how the 2016, 2017, and 2018 models performed on separate data, such as 2019 or 2020 EDGE data. Measuring a model based on the data used to calibrate it can result in overfitting of parameters and can create the illusion of performance improvements that do not hold when applied to different data. One common practice to avoid these concerns is splitting the sample into a training set and a test set. The model is calibrated using the training set and evaluated using the test set. We understand that CMS tested splitting the sample for calibration and validation but elected to report the full set numbers, citing minor levels of variation in the predictive ratios between the subsamples. We note the appeal of generating coefficients with a larger calibration set but are curious as to the degree of variation experienced in the results and how the training and test sets were selected. Given the degree of variation in the underlying data across the country, we expect that some small variation is present—if results are truly unchanged to any noticeable degree, then we would want to ensure there was no technical mistake in the training and test set procedures.

While not specifically part of the comment paper, we note that CMS' lack of access to geographic region data sharply limits CMS' ability to evaluate the performance of the risk adjuster and to evaluate the extent to which certain features identified in analysis of this national risk score data may be unique to one region or another. We would support CMS' efforts to expand the scope of data collection and retention, particularly for geographic data, even if such information was not disclosed in future publications of the EDGE data research dataset and was only used to facilitate more complete evaluations of risk score or the risk adjustment methodology updates. For instance, such data would allow CMS to evaluate predictive ratios for different cohorts in different states, which could help federal regulators understand where specific modifications to the risk adjuster may be broadly accurate but where state-level patterns of practice and care may introduce additional undesired variation.

Comments on specific changes

CMS evaluated changes intended to improve the risk adjusters' performance for enrollees in particular expenditure deciles in response to concerns that the current model calibration underpredicts expenditures for the lowest-cost and highest-cost enrollees. As outlined in the proposed 2022 Payment Notice, two changes target higher-cost enrollees and one change targets lower-cost enrollees.

We appreciated the additional explanation surrounding enrollment duration factors, particularly regarding the individual and small group markets and the reasons for the lack of full-year enrollment. We remain concerned that the reasons for partial-year enrollment in the individual and small group markets differ in ways that would result in variations in risk scores by market, though this potential impact is not immediately apparent in the values published in the CMS report. For instance, the effect on small group health plans may be more pronounced for groups with a renewal date in the second half of the year when an employee changes issuers, where the new issuer may experience reduced claim levels. It is unclear whether this situation is prevalent enough to influence the overall values shown in the report.

Similarly, we appreciated the detail surrounding the interacted Hierarchical Condition Categories (HCC) counts approach to severity. We recognize that CMS has selected conditions for this interacted HCC counts approach to limit the ability for issuers to over-code conditions, but we note that there is still at least some potential for issuers to target risk score coding efforts for these specific conditions. Given the significant benefit to each additional eligible HCC identified, this behavior could produce meaningful shifts to risk transfers. We also note this may create more challenges in the risk adjustment data validation process due to the increased importance of this list of HCCs on overall risk scores and risk transfers and would appreciate further guidance on the interaction of condition identification and error rates on these interacted HCC severity indicators.

We note that these two proposals in tandem improve CMS' measure of the predictive ratios for both low- and high-cost deciles, as well as the model's R-squared value. Based on the data provided, we are largely supportive of implementation of both proposals in tandem, though we note small group market issuers with non-calendar-year business may prefer the original enrollment duration factor approach due to the timing issue noted above. While we would have liked to see more analysis on independent data, the resulting factors appear sound on a theoretical basis and in terms of measured performance against the model calibration data.

With regards to the two-stage model, we appreciate the improvement in predictive ratios, particularly in combination with the interacted HCC counts severity factors and modified enrollment duration factors. However, we note that the methodology is somewhat contrived and modestly reduces overall model fit and accuracy when measured against the data used to calibrate the model. It is particularly important to review the implications to payment transfers when evaluating the appropriateness of this specific change.

Value of practicality relative to theoretical soundness

The primary goal of risk adjustment is premium stabilization. Premiums are most stable when risk adjustment appropriately offsets the costs of risk selection so that an insurer is ambivalent to any characteristics (such as age, income level, health status, or preferred metal tier) of any individual who chooses to enroll. In the absence of modeled transfer results, it is extremely challenging to evaluate how the model changes may create more or less stable financial outcomes. Appropriate financial outcomes are ultimately measured not by accuracy of the risk adjuster as measured against the data used to calibrate the model, but by how the updated risk scores work in concert with other elements of the transfer formula to limit the financial impact of risk selection on issuers. A sound theoretical framework for the risk adjuster can help instill confidence in the marketplace that the model is less likely to produce inaccurate results when used to adjust actual health plan experience. Such a structure should align the methodology determining the risk scores with the actual factors that influence risk assumed by health plans. The lack of such an alignment, though, should not necessarily prohibit any modification improving the risk adjuster, provided the methods produce meaningful improvements to the true predictive power of the model.

Implications for pricing

Both anecdotal evidence and Affordable Care Act (ACA) risk adjustment research² indicate that the ACA's risk adjustment structure generally overcompensates issuers for enrollees with risk adjustment-eligible conditions and undercompensates issuers for enrollees without risk adjustment-eligible conditions. This may not be optimal from a risk stabilization perspective, as material misalignments between risk adjustment results and the underlying risk of the population could result in issuers incorporating an additional risk premium into their rate filings to offset the increased uncertainty in financial outcomes. Higher premium levels could decrease the health insurance value for the healthiest enrollees, increasing the overall average health status of the single risk pool and, thus, potentially raising gross premiums. We note that this value dynamic is somewhat different for lower-income individuals as a result of the ACA's premium subsidies, particularly the enhanced subsidies offered in 2021 and 2022 pursuant to the American Rescue Plan Act. As increases in risk premium have limited effect on the net premium for individuals with subsidies, any influence of increased morbidity is likely to be muted, particularly because most individuals enrolled in the ACA's individual market are currently subsidized. However, a meaningful number of enrollees still lack access to premium subsidies, and the availability of affordable coverage to the unsubsidized is an important consideration.

With this in mind, we are commenting on the statement that one benefit of the proposed model changes would be the additional incentives for issuers to “create plans that encourage [healthier] individuals to enroll in coverage and improve [ACA] risk pools.” We note that plan designs and incentives in the ACA market (including performance of the risk adjuster with regards to enrollees receiving cost-sharing reductions [CSRs]) currently place significant incentives for an issuer to offer the lowest-cost or second-lowest-cost silver plan. Further, healthier individuals are generally expected to choose less expensive health plans, which aligns with issuer incentives to offer lower-cost silver plans. As such, the market currently does not particularly discourage healthy individuals (particularly healthy lower-income individuals participating in the individual market whose premium would not change due to premium subsidies) from enrolling in these plans. It can be argued that the current pattern of overcompensating for less healthy individuals (except for, as noted, the highest-cost individuals) serves as an incentive for issuers to continue to offer robust coverage despite incentives to offer the lowest-cost products; these products typically have the leanest benefits and/or narrowest networks.

We do not take a stance on which of these two arguments carries more weight. However, we note that a risk adjustment model favoring healthier enrollees over less-healthy enrollees would provide a meaningful incentive for insurers to engage in some form of selection based on health status, which is contrary to the general market reform principles contained in the ACA. We do not believe this was CMS' point regarding the statement, however, and we expect the intention was to reinforce the principle that a model that more appropriately accounts for health status will limit the disincentives for issuers to attract or exclude a specific population based on health status. We generally support changes that improve the alignment of results of the risk adjustment program with the costs incurred for individuals of any health status. Generally, we

² For example, see [“When adverse selection isn’t: Which members are likely to be profitable \(or not\) in markets regulated by the ACA”](#); Milliman Healthcare Reform Briefing Paper; December 2013.

would not favor reforms that make it untenable for issuers with high-cost enrollees to offer a competitively priced product.

Comments on CSR requests

Appendix A of the report addresses concerns and potential options surrounding risk adjustment related to individuals receiving cost-sharing reductions in the individual market. Some of these changes are driven by the October 2017 decision to cease payment of federal subsidies for these enrollees, and we note that further changes to federal policy regarding subsidy payments could significantly influence the appropriateness of modifications to the risk adjuster. However, we also note that not all dynamics surrounding CSR pricing are directly related to payment or nonpayment of these subsidies. Some of these changes are related to how CSR eligibility criteria may influence the health care service utilization of these enrollees. Additionally, because silver plans set the benchmark used to determine premium subsidies for all enrollees and because rate levels affect the enrollment decisions for a large portion of the individual market, competitive forces also place significant pressure on premiums.

Comments on the observed utilization of CSR enrollees

First, we provide general comments regarding CMS' data, particularly related to the relative utilization levels of CSR enrollees versus non-CSR enrollees. We note that this data combined with the additional utilization assumed by the risk score model would be expected to result in considerable risk adjustment advantages for plans with significant enrollment in the non-American Indian/Alaska Native silver CSRs should the federal government return to funding CSRs or whether CSR risk scores are further modified to directly incorporate CSR actuarial values without making any additional adjustments to the CSR induced demand factor. We also note that the current CSR induced demand factors appear to roughly offset the additional liability plans incur as a result of the lack of federal funding for CSR, which in turn implies that the "with risk" portion of the risk adjustment transfer formula reasonably captures the relative plan liability for CSR and non-CSR plans. However, we also note that this is not the intent of the current formula, and appreciate CMS' further discussions in this paper about formalizing this result.

Rating term analysis

While the "with risk" component of the risk adjustment transfer formula appears to be working reasonably well, it only represents one-half of the transfer formula. Currently, the "without risk" component has no special modifications for CSRs. CMS notes that the use of the 70% actuarial value (AV) for silver coverage could result in plans with silver enrollees receiving higher risk adjustment receipts or paying lower risk adjustment charges than if the AV reflects the average liability considering the CSR variants, which are being built into silver rates in many states. This in turn puts significant pressure on plans participating in the individual market to be the lowest-cost or second-lowest-cost silver plan in order to obtain these enrollees, as their favorable risk adjustment experience would then ultimately make it possible to reduce premiums further while putting additional strain on financial performance of issuers who fail to meet this threshold. This likely creates some subsidization of silver rates with premiums from other metal tiers, distorting

the underlying actuarial value relationship between these plans. However, modifications to this “without risk” component may still result in a misalignment with observed premium relativities due to the competitive considerations that tend to be particularly associated with silver plan pricing.

Potential equity considerations of modifying the CSR-induced utilization factors

At the same time, we have a concern with a statement made in the technical paper. Specifically, CMS stated the data provided in the paper do not “[imply] a lack of evidence of higher induced demand associated with receipt of CSRs for most CSR enrollees.” The data, in actuality, do imply a lack of evidence that CSR enrollees have higher induced utilization than non-CSR enrollees, but there is no viable mechanism available to associate this solely with the receipt of CSRs. In general, lower-income individuals have been shown to have a significantly higher threshold for determining when a service is sufficiently valuable to be utilized. As such, we would generally anticipate much lower utilization among low-income enrollees when they are enrolled in the same health benefit as someone with higher income. In order to determine whether solely the receipt of a CSR plan triggers increased utilization, there must be a comparison of utilization for similarly situated individuals, which is impractical given the automatic eligibility for and availability of CSRs to ACA market participants in silver plans at these low income levels. The statement that CSR subsidies themselves do not cause induced demand seems unsupported and may actually be a proxy for a socioeconomic statement that lower-income individuals are less likely to seek health care. This is not to say that some adjustment to these factors is inappropriate at some level. We recognize that a failure to adjust CSR-induced utilization factors would retain the subsidization incentives we note above, regardless of the drivers of the utilization differences. However, this would only be viable to the extent that CMS creates a socioeconomic adjustment in the risk adjuster that recognizes the reduced propensity of lower-income individuals to seek care and the related equity issues associated with that statement.

Using the risk adjuster to address plan liability increases due to the lack of CSR funding

Another significant element of the technical paper addresses ways in which the risk adjuster could be used to recognize the higher actuarial value for CSR plans that results from the failure of the federal government to provide funding for CSR subsidies. We reiterate our previous point that any such adjustments are only appropriate as long as CSRs remain unfunded and unpaid.

We note that customizing the induced demand factors to reflect the additional paid claims liability plans may owe for CSR plan variations is in essence a shift from developing a CSR adjustment based on the influence of allowed claims (i.e., induced demand) to developing the factor on a paid claims basis. This is generally appropriate, given the goal of the risk adjuster to predict plan paid claims liability. We note that the actuarial value of CSR plans, like the actuarial value of regular plans, may vary from statutory actuarial values. These variations may additionally be driven by lower relative utilization and the accompanying socioeconomic factors that contribute to it. This creates challenges for developing an adjustment that accurately accounts for any variation in cost relative to the standard silver risk score model. Additionally, CSR plan designs are likely to affect age-gender and HCC factors in different ways, so that a

single multiplicative adjustment to reflect increased plan liability will inevitably overpredict some individuals and underpredict others. Given the influence of the maximum out-of-pocket (MOOP) provisions, it is likely that lower-cost conditions would be somewhat underpredicted and higher-cost conditions would be somewhat overpredicted. In this sense, use of the platinum risk score model may more accurately capture factor-level variations, even as it cannot reflect differences between a 94% and 87% silver CSR variation. Another option may be to combine the two approaches—use a platinum risk score model but make smaller adjustments via the induced demand factor to reflect differences between the CSR variations.

With regards to modifications to the “without risk” component of the risk adjustment transfer formula, we note that there are a variety of factors that influence the relativity of premiums by metal level. Most states currently use some variation of the silver loading practice, which increases silver plan premiums either for all silver plans or for all silver qualified health plans (QHPs) to reflect the additional plan liability associated with CSRs. However, many states either permit multiple approaches, require the additional liability to be spread across all plans, or else do not permit a load of any kind. The approach outlined by CMS of using a “national average CSR mix” may correspondingly not be an appropriate solution for all states, as it aligns best with an “all silver plans” silver load. Furthermore, even in states adopting silver loading, a different mix of CSR variations would be expected due to a variety of factors, including the presence of a basic health plan, Medicaid eligibility levels, and income distribution within the state. CMS might consider separate adjustments for specific states or groups of states depending on CSR loading practices and enrollment levels. This may be complicated by CMS deference to state authorities on CSR load rating practices, as states can and do change their regulatory approach. We are in general alignment with CMS’ statement that an issuer-specific adjustment to actuarial value would be inappropriate, as it implies that an issuer would be able to retrospectively adjust premiums to reflect CSR enrollment. However, we also note that premium variation between metal tiers may be driven by factors beyond any relative risk adjustment-induced profitability of CSR enrollees, so that a state-specific CSR adjustment is unlikely to fully capture variation in premiums between plans.

Thank you for the opportunity to provide input on the *HHS-Operated Risk Adjustment Technical Paper on Possible Model Changes*. We welcome the opportunity to speak with you in more detail and answer any questions you have regarding these comments. If you have any questions or would like to discuss further, please contact Matthew Williams, the Academy’s senior health policy analyst, at williams@actuary.org.

Sincerely,

Jason Karcher, MAAA, FSA
 Chairperson, Risk Sharing Subcommittee
 American Academy of Actuaries

CC: Jeff Wu, Deputy Director for Policy, Center for Consumer Information & Insurance Oversight