



AMERICAN ACADEMY *of* ACTUARIES

---

August 20, 2010

Steven B. Larsen  
Deputy Director, Office of Consumer Information and Insurer Oversight  
Department of Health and Human Services  
200 Independence Avenue, SW  
Washington, DC 20201

Re: Credibility Approaches in PPACA Section 2718

Dear Mr. Larsen:

Thank you again for taking the time to discuss the topic of credibility with members of the American Academy of Actuaries'<sup>1</sup> Medical Loss Ratio Regulation Work Group during our phone call on July 9. We hope we were helpful in addressing the questions you had about credibility in the context of calculating rebates under the minimum loss ratio (MLR) requirements adopted in the Patient Protection and Affordable Care Act (PPACA), which created Sec. 2718 of the Public Health Service Act.

During the call you expressed an interest in our group's views on the then-current draft of Issue Resolution Document IRD044 from the National Association of Insurance Commissioners' (NAIC) Actuarial PPACA Subgroup (subgroup). The IRD044 proposal would create a hierarchy for applying credibility and pooling techniques in the implementation of Sec. 2718 rebates. IRD044 has been under revision since our July 9 phone call and, as of when this letter was being drafted, still had not been adopted by the NAIC. The August 5 version of IRD044 is attached to this letter for your reference.

We can, however, offer some general comments regarding the subgroup's goal of establishing a hierarchy of approaches for a health insurance issuer's MLR to have the greatest possible credibility in a given state for a given line of business (i.e., individual, small group, or large group). The Academy's work group is supportive, in general, of methods that would lead to greater credibility in the calculation of medical loss ratios under Sec. 2718. By "greater credibility," we mean that the MLR would be less subject to wide variations due to random fluctuation.

The most effective way to increase credibility of a block that is not "fully credible" by itself is to aggregate its experience with the experience of other blocks with similar characteristics or of another block that is fully credible by itself. In general, "similar characteristics" would include reasonably similar rating approaches and pricing MLR targets, similar administration, and similar expectations with respect to cost and utilization of health care services.

---

<sup>1</sup> The American Academy of Actuaries is a 17,000-member professional association whose mission is to serve the public on behalf of the U.S. actuarial profession. The Academy assists 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.

The earlier versions of IRD044 incorporated several forms of aggregation that have been removed since our July 9 call:

- Aggregation for an enterprise across that enterprise's legal entities in a given state;
- Aggregation for a legal entity across all states in which it issues policies or across all states for which the experience was combined for rating purposes;
- Aggregation for an enterprise across all states for all legal entities in which the company issues policies.

The intended objective of the current working version of IRD044 is to maximize credibility without combining states or legal entities within a state, instead combining up to three years of experience rather than looking at each year alone.

In our work group's letters to the NAIC subgroup on May 12 and May 20, we provided tables that illustrate how a credibility tolerance adjustment would increase as membership decreases. Such a technical table would be appropriate to use when calculating whether some amount of rebate is due. In fact, such a table and method is already in use for Medicare Supplement refund provisions. This approach, in effect, would assign partial credibility when membership does not reach the full-credibility membership level. This is consistent with long-standing actuarial practice with respect to calculating loss ratios for pricing purposes. Factors can be set in a straightforward and easy-to-apply manner to adjust for both the size of the block and the mix of benefits of the plans within the line of business. We believe such a table would, in comparison to other credibility approaches, be the simplest to administer and the easiest to explain to consumers, and it can be combined with the various methods of aggregation and adjustment discussed throughout this letter.

We have the following comments on the proposed approaches to increase credibility in the Aug. 5 version of IRD044.

*Large-claim pooling*—The concept of large-claim pooling involves all claim dollars in excess of a certain calendar-year amount for an individual being excluded from the numerator of the MLR calculation and replaced with a "pooling charge" that would be sufficient to cover all such claims above the pooling point in the universe of experience under consideration.<sup>2</sup> For example, in exchange for excluding all claims above \$60,000, a pooling charge of \$X per member per month would be added to the numerator in calculating the MLR. The amount \$X would be determined such that \$X times the total membership in the experience universe would equal the total claims expected to be in excess of \$60,000.

Large-claims pooling presents advantages, disadvantages, and new unanswered questions:

- It would reduce random fluctuations for aggregations of experience of all membership sizes, which, in turn, would reduce the credibility adjustments that would be appropriate. With neither full credibility nor the pooling points or charges yet defined, it is unclear the extent to which the use of pooling would lower the full credibility membership threshold.

---

<sup>2</sup> At this time, the suggested "universe" in draft IRD044 would be all legal entities within an enterprise regardless of state or line of business.

- It would create additional administrative complexity, because companies would need to calculate their experience by aggregating each member's claims for the year to determine whether claims for a member exceeded the pooling point.
- In states in which medical costs are higher, a greater percentage of members (covered on a per-service basis) would reach the pooling point than in a state with lower medical costs. Therefore, at least theoretically, the pooling charge should vary by state or possibly by a smaller region within a state (e.g., New York City versus upstate New York).
- There are a variety of other questions that would have to be addressed to implement large claims pooling properly. For example:
  - How would one properly integrate large-claims pooling with any individual XOL stop loss reinsurance and with the PPACA risk-equalization mechanisms?
  - Would the pooling charge be specific to each health insurance issuer, or would it be a single charge for all companies across each market segment? Would there be different pooling charges for each segment or a single charge for all segments?

If desired, our work group would be glad to assist in identifying and proposing answers to these, and other, questions.

It is our understanding that currently the subgroup is proposing that pooling apply to all aggregations of experience within an enterprise. We agree that this would lead to a more equitable approach if the subgroup adopted a credibility table with pooling that also applied to all aggregations of experience within a given enterprise, with the tolerance adjustment determined via that table. This would have the advantage of using an enterprise's entire experience to determine the pooling charge for a given pooling point, although it would increase the administrative burden for companies that achieve full credibility without pooling.

*Combining years of experience*—We agree with the NAIC's contention that combining years of experience would increase credibility. For example, a credibility adjustment for three years of experience with 25,000 members could be considered equivalent to one year's experience with 75,000 members. (The longer the span of time of experience being combined, however, the less likely the pool of experience is homogenous, particularly with respect to population demographics and plan designs, which are more likely to change materially over a three-year span than over a one-year span.) The subgroup suggests limiting to three the number of years for which experience is combined. We agree with this suggestion, given it is consistent with the PPACA statute.

*Other suggestions*—Some of the other rules being suggested by the subgroup may be less equitable. It has been suggested, for example, that if an entity does not have full credibility after three years yet has partial credibility in each of the three years, and if the entity has a MLR less than target in each of the three years, then for rebate purposes the entity should be treated as having full credibility, in which case a rebate would be paid. However, a company that priced for an 80 percent MLR could certainly experience three years of random fluctuation below 80 percent. (Indeed, assuming a normal curve of loss ratios around the 80 percent target, such a result would happen one-eighth of the time.) Such a result would lead to rebate payment (without any tolerance adjustment) based on all three years being close to

but below the MLR requirement. By contrast, if instead a similar company had two years in which the MLR was well below the MLR requirement but one year in which the MLR was just above the requirement, that company would be allowed to include the tolerance adjustment. This difference in treatment between the two companies does not appear to be equitable.

There are other concerns regarding equity that, thus far, are not being addressed adequately.

For example, the loss ratios of very high-deductible policies (e.g., Health Savings Accounts (HSAs)) with deductibles of \$5,000 and higher for families) have significantly higher volatility than policies with a \$1,000 deductible. Credibility factors should therefore vary in some fashion by deductible level, to appropriately accommodate this segment of the individual insurance market. Three or four deductible groupings could be used, as a compromise between technical accuracy and administrative complexity.

MLR volatility of small blocks of business is also driven by uncertainties in setting premium rates (in addition to the statistical fluctuations of claims). In setting premium rates, for example, a small block is similar to a random sample drawn from a population whose parameters are used in setting the price of the sample and is therefore subject to sampling error. Credibility factors derived from an analysis of statistical fluctuation only, such as the tables in our two May letters, would need to be set at a very high confidence level in order to compensate for other sources of MLR volatility.

Finally, in addition to equity concerns, rules proposed to address credibility should be considered in light of their administrative complexity, as well as from a consumer perspective on the degree of complexity as to when a rebate may be payable.

\*\*\*\*\*

We appreciate this opportunity to provide further thoughts on MLR credibility, and we would be willing to expand on these ideas once the subgroup's approach to MLR credibility is more completely fleshed out. If you have any questions, please contact Heather Jerbi, the Academy's senior federal health policy analyst, at [jerbi@actuary.org](mailto:jerbi@actuary.org) or 202.785.7869.

Sincerely,

A handwritten signature in black ink, appearing to read "Rowen B. Bell". The signature is written in a cursive style with a large, looped initial "R".

Rowen B. Bell, FSA, MAAA  
Chair, Medical Loss Ratio Regulation Work Group  
American Academy of Actuaries

Cc: Steven Ostlund, Chair, NAIC Actuarial PPACA Subgroup

This is a DRAFT and is Exposed for Comment – It Does Not Represent the Position of the NAIC.

## A&HWG PPACA Actuarial Subgroup Issue Resolution Document

### IRD044

#### Issue:

Should a hierarchy for application of credibility and pooling techniques be used?

#### Proposed Resolution:

Preliminary resolution 8/5:

A hierarchy for application of credibility and pooling techniques should be used as outlined in the evaluation below.

#### Exceptions:

None identified.

#### Description:

2718(b) references providing an annual rebate of premium in each plan year at the individual, small group and large group levels, but also requires the NAIC to establish standardized methodologies that take into account the special circumstances of smaller plans. If various credibility adjustments are used for this purpose, the form of the adjustments and order of application need to be determined.

#### Documentation in support:

From the July 2, 2010, letter from America's Health Insurance Plans (AHIP):

"We believe that a hierarchy will be needed."

From the July 2, 2010, e-mail from Assurant Vice President Brian Rees:

##### **"IRD044 - Credibility Hierarchy**

I fully agree with the outlined list of objectives or factors to consider with respect to the selection of a hierarchy. If I understand the hierarchy and associated examples correctly, it is the intent that a rebate calculation would be done at the end of each year and if the calculation indicates that the experience MLR less the credibility adjustment is below the MLR Standard, then a rebate would be paid. This would be done for each year separately.

... Allowing the hierarchy to extend to 5 years and allowing for interim rebates would likely add to the problems outlined above. There will be additional administrative expense in tracking policyholders over

the 5 years in order to distribute any rebate. It would be conceivable that a policyholder who was in force in year 1 but lapsed before year 2 could receive interim rebates for all 5 years if the ultimate results called for a rebate even though they only had coverage during year 1, due to the pooling across years. It would be difficult to argue that a policyholder's experience that contributed to the pooled results shouldn't participate in the outcome of those results. This would not appear to be a logical conclusion of PPACA for this to occur.

### **Documentation in opposition:**

From the July 2, 2010, e-mail from Assurant Vice President Brian Rees:

It is my belief that paying a rebate before reaching 100% credibility or reaching the end of the hierarchy steps could result in the pooling concept is being violated. Payment of interim rebates before the full calculation period is completed will not allow for the offsetting nature of the pooling mechanism to work properly and the results will be unequitable to the carriers. I will illustrate via the following example using Company F:

This is a situation where the pooled experience in year 1 is less than 100%. As illustrated, assume that the 2011 experience MLR is 77% and the credibility adjustment is 2%. The hierarchy rules would require a rebate of 1% to be paid ( $80\% - 2\% - 77\%$ ). The problem will arise in years 2+ if the experience MLR is greater than the MLR Standard. Assume the 2012 experience MLR is 83% and the adjustment is 0% (i.e., 2011-2012 is fully credible). The calculated rebate would be 0% ( $80\% - 0\% - 80\%$  with the 2nd 80% being the average of 77% and 83% by combining 2011 and 2012). However, there was already a rebate of 1% paid from 2011 which can't be recovered. In essence, the MLR standard has been increased above 80% (in this example, it is effectively 80.5% due to a rebate of 1% for 1/2 the time period). This is above the level set in PPACA. The pooling across years is negated by the payment of the rebate in earlier year. If the experience MLRs were to be reversed such that 2011 = 83% and 2012 = 77%, the rebate in 2011 would be 0% and the rebate in 2012 would also be 0% ( $80\% - 0\% - 80\%$ , the second 80% being the average of 83% and 77%) - \$0 rebate from 2011. This generates an inequitable result purely due to timing. If pooling were to work properly in this situation, the results would be the same irregardless of the timing of the experience since the pooling is being done for credibility purposes.

This situation could occur in any scenario where multi-year periods are used for pooling and rebates payable during the interim periods. In addition to paying the non-recoverable rebates, the carrier also incurs the administrative expenses of distribution of the rebates in an era where there will be significant expense reduction pressure.

There would be two, possibly three solutions to this inequality:

1. Do not pay any rebates until the full credibility has been reached or the end of the hierarchy has been reached - This would allow for equitable treatment between the policyholders and the carriers;
2. Calculate the rebates at each year but hold them in a reserve to allowed to be used in offsets in subsequent years in the rebate hierarchy calculations - This would allow for offsetting to occur as the concept of pooling is intended since reserves are available while rebates paid out are not recoverable.
3. A third option of not allowing interyear pooling (although this may not allow for achieving higher levels of credible results)

This will also raise the issue of equity for policyholders across the pooled years. The same policyholder in year 1 who lapsed but had favorable albeit statistically driven results in year 1 received a rebate than a policyholder who was in force in year 2 who received no rebate due to the pooled results.”

From the July 6, 2010, letter from AIS Risk Consultants:

“2. Pooling / combining of business should only be performed to the extent necessary to make a block of business fully credible.

For example, if the large claim experience of a block of business is 30% credible, then in determining the MLR that large claim experience should be given 30% weight, and the next level of experience (i.e., pooled large claim experience) should be given 70% weight.

Similarly, if the experience in a state is 80% credible, then in determining the MLR, that state’s experience should be given 80% weight, and the next level of experience (i.e., pooled states experience) should be given 20% weight. A given state that has favorable experience which is not fully credible should not have that experience watered down by other states.

If a block of business for which the experience is not fully credible is completely subsumed into pooled experience, then inappropriate outcomes could result. One such possible scenario is where 2011 and 2012 are both 90% credible, with loss ratios of 83% and 77%, respectively, with a combined loss ratio of 80%. If combined experience is used because 2012 is not completely credible, then with a combined loss ratio of 80% no rebate would be payable. However, if 2012 experience is given its credibility weight of 90%, the credibility weighted loss ratio is 77.6% ( $= 77\% \times 90\% + 83\% \times 10\%$ ), so that a rebate of 2.4% would be payable. The later result seems more consistent with the intent of the legislation.

In summary, the experience for a block of business should not be completely diluted into a large grouping of business if that block of business has some credibility on its own.”

### **Evaluation:**

1. A hierarchy of approaches, detailed below, will be used to address blocks of business that are not fully credible.
2. Approaches considered most desirable or least objectionable will be higher on the list and will be applied first. Desirability of various approaches will be based on the following factors. These are listed in order of importance, although some may have equal importance.
  - a. Rebates should not be required based on experience that is not even partially credible.
  - b. Combining states with varying degrees of rate regulation should be avoided.
  - c. Methodologies that result in excessive administrative burdens should be avoided.
  - d. Inequities among policyholders should be minimized.
  - e. Situations in which there is no possibility of a rebate should be minimized.
  - f. Situations in which rebates are delayed should be minimized.

- g. Inequities among carriers should be minimized.
3. To the extent the block is still not fully credible after applying the first approach on the list, the second approach will be applied and so on.
4. The hierarchy will be as follows:
- a. Pool large claims across all states or across non-credible states (see IRD020 for details).
  - b. If still not fully credible, apply the credibility adjustment (see IRD014 for details) and proceed to step (c).
  - c. If 2012 experience was not fully credible, combine 2011 and 2012 experience for the 2012 calculation. (See IRD061.) (Note: After 2014, when three-year averaging is required across the board, all calculations will be based on three years combined. See IRD069.)
  - d. If still not fully credible after three years, but at least partially credible, and if the medical loss ratio in each of the three years, before applying the credibility adjustment, was below the applicable standard and at least partially credible, then rebates must be paid based on the pooled three-year experience with no credibility adjustment. Rationale: If rates at a level that would result in an MLR equal to the standard but for statistical fluctuations. Then the MLR should meet or exceed the standard 50% of the time. In that situation, the probability that the MLR will be below the standard three years in a row is one in eight. Therefore the presumption is that if the MLR is below the standard three years in a row, it is because the rates reflect a lower MLR, whether intentionally or not. **Open Issue: If this rule results in no credibility adjustment in Year Y, should the rule still apply again in years Y+1 and Y+2 or be suspended until a three new years of experience is accumulated.**
  - e. If still not fully credible after three years, but at least partially credible, and if the medical loss ratio in at least one of the three years, before applying the credibility adjustment, was equal to or greater than the applicable standard or was not even partially credible, then a credibility adjustment will be applied based on the three years and any indicated rebate will be paid.
  - f. If the block is so small that it is still not even partially credible after three years, no rebate is payable.

#### Exceptions References:

#### Attachments:

W:\National Meetings\2010\Summer\fh\lha\ahwg\Rebate IRD\IRD044.doc