SOCIETY OF ACTUARIES

FINANCIAL REPORTING SECTION

VM-20 Game Changer: Mortality Aggregation

By Tim Cardinal

Mortality is a key risk factor receiving considerable attention and effort for companies that have or are about to implement VM-20. The 2020 edition[1] of the Valuation Manual incorporates APF 2018-17 which addresses mortality data aggregation and calculation of company experience rates. It is possible that a company that satisfied requirements in the 2017 through 2019 editions will need to reconsider the 2020 edition requirements.

The most memorable aspect is that VM-20 permits data aggregation to enhance credibility. For example, blocks considered for aggregation might include:

- different underwriting eras, such as thresholds associated with risk characteristics or risk class structures (e.g., 5-class with 6-class);

- different products;

- different distribution channels;

- accelerated underwriting with traditional underwriting; and

- internal data with external data.

In the absence of aggregation, lower credibility means larger margins, which means larger reserves.

The accompanying requirements have perhaps not received as much attention. If the Dire Straits 1980s classic "Money for Nothing" were about VM-20 mortality aggregation rather than about MTV, the lyric would be, "credibility for nothing and your margins for free." However, credibility through data aggregation is not "free." Permission to aggregate hinges on "if similar"—there needs to be similar underwriting processes and similar mortality.

In particular, a current hot topic is aggregating accelerated underwriting with traditional underwriting. Adequately addressing similarity requirements is not free. Does an accelerated underwriting program produce similar outcomes? What is similar? *A/E*s within five percent? Seven percent?

Actuaries are accustomed to establishing and supporting that models, methods and assumptions are reasonable in the context of materiality. Reasonable does not imply similar. Once a company determines "what" is aggregated, a company must demonstrate "why" they can aggregate the "what." I make no attempt to evaluate whether or not it is desirable or permissible to aggregate various blocks of business. This article focuses on similarity requirements but does not advocate specifications.

First, I cite pertinent VM-20 language permitting data aggregation, adding bold for emphasis. I recommend you read the original in its entirety, especially all of 9.C.

Second, I discuss considerations per the necessary "if" condition. These considerations are entirely my opinions. Third, I discuss the requirements for setting aggregate company mortality experience rates per 9.C.2.d.vi.

MORTALITY DATA AGGREGATION: VM-20 LANGUAGE

Previous editions of the valuation manual permitted data aggregation. VM-20 Section 9.A.6.a states: "For risk factors (such as mortality) to which statistical credibility theory may be appropriately applied, the company shall establish anticipated experience assumptions for the risk factor by combining relevant company experience with industry experience data, tables or other applicable data in a manner that is consistent with credibility theory and accepted actuarial practice."

The takeaway is that company rates are not simply the aggregate company experience rates. Rather, company experience is combined and weighted with other data using credibility techniques.

The 2020 edition clarifies and strengthens the conditions to combine data. First, 9.C.2.b defines company experience data as being derived from three sources.

"Company experience data shall be based on experience from the following sources:

i. Actual company experience for books of business within the mortality segment.

ii. Experience from other books of business within the company **with similar underwriting.**

iii. Experience data from other sources, … Data from other sources is appropriate **if the source has underwriting and expected mortality experience characteristics that are similar to policies in the mortality segment.** …"

Data belonging to the mortality segment does not have additional requirements. Data internal to the company requires **similar** underwriting. Data external to the company requires **similar** underwriting and similar expected mortality experience characteristics.

Section 9.C.2.d. reaffirms the permission and adds further to the similarity requirements. It states: "The company may base mortality on the aggregate company experience for a group of mortality segments when determining the company experience mortality rates for each of the individual mortality segments in the group **if the mortality segments were subject to the same or similar underwriting processes. …"**

Sections 9.C.d.i and 9.C.d.ii define underwriting processes as processes by which the company "determines which risks to accept and to which risk class each policy is assigned, including any impacts on these determinations due to distribution systems and target markets." Sections 9.C.d.iii and 9.C.d.iv address a process that is expected (d.iii) or has been shown (d.iv) **"to produce similar mortality"** to that of a previously established underwriting process. Section 9.C.d.iii further states: "… may be treated as similar to the previously established underwriting process if these expectations regarding mortality are supported by relevant, pursuant to Section 9.A.6, third-party proprietary experience studies (such as those of reinsurers or consulting firms. …" Meanwhile 9.C.d.iv requires a retrospective demonstration using statistical analyses, predictive model back testing, or other modeling methods. The common requirement in the two subsections is similar mortality.

SIMILARITY

Consider internal and external blocks of business *A, B, C, D* and so on. There are three separate similarity requirements:

a. Mortality experience characteristics (e.g., demographics, markets, products);
b. underwriting processes; and
c. mortality experience (i.e., outcomes).

How do we demonstrate that the key conditions of "if similar" are satisfied?

We start with some definitions of similar: having traits or characteristics in common, like in form, appearance, size, qualities, relations, etc.; having a resemblance in appearance or nature as to something implied or specified; alike in some respects though not identical.[2]

To a large degree, similarity is subjective and a matter of perspective. Perspective determines which traits to consider. From a 30,000-foot view, a cat is similar to a dog and to a fish—they are all animals. From a 10,000-foot view, a cat is similar to a dog but not to a fish. From a 1,000-foot view, a German shepherd is similar to a Labrador retriever but not to a poodle nor to a Bombay cat. How do we demonstrate our assertions? How do we reconcile that one assertion implies a dog is similar to a cat whereas another assertion implies a dog is not similar to a cat?

Is *A* similar to *B* (i.e., $A{\sim}B$)? To ascertain similarity, a list including subjective (qualitative) and objective (quantitative) criteria can be enumerated and evaluated. Subjective criteria could appeal to intuition or be based on a list of traits that are common to *A* and *B* and traits that are different. We could first develop a list of pre-defined traits—such as size, color, shape, weight, age, mammal, number of legs—and then given any two objects evaluate each trait on the list.

To a degree, similarity is transitive, but there is a slippery slope. We can start with an object A_1 and select five to 10 traits and by making one small change to one trait, derive similar object $A_2{\sim}A_1$. With each successive change, we assert similarity: $A_k{\sim}A_{k-1}$. But through a series of 10+ changes, we end up with an object considered entirely different: $A_1{\sim}A_2, A_2{\sim}A_3, …, A_{99}{\sim}A_{100}$. Is $A_1{\sim}A_{100}$?

We can specify a criterion by defining a "distance" function and stating that if the distance between X and Y is less than *r*, then *X~Y*. In Figure 1, we can visually see that *A~B*, but A and *C* are not similar.

Figure 1
Comparing Objects by Distance
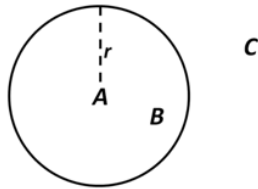
We can make the criteria for dogs, cats and fish objective. In topology, one defines a distance function based on DNA—the differences between sequences of nucleotide bases *A*, *C*, *G* and *T*. Then based on our perspective, we choose an *r* to assert *X~Y*.

> **A Company Wishing to Aggregate Blocks of Business Should Specify What**
> **Similar Means. A Company Can Then Assert *A~B* by Demonstrating the Specified Criteria Are Met.**

Ideally, we could use a perfect existing set of criteria with off-the-shelf distance functions for each of the three similarity requirements. However, each requirement has different challenges in choosing criteria and defining distance functions. Although we can quantify qualitative criteria, underlying subjectivity remains present.

My opinion is that similarity requirements a) and b) have more subjectivity that allows a less stringent, less precise, more qualitative set of criteria, thus providing more wiggle room to demonstrate similarity. Meanwhile requirement c) is more quantitative, resulting in a more precise requirement. I also believe requirements a) and b) will have larger "*r*"—perhaps requirements a), b) and c) could be associated with 10,000-, 20,000- and 1,000-foot perspectives, respectively.

Requirements a) and b) are ex-ante while requirement c) is ex-post. It is necessary, but not sufficient, that ex-ante requirements are similar. No matter how liberal the ex-ante similarity criteria are, there is no escaping the ex-post requirement that outcomes are similar. A company might assert via weak criteria that a) and b) are similar but the real proof is when experience comes through and the outcomes are close.

Section 9.C.2.d (d.iii and d.iv in particular) suggests that the effort to demonstrate similarity should be proportional and reasonable. Qualitative and quantitative criteria for similarity should make sense and be defendable. The bar to clear this might range from little more than asserting "similarity is obvious" to well-thought-out criteria for similarity accompanied by

qualitative and quantitative analyses of underwriting processes and experience such as retrospective analyses. For example, process changes such as 1) adding one threshold, 2) changing the risk class structure by splitting a class (e.g., super-preferred), and 3) introducing accelerated underwriting represent a wide spectrum to assess similarity requirements.

Let's briefly consider each similarity requirement. First, we can form a profile of mortality experience characteristics such as:

*   Sales mix across various attributes such as sex, face amount, risk class, issue age;

*   distribution channels and target markets; and

*   product design features that might affect behavior such as conversions and guarantees.

Sales mix by attribute is readily quantifiable, but the issue is suitable "*r*." For example, is a 60 percent/40 percent male/female mix similar to a 45 percent/55 percent mix or to a 30 percent/70 percent mix?

Second, underwriting processes can be complex and challenging to compare. A process is a series of operations, actions, changes or functions. Underwriting processes are distinct and separate from the result, that is, risk classifications. Underwriting guidelines and practices are more readily compared qualitatively but can be compared quantitatively. Similarity criteria might include the length and nature of questions, use of medical exams, and so forth, for dozens of other underwriting process "traits."

Third, mortality experience can be quantified as A/Es. Provided the Es are based on the same tables, then all that remains is defining a suitable "*r*." For example, is an 80 percent A/E similar to an 85 percent A/E or to a 90 percent A/E? Additional considerations arise when different tables or various sets of adjustment factors are used.

COMPANY EXPERIENCE MORTALITY RATES

VM-20 sections 9.A.6.a and 9.C.2.d.vi require that company experience mortality rates are "informed" by the aggregate experience using credibility or other techniques. Mortality segment rates are a blend and not simply the aggregate rates or individual segment rates. The essential concept underlying methods such as credibility and meta-analysis is that we can better understand, estimate and predict one group by aggregating many groups.

Section 9.C.2.d.vi.a is called the "top-down" approach, and section 9.C.2.d.vi.b is called the "bottom-up" approach. There has been enough confusion that the NAIC published examples of both approaches at *https://www.naic.org/documents/pbr_data_*

mortality_aggregation.xlsx. I do not go into detail here and refer the reader to the spreadsheet for basic examples.

The top-down approach uses predefined expected relativities between mortality segments determined from a reliable and applicable external source to subdivide the aggregate experience into mortality segments. The bottom-up approach adjusts the experience of each mortality segment by credibility weighting the individual mortality segment experience with the aggregate company experience for the group. More complex contexts might use a hybrid-approach that is a combination of steps consisting of top-down and bottom-up approaches.

Once techniques have been applied, section 9.C.2.d.vi requires company experience mortality rates to be increased if necessary to conserve deaths. It states that "the company must ensure that when the mortality segments are weighted together, the total amount of expected claims is not less than the aggregate company experience data for the group."

In general, VM-20 requires additional margins where there is greater uncertainty. Section 9.C.6.d.v makes this explicit in regard to mortality aggregation. It states: "To the extent that, when treating an underwriting process as similar, the judgment of the similarity of expected mortality or the estimate of the expected difference in mortality increases uncertainty in the mortality assumption, the margin applicable to the mortality assumption should be increased pursuant to Section 9.C.6.d."

CONCLUSION

To aggregate one or more blocks, a company must demonstrate three similarity requirements: Mortality experience characteristics, underwriting processes, and mortality experience. While we might each have our opinions, similarity is not only in the eye of the beholder (company); the final verdict is by the regulators. ◼

Tim Cardinal, FSA, CERA, MAAA, MBA, is principal at Cardinalis 1 Consulting LLC in Ohio. He can be reached at *tim.cardinal@cardinalis1.com.*

ENDNOTES

1   National Association of Insurance Commissioners. Valuation Manual Jan. 1, 2020 Edition. August 2019. Accessed online: *https://www.naic.org/documents/pbr_data_val_2020_edition_redline.pdf*

2   Used Wiktionary as starting point for definitions of "similar" and "process." Accessed online: *https://www.wiktionary.org*