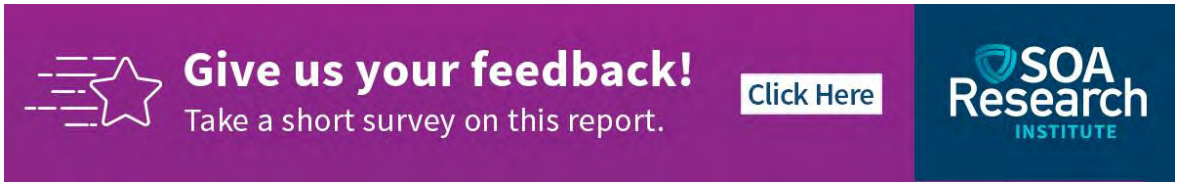
The Impact of Data Limitations On the Statistical Reliability of Mortality Improvement Rate Estimates

DECEMBER | 2024

The Impact of Data Limitations on the Statistical Reliability of Mortality Improvement Rate Estimates

AUTHOR Society of Actuaries Research Institute
Mortality Improvement Reliability Committee



CONTENTS

The Impact of Data Limitations on the Statistical Reliability of Mortality Improvement Rate Estimates

Section 1: Overview

Estimating historical mortality trends—which this report refers to as "mortality improvement rates"—is statistically challenging. The frequently cited rule-of-thumb that 1,082 deaths are needed to achieve a credible mortality rate estimate is not applicable to improvement rate estimation. This rule-of-thumb assumes that a single rate, rather than the ratio of two rates, is being estimated. To produce a statistically meaningful estimate of a mortality improvement rate, a larger pool of data is required. This report examines how the number of persons ("lives") covered in a mortality study, and the length of the study period (in years), affects the reliability of improvement rate estimates. In addition, the report examines practical strategies for addressing data limitations, such as estimating improvement rates by age group instead of by single age, as well as smoothing results across single ages.

This report quantifies mortality improvement using a geometric rate of change. Across two consecutive years of mortality rates, the geometric rate of mortality improvement is equal to 100% minus the year-two mortality rate divided by the year-one mortality rate. For example, given mortality rates of 1.00% and 0.98% in 2016 and 2017 respectively, the mortality improvement rate is 100% minus 0.98% divided by 1.00%, which yields a result of 2.0%. A positive rate of improvement indicates that mortality rates are declining across time, while a negative improvement rate indicates that mortality rates are increasing.

With respect to credibility, a confidence interval for a mortality rate estimate can easily be calculated using basic statistical theory. For example, using simple mathematics, one can demonstrate that given 1,082 observed deaths, and given a small value of "q" (the observed mortality rate), there is a 90% chance that the true value of q falls within 5% of the observed value.

In contrast, the credibility of an estimated improvement rate ("MI rate") is more challenging to determine from first principles, particularly if three or more years of estimated mortality rates are used to produce the MI rate estimate. An exact formula for the credibility of an estimated MI rate is not possible in most situations, and approximations must therefore be introduced. An alternative approach—used to produce the results in this report—is stochastic simulation. As described briefly in Section 3.1 of this report, and in detail in the Appendix, stochastic simulation of mortality rates across many independent trials can facilitate the construction of MI rate confidence intervals.

The advantage of simulation is that it sidesteps the need for complex mathematics (that may involve approximations), providing an easy means to assess the reliability of MI estimates. However, confidence intervals generated from simulation differ from traditional confidential intervals. A traditional confidence interval ("CI") is a concrete mathematical truth that originates from first principles. In contrast, a CI produced via simulation is an estimate that gradually converges with the true value as the number of stochastic trials goes to infinity. For a simulation approach, the observed data alone is insufficient; rather, a simulation is required to assess the statistical reliability of the observed data.

Throughout this report, the term "confidence interval" refers to a distribution of MI errors generated via stochastic simulation, and margin-of-error (MOE) refers to half of the width of the confidence interval.

Key findings of this report include the following:

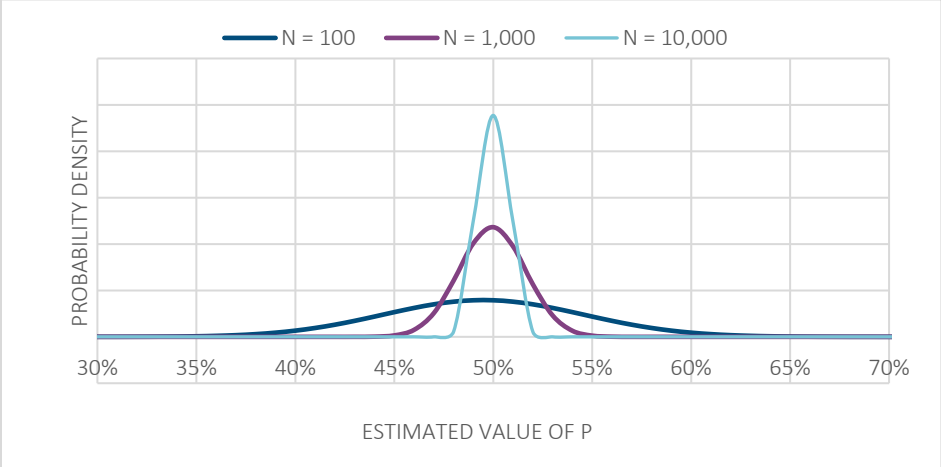
- The MOE for an MI rate estimate is inversely proportional to the number of years in the study, and inversely proportional to the square root of observed deaths (recall that a confidence interval is the estimated value, plus/minus the MOE).
- Given a mortality dataset consisting of only two consecutive years of data and 1,082 observed deaths per year, the resulting MI rate estimate will have an MOE of about 7.0% at 90% statistical confidence (note that the MOE varies slightly as a function of assumed MI rate, but this variation is trivial). For example, if the estimated MI rate is 1.0%, then there is a 90% chance that the true MI value falls between negative 6.0% and positive 8.0%. Clearly this estimate lacks statistical robustness, and would, therefore, be a poor guide for setting a mortality improvement assumption.
- A reliable MI rate estimate—that is, an estimate with a small MOE—requires a large dataset in terms of observed deaths. To achieve an MOE of 0.5% at 90% confidence using five years of recent data would require over 10,000 observed deaths per year. To decrease the MOE to 0.1% (given five years of data) would require over 250,000 observed deaths per year.
- Smoothing or pooling mortality data across ages is a helpful strategy for addressing data limitations, but mortality researchers must be mindful that MI rates are not constant as a function of age. Consequently, there are logical limits to the extent to which smoothing or pooling can be employed.
- Given the challenge of developing reliable MI estimates, practitioners may wish to examine the MI estimates generated by the SOA Research Institute's Mortality Improvement Model, as well as estimates generated by the U.S. Individual Life Mortality Improvement Analysis Tool, each of which make use of datasets that cover tens of millions of lives.

Section 2: The Statistical Reliability of Mortality Rate Estimates

A fair coin flipped N times is not likely to produce an estimate of the probability of heads ("p") equal to exactly 50%. However, the larger the value of N, the greater the chance that the estimate of p—denoted as $\hat{p}$—will be close to its true value. This is illustrated in Figure 1, which presents the probability distribution of $\hat{p}$ for three different values of N:

Figure 1

PROBABILITY DISTRIBUTION OF $\hat{P}$ FOR A FAIR COIN, AS A FUNCTION OF NUMBER OF TOSSES



The greater the number of tosses used to estimate p, the tighter the distribution of potential estimates. The distribution of $\hat{p}$ has a mean equal to the true value of p, and a standard deviation that is inversely proportional to the square root of N. Therefore, increasing N from 100 to 10,000 reduces the standard deviation of the estimator by 90% (because the square root of 10,000 is 10 times larger than the square root of 100).

The same mathematical logic is applicable to the estimation of a mortality rate. Consider N lives of the same age, sex, and initial health status observed across a 12-month period. For analytical purposes, it is convenient to assume that each life has the same mortality risk (just as each coin toss has the same probability of heads), and to assume that the observations are independent. The mortality rate (q) is estimated as the number of observed deaths divided by N. As in the coin toss example, the estimator, $\hat{q}$, has a mean equal to the true value of q, and a standard deviation that is inversely proportional to the square root of N.

Of particular importance is the ratio of the estimator's standard deviation to its expected value. The lower this ratio, the lower will be the potential error of the estimator expressed as a percentage of the true value. The ratio of the estimator's standard deviation to its expected value varies inversely with the square root of the expected number of deaths, as illustrated in Table 1 (notice that quadrupling the number of deaths leads to a halving of the ratio in row five). In addition, Table 1 illustrates how this ratio affects the statistical reliability of the estimator.

The results in Table 1 were calculated using the Excel/VBA workbook that accompanies this report. Using the workbook, the results can be replicated by pressing the button labeled "Load Scenarios Used in the Report." After loading the scenarios, the inputs and outputs associated with Table 1 will appear in columns "F" through "K" of the workbook.

Table 1

RELIABIITY OF MORTALITY RATE ESTIMATES AS A FUNCTION OF DEATH COUNT

		A1	B1	C1	A2	B2	C2
1	"True" Mortality Rate (q)	1.00%	1.00%	1.00%	2.00%	2.00%	2.00%
2	Lives	25,000	100,000	400,000	12,500	50,000	200,000
3	Expected Deaths	250	1,000	4,000	250	1,000	4,000
4	Stdev of Deaths	16	31	63	16	31	63
5	Stdev / Expected Deaths	6.3%	3.1%	1.6%	6.3%	3.1%	1.6%
6	Prob of Error < 1.0 Stdevs	68.3%	68.3%	68.3%	68.3%	68.3%	68.3%
7	Prob of Error < 10% of q	88.8%	99.9%	100.0%	89.0%	99.9%	100.0%
8	Prob of Error < 5% of q	57.3%	88.8%	99.9%	57.5%	89.0%	99.9%
9	Prob of Error < 1% of q	12.6%	24.9%	47.5%	12.7%	25.1%	47.7%
10	Prob of Error < 0.5% of q	6.3%	12.6%	24.9%	6.4%	12.7%	25.1%
11	Prob of Error < 0.1% of q	1.3%	2.5%	5.1%	1.3%	2.5%	5.1%

Rows 6 through 11 refer to the absolute value of the estimator's relative error. For example, row 8 displays the probability that the estimated mortality rate falls between 95% and 105% of the true mortality rate.

Table 1 contains six columns of results. For columns A1, B1, and C1, the true mortality rate is 1%, while for columns A2, B2, and C2, the true mortality rate is 2%. While the "2" columns have double the mortality rate of the corresponding "1" columns, they have half the number of lives. Consequently, A1 and A2 have the same number of expected deaths, and the same is true for B1 and B2, and for C1 and C2. Furthermore, the values in rows five through eleven are nearly identical for columns A1 and A2, B1 and B2, and C1 and C2 (a result that will be explained shortly).

The greater the number of expected deaths (the product of rows one and two of Table 1), the greater the reliability of the estimated mortality rate (assuming that "reliability" is evaluated by examining percentage errors). This relationship arises because, as the number of expected deaths increases, the standard deviation of deaths declines relative to the expected number of deaths (see row five). A smaller relative standard deviation translates into higher probabilities that the estimate will fall close to the true value. For example, as indicated on row nine of the Table, with 250 expected deaths there is only a 12.6% chance that the estimated mortality rate will fall within 1% of the true value. This probability rises to nearly 25% given 1,000 expected deaths, and to almost 50% given 4,000 expected deaths.

Given 1,000 expected deaths (columns B1 and B2), there is an 89% probability (row eight) that the estimate will fall within 5% of its true value. An oft-cited rule-of-thumb is that 1,082 deaths are needed for full credibility. While not shown in the table, 1,082 deaths leads to a 90% confidence of an error of less than 5% (which is the situation that underlies the derivation of this value).

Regardless of the number of expected deaths, the estimator ($\hat{q}$) will always have a non-zero standard deviation. Therefore, in a sense, the "true" value of q is unknowable: it is impossible to arrive at estimates that have no uncertainty. The goal of a mortality study is not to eliminate uncertainty, but rather to develop estimates that are sufficiently reliable or credible for their intended application.

Section 3: The Statistical Reliability of Mortality Improvement Rate Estimates

3.1 A SIMULATION TECHNIQUE FOR CALCULATING MI RATE CONFIDENCE INTERVALS

Calculating an MI rate confidence interval is more complicated than calculating a confidence interval for a point-in-time mortality rate estimate. Working with the ratio of two random variables is complex, often leading to a need for approximations. Rather than derive an approximate formula, this section of the report focuses on the results of simulations. The MI rate confidence intervals presented herein were calculated using an Excel/VBA workbook developed by the SOA Research Institute. The workbook is available for download on this report's web page. An overview of the workbook is provided in the report's Appendix.

Briefly, the workbook has parameters to specify the number of years of data in a mortality study, the exposure in each year, the initial mortality rate, and the MI rate. Using these parameters, a simulation is performed consisting of 100,000 independent stochastic trials. For each trial, a time series of mortality rates is simulated. The simulated mortality rates are independently generated and uncorrelated, each following a normal distribution with a mean equal to the assumed mortality rate (adjusted across time to reflect the assumed MI rate) and a standard deviation that is inversely proportional to the square root of exposure. From the simulated mortality rates, an MI rate is calculated. This rate can be viewed as an estimate, distinct from the assumed MI rate that underpins the simulation. For each trial, the gap between the assumed MI rate and the estimated MI rate is recorded. This gap constitutes an MI rate estimation error. Across 100,000 stochastic trials, the distribution of errors facilitates the construction of MI rate confidence intervals.

An example is helpful for illustrating this concept. Using the Excel/VBA workbook, the following scenario was simulated across 100,000 stochastic trials: two consecutive years of data, each with exposure of one million; an assumed mortality rate of 1% in the first year, and an assumed MI rate of 2%. For each stochastic trial, these parameters are used to simulate two consecutive years of mortality rates, and an MI rate estimate is estimated from the simulated results. The estimated MI rate is equal to 100% minus the year-two simulated mortality rate divided by the year-one simulated mortality rate. For this example, across 100,000 stochastic trials, 90% of the estimated MI rates fall in the range of 0.4% to 3.6%. Relative to the true MI rate (assumed to be 2%, in this example), this interval runs from negative 1.6% to positive 1.6%. Thus, the margin-of-error (MOE) at 90% confidence is 1.6%. This MOE can serve as an uncertainty estimate for a real-world analysis that shares the same data characteristics (two years of data, exposure of one million, initial observed mortality rate of 1%, and an estimated MI rate of 2%).

3.2 RELIABILITY OF MI RATE ESTIMATES GIVEN TWO CONSECUTIVE YEARS OF MORTALITY DATA

To estimate the rate at which mortality rates are changing across time, mortality estimates for different points in time must be compared against each other. The greater the number of years of data available, the greater the range of options and techniques for estimating improvement rates. As a starting point for discussion, it is useful to consider a simple case in which only two consecutive years of data are available. Given two years of data, the rate of annual improvement (MI) will be estimated as follows:

$$\widehat{MI}_{year} = 1 - \hat{q}_{year} / \hat{q}_{year-1}$$

For example, suppose that the estimated mortality rate for 65-year-old females is 1.00% in 2021 and 0.98% in 2022. Given this data, the estimated improvement rate would be 2%. The positive improvement rate indicates that the mortality rate declined across time (while a negative improvement rate would indicate that the mortality rate increased).

The MI confidence intervals presented in Table 2 assume (1) two consecutive years of data, each of which is used to derive a point-in-time mortality rate estimate, (2) the number of observed lives is equal across the two years of data, and (3) the observed lives are assumed to have roughly the same mortality risk (for example, the lives could consist of individuals of the same age, sex, and initial health status). All lives, both within and between years, are assumed to be independent.

Table 2

RELIABIITY OF MORTALITY IMPROVEMENT RATE ESTIMATES AS A FUNCTION OF ANNUAL DEATH COUNT

	Expected Deaths	0.25K	1K	4K	16K	64K	256K	1024K
1	Stdev of MI estimator	8.99%	4.46%	2.22%	1.11%	0.56%	0.28%	0.14%
2	Margin-of-Error @ 90%	14.75%	7.33%	3.66%	1.83%	0.91%	0.46%	0.23%
3	Prob of MI Error < 10%	73.87%	97.48%	100.00%	100.00%	100.00%	100.00%	100.00%
4	Prob of MI Error < 5%	42.59%	73.88%	97.53%	100.00%	100.00%	100.00%	100.00%
5	Prob of MI Error < 1%	8.95%	17.78%	34.69%	63.13%	92.79%	99.97%	100.00%
6	Prob of MI Error < 0.5%	4.48%	8.95%	17.78%	34.69%	63.13%	92.79%	99.97%
7	Prob of MI Error < 0.1%	0.91%	1.79%	3.59%	7.16%	14.27%	28.08%	52.80%

The results in this table assume two consecutive years of mortality data, each of which is used to estimate a mortality rate. An improvement rate is obtained by comparing the two mortality rates. The MOE is half of the corresponding confidence interval. An MOE of 1% at 90% confidence means that there is a 90% probability that the true MI rate falls within 1% of the estimated rate; therefore, if the estimated MI rate is 2%, then 90% confidence interval would run from 1% to 3%.

The results in Table 2 were generated using the Excel/VBA workbook that accompanies this report. To examine the inputs and outputs associated with Table 2, press the button labeled “Load Scenarios Used in the Report.” After pressing the button, the inputs and outputs for Table 2 will appear in columns “M” through “S.” Because results are stochastically generated using sequences of random numbers, and because these sequences change from one run to the next, simulation results will differ across runs. These differences, however, will be quite small because the workbook uses a minimum of 100,000 stochastic trials to analyze each scenario. The large number of trials ensures that the distribution of simulated results is statistically robust.

The results in Table 2 were generated using an assumed MI rate of 2%. A change to the assumed MI rate leads to shifts in results, but the shifts are quite small. For example, in the “4K” column, changing the assumed MI rate from 2% to 0% causes the MOE (at 90% confidence) to drop from 3.66% to 3.61%.

Row one of Table 2 reveals that the standard deviation of the MI rate estimator is roughly halved if the number of deaths is quadrupled. Or, to put it differently, the standard deviation of the MI estimator is proportional to the inverse of the square root of deaths. This is the same relationship observed in Table 1 with respect to the estimator for a mortality rate. However, for an MI rate estimator, this relationship is approximate rather than exact.

To better understand the results in Table 2, it is useful to consider a numerical example. Suppose that a two-year mortality study has an exposure of 100,000 lives in each year. Suppose that 1,010 deaths occur in the first year, yielding a mortality rate estimate of 1.01%, and 990 deaths in the second year, yielding a mortality rate estimate of 0.99%. The estimated rate of improvement is 2% (100% minus 0.99% / 1.01%).

Because the number of deaths in each of the two years is close to 1,000, column “1K” in Table 2 can be used to assess the reliability of the MI estimate of 2%. According to row one of column 1k, the standard deviation of the MI rate estimator is about 4.5%. Row two indicates that there is a 90% chance that the true MI rate lies within 7.3% of the estimated value of 2%. Therefore, the 90% confidence interval runs from negative 5.3% to positive 9.3%.

Rows three through seven in Table 2 provide additional insight into the reliability of an MI rate estimate. Applying the information in these rows to an MI rate estimate of 2% leads to the following confidence intervals:

- 1.8% chance that the true MI rate falls between 1.9% and 2.1%.
- 8.9% chance that the true MI rate falls between 1.5% and 2.5%.
- 17.8% chance that the true MI rate falls between 1.0% and 3.0%.
- 73.9% chance that the true MI rate falls between –3.0% and 7.0%.
- 97.5% chance that the true MI rate falls between –8.0% and 12.0%.

Suppose that an insurer wishes to develop an MI rate that has a 90% chance of falling within 0.5% of the true level. Given this objective, 1,000 deaths are clearly insufficient because it provides merely an 8.9% chance of the estimated MI error being less than 0.5%. While not specifically shown in Table 2, the required number of deaths (per year) is greater than 200,000. This is a sobering result given that there are only about three million deaths in total each year in the United States.

3.3 OPTIONS FOR ESTIMATING MI RATES GIVEN THREE OR MORE YEARS OF DATA

The MI reliability metrics presented in Table 2 are based on merely two years of consecutive data. Given a longer study period, tighter confidence intervals can be achieved. The longer the study period, the greater the range of options and techniques for estimating MI rates. This report considers two options:

1. Using the first and last year of data from an N-year study period to calculate an MI rate, while discarding the data from all other years. For example, if data is available from 2013 to 2018, MI rates would be calculated as follows: MI rate = $1 - (\hat{q}_{2018} / \hat{q}_{2013})^{(1/5)}$

2. Using all N-years of data as follows: take the natural log of each annual mortality rate; fit a regression to the time series of logged mortality rates; extract the slope of the regression line; the estimated MI rate is equal to the negative of the slope.

Option one is explored in Section 3.4 and option two is explored in Section 3.5. Option one is easy to implement, and the first and last years of a mortality study are usually the most important with respect to understanding a mortality trend. However, option two leads to tighter MI confidence intervals because it utilizes all available data.

3.4 ESTIMATION RELIABLITY USING FIRST AND LAST YEARS OF AN N-YEAR STUDY PERIOD

Given N years of data, a simple MI estimation approach is to discard all data except for the first and last years, and to calculate the MI rate as follows:

$$\text{Estimated MI Rate} = 100\% - (\hat{q}_{\text{last year}} / \hat{q}_{\text{first year}})^{(1/N)} \quad [N = \text{last year} - \text{first year}]$$

Using this approach leads to the results shown in Tables 3 and 4. In each table, the “time interval” refers to the number of years separating two points in time. If the interval is one, then two consecutive years of data are used (for example, 2018 and 2019). If the interval is four years, then the two years of data are separated by a gap of four years (for example, 2015 and 2019).

The results in Tables 3 and 4 were calculated using the Excel/VBA workbook that accompanies this report. To examine the inputs and outputs associated with tables, press the button labeled "Load Scenarios Used in the Report." After pressing this button, the inputs and outputs associated with column "64K" of Tables 3 and 4 will appear in columns "U" through "Y" of the workbook. If you wish to produce the results for other columns of these tables, simply modify the exposure assumption in columns U through Y accordingly. For example, to replicate the results in column "16K," the exposure levels used for the "64K" column must be reduced by 75%.

Table 3

MARGIN-OF-ERROR AT 90% CONFIDENCE FOR MORTALITY IMPROVEMENT RATE ESTIMATES

Time Interval (Years)	Deaths Per Year						
	0.25K	1K	4K	16K	64K	256K	1024K
1	14.755%	7.333%	3.661%	1.830%	0.915%	0.457%	0.229%
2	7.357%	3.664%	1.830%	0.915%	0.457%	0.229%	0.114%
4	3.676%	1.832%	0.915%	0.457%	0.229%	0.114%	0.057%
8	1.838%	0.916%	0.457%	0.229%	0.114%	0.057%	0.029%
16	0.919%	0.458%	0.229%	0.114%	0.057%	0.029%	0.014%

The results in this table are based on two years of mortality rate data, separated by the specified time interval.

Table 4

PROBABILILITY THAT MORTALITY IMPROVEMENT RATE ESTIMATE IS WITHIN 0.5% OF TRUE VALUE

Time Interval (Years)	Deaths Per Year						
	0.25K	1K	4K	16K	64K	256K	1024K
1	4.48%	8.95%	17.78%	34.69%	63.13%	92.79%	99.97%
2	8.95%	17.78%	34.69%	63.13%	92.79%	99.97%	100.00%
4	17.78%	34.69%	63.13%	92.78%	99.97%	100.00%	100.00%
8	34.68%	63.12%	92.78%	99.97%	100.00%	100.00%	100.00%
16	63.07%	92.75%	99.97%	100.00%	100.00%	100.00%	100.00%

The results in this table are based on two years of mortality rate data, separated by the specified time interval.

The results in Table 3 reveal that the MOE is inversely proportional to the square root of the length of the period analyzed. Keep in mind that only the first and last years of the data are used to compute the MI rate. A doubling of the time interval leads to a halving of the MOE. In addition, as discussed earlier in this report, the MOE is inversely proportional to the square root of deaths—that is, quadrupling the number of deaths leads to approximately a 50% reduction of the MOE. These relationships are summarized in the following equation, with "K" a parameter that varies by level-of-statistical confidence:

$$\text{Margin-of-Error for MI Estimate} = K / [\text{Years} * \text{Deaths}^{0.5}]$$

A caveat is that the greater the length of the period (in years), the less relevant the MI estimate is with respect to recent mortality improvement. A lengthy period will produce an MI estimate that summarizes long-run rather than short-run experience. If one's goal is to estimate *recent* mortality improvement rates, then the analyzed period must be relatively short if the estimation technique assumes a constant MI rate.

Alternatively, a more complex estimation approach could be used that produces MI estimates that change across time. Such an approach is used to develop the MP projection scale. However, this report focuses solely on simple approaches that assume MI rates that do not vary across time.

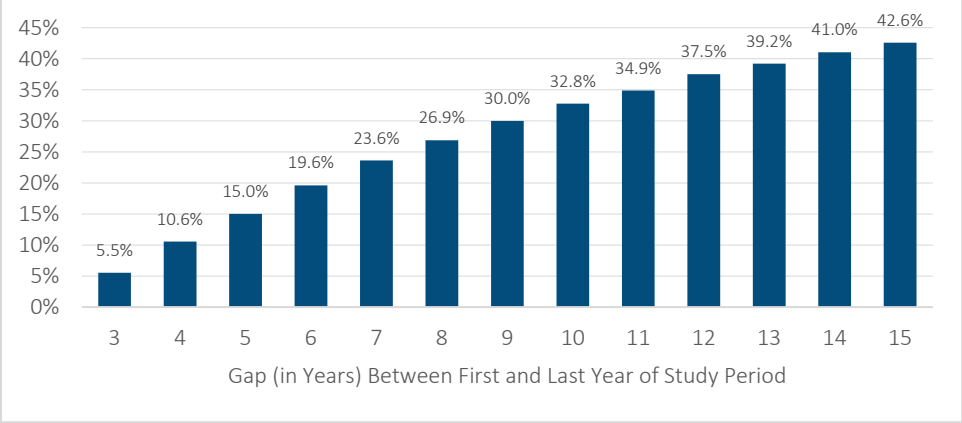
3.5 ESTIMATION RELIABILITY USING ALL YEARS OF DATA FROM AN N-YEAR STUDY PERIOD

In Section 3.4, MI rate estimates and confidence intervals were calculated using the first and last years of data from N-year periods. Alternatively, instead of discarding the data between a period's end points, all years of data may be used as follows: (1) take the natural log of each annual mortality rate; (2) fit a regression to the time series of logged mortality rates; (3) extract the slope of the regression line; (4) and compute the estimated MI rate as 100% minus "e" raised to the result of step four ("e" is "Euler's number, approximately equal to 2.718). For example, if step three produces a value of negative 2%, then the MI rate is equal to 100% minus $e^{-2\%}$, which yields a result of 1.98%.

As illustrated in Figure 2, this approach leads to a reduction in the standard deviation of the MI estimator, relative to an approach in which only the first and last years of data are used for an N-year period. The lower the standard deviation of the MI estimator, the greater the reliability of MI estimates.

Figure 2

PERCENT REDUCTION IN STANDARD DEVIATION OF MORTALITY IMPROVEMENT RATE ESTIMATOR IF ALL YEARS OF AN N-YEAR PERIOD ARE USED RATHER THAN ONLY THE FIRST AND FINAL YEARS



The results in this table were determined via stochastic simulation. For 100,000 stochastic trials, synthetic mortality data was generated under the assumption the "true" rate of mortality improvement is constant. For each stochastic scenario, the MI rate was estimated using (1) only the first and last years of synthetic data and (2) all years of data.

The results in Figure 2 were calculated using the Excel/VBA workbook that accompanies this report. To examine the inputs and outputs associated with tables, press the "Load Scenarios" button. After pressing this button, the inputs and outputs associated with the 10-year column of Figure 2 will appear in columns "AA" and "AB" of the workbook. To produce the other results shown in Figure 2, adjust the parameter that specifies the gap between the first and last years of the simulated data.

Suppose that data is available from 2014 to 2019, so the distance between the first and last time point is five years. For a five-year time interval, using all years of data leads to a 15% reduction in the MI estimator as opposed to simply using the first and last years of data. This translates into a 15% reduction in the width of MI rate confidence intervals.

3.6 PRACTICAL CONSIDERATIONS

There are several factors to consider when using multiple years of experience data to estimate mortality improvement.

One consideration is the homogeneity of the data. Across time, the demographic makeup of a dataset may change. For example, the average face amount of the policies may have increased substantially, or underwriting practices may have changed. Because of these changes, the observed mortality improvement may include a non-biometric element. Therefore, the practitioner should consider whether it is advisable to normalize the data if the objective is to estimate mortality improvement resulting from biometric causes only.

Similarly, if ages are pooled into groups, the average age of a group may change across time. For example, if ages 60 to 69 are pooled together, the group's average age may shift from 64.0 to 64.5 across a study period. An upward shift in the average age will tend to push the pooled mortality rate upwards, while a downward shift will have the opposite effect. This type of change should not be mistaken as mortality improvement or deterioration, because it is unrelated to shifts in age-specific mortality rates. Therefore, pooled mortality rates should be normalized so that they reflect a constant age distribution across time (a practice referred to as "age standardization").

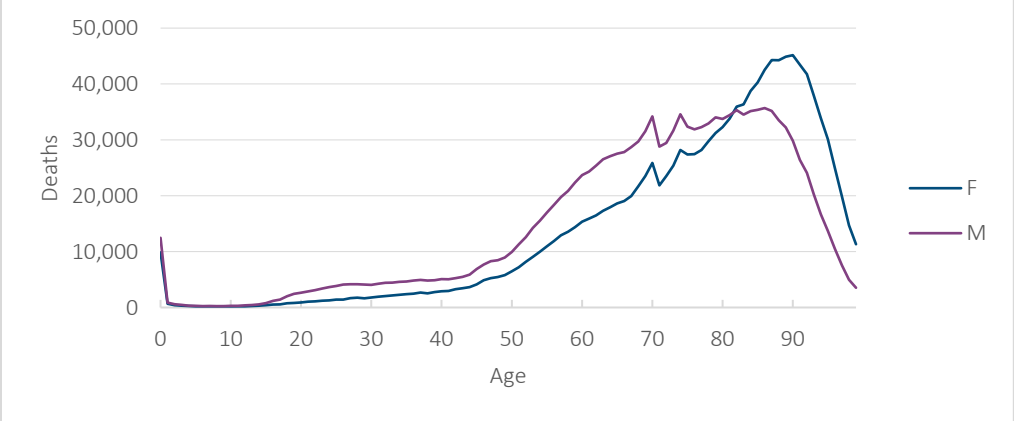
A final consideration concerns the fact that the first and last years of mortality experience are of primary importance in estimating mortality improvement. The mortality rates for these years should be analyzed to see if they are anomalous vis-a-vis the other years of data. It may be necessary to exclude or adjust the experience of these years.

Section 4: An Illustrative Example Using National-Level Data

Given that robust MI rate estimates demand a large dataset, it is worthwhile to assess what level of reliability can be achieved using a dataset that captures all deaths in the United States (U.S.). To this end, Figure 3 shows total deaths in the U.S. in 2017, separately by age and sex, using data from the Centers for Disease Control and Prevention (CDC).

Figure 3

TOTAL DEATHS IN THE UNITED STATES IN 2017, BY SEX AND AGE



Source: the data used in this graph was downloaded from Centers for Disease Control and Prevention.

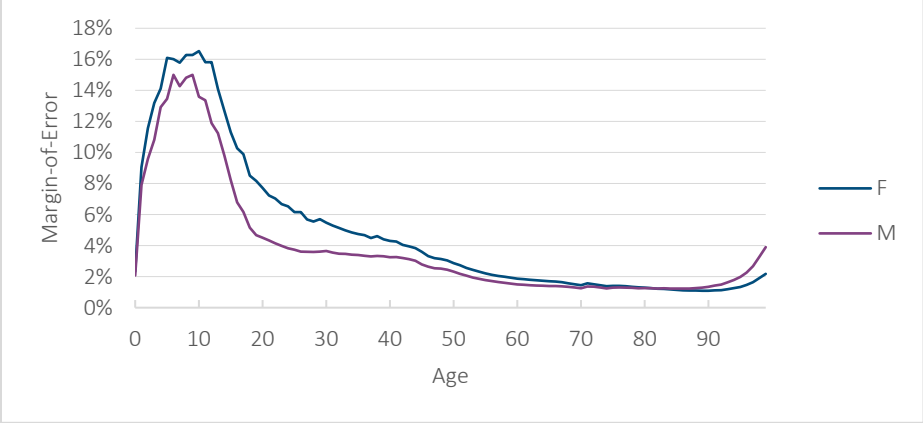
Figures 4A and 4B translate the death counts in Figure 3 into MI rate margins-of-error at a 90% confidence level. The calculations assume that the study period consists of merely two years (e.g., 2016 and 2017), and

assumes that MI rates are calculated by single age, without pooling or smoothing across ages. Figure 4A runs from age 0 to 99, while Figure 4B presents the same data, but focuses on ages 40 to 99.

The results in Figures 4A and 4B reveal that using a short study period of merely two years and performing MI rate calculations separately by single age (without smoothing), leads to statistically unreliable MI estimates. Even from ages 75 to 90 (where most deaths occur), the MOEs exceed 1%. If, for example, the estimated improvement rate at age 80 is 2%, and the MOE is 1%, then there is a 90% probability that the true MI rate falls between 1% and 3%. This is a broad range, indicative of substantial uncertainty.

Figure 4A

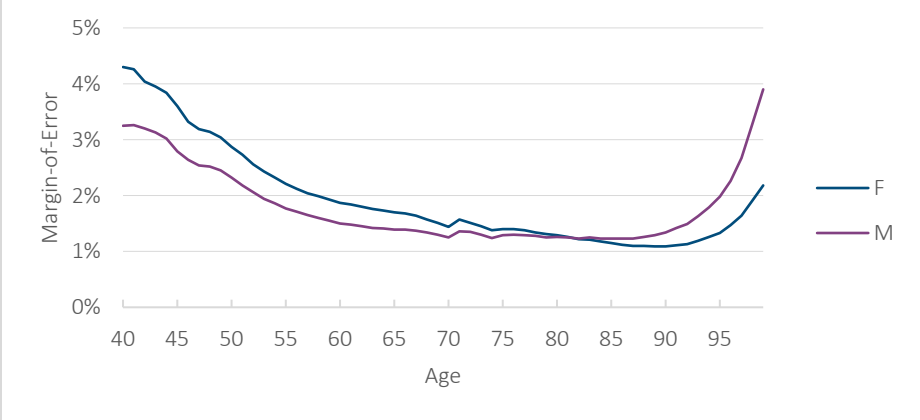
MARGIN-OF-ERROR AT 90% CONFIDENCE FOR MORTALITY IMPROVEMENT RATE ESTIMATES BASED ON TWO CONSECUTIVE YEARS OF NATIONAL-LEVEL DATA



These results assume that two consecutive years of national-level U.S. population data are used to produce two sets of mortality rates. The rates are compared to produce MI estimates by single age, without any smoothing.

Figure 4B

MARGIN-OF-ERROR AT 90% CONFIDENCE FOR MORTALITY IMPROVEMENT RATE ESTIMATES BASED ON TWO CONSECUTIVE YEARS OF NATIONAL-LEVEL DATA



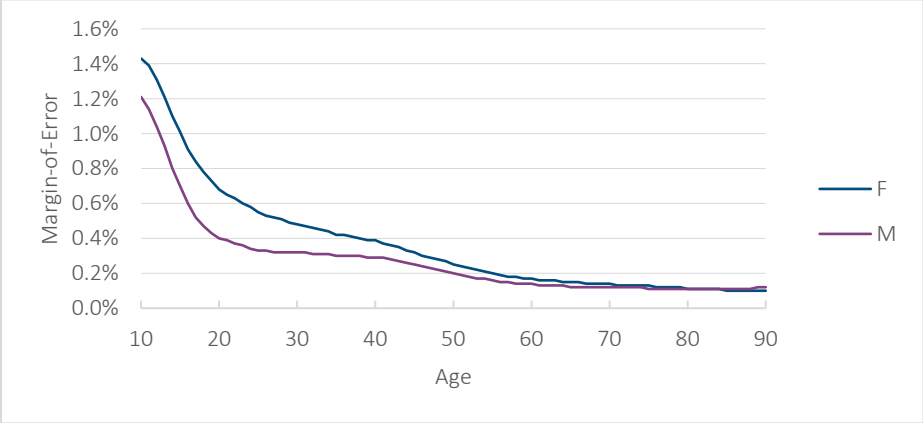
The results in Figure 4B are identical to those in Figure 4A but restricted to a narrower age range.

To reduce the level of uncertainty, two options are available: pooling or smoothing mortality data by age, and/or increasing the length of the study period. Care must be taken not to abuse these techniques by extending them beyond a reasonable limit. For example, one could aggressively assume that improvement rates are constant across all ages, but such an assumption runs contrary to empirical evidence that shows MI differences by age. With respect to the number of years including in an MI study, the longer the period, the less relevant will be the MI rate estimates with respect to quantifying recent, short-term trends.

The results in Figures 5A and 5B are based on the following MI estimation approach: (1) data from 2014 and 2019 is used, providing a time interval of five years across which to estimate improvement (only the first and last years of this period are used); and (2) mortality rates for each age "x" are calculated by pooling data from "x - 2" to "x + 2". For example, the mortality rate for age 80 is determined by pooling data from age 78 through 82. Each age-specific MI rate estimate reflects data from five age cohorts, and data from 2014 and 2019.

Figure 5A

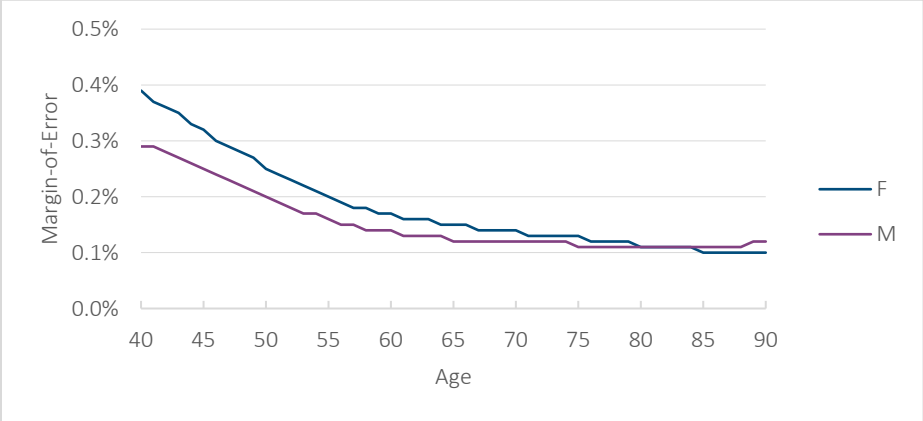
MARGIN-OF-ERROR AT 90% CONFIDENCE FOR MORTALITY IMPROVEMENT RATE ESTIMATES USING DATA FROM A 5-YEAR TIME INTERVAL AND SMOOTHING BY 5-YEAR AGE GROUPS



These results assume that national-level data from 2014 and 2019 is used to produce two sets of mortality rates by single age. For each age "x", the mortality rate is estimated using pooled data from ages "x-2" to "x+2". The resulting mortality rates for 2014 and 2019 are then used to compute MI estimates.

Figure 5B

MARGIN-OF-ERROR AT 90% CONFIDENCE FOR MORTALITY IMPROVEMENT RATE ESTIMATES USING DATA FROM A 5-YEAR TIME INTERVAL AND SMOOTHING BY 5-YEAR AGE GROUPS



The results in Figure 5B are identical to those in Figure 5A but restricted to a narrower age range.

The MOEs in Figures 5A and 5B are over 90% smaller than the corresponding MOEs in Figures 4A and 4B. Recall that MOEs are inversely correlated with the time interval, and inversely correlated with the square root of deaths. Figures 5A/5B use a time interval that is five times the length of the interval used for Figures 4A/4B. This reduces the MOEs to 20% of their prior size. In addition, pooling five adjacent ages together increases the number of deaths in each MI calculation by a factor of five, which, considered by itself, reduces the MOEs to 45% of their prior size [45% = 1 / sqrt(5)]. The product of 20% and 45% is 9%; thus,

the collective effect of using the five-year time interval and pooling five adjacent ages together leads to over a 90% reduction in MOEs).

For ages 70 to 90, the MOEs are around 0.1%. If the MI estimate is 2.0% at age 80, for example, then there is a 90% chance that the true MI falls between 1.9% and 2.1%. This is a narrow range indicative of a high level of statistical reliability.

The results in Figures 4A, 4B, 5A, and 5B assume that the death and exposure data that feeds into MI calculations is accurate. Data that is estimated—for example, national-level exposure data from the U.S. Census—is subject to some uncertainty. This adds an additional layer of uncertainty to MI estimates, but it may be difficult to quantify its effects on MI reliability.

Appendix: An Overview of the SOA's Excel Tool for Analyzing MI Reliability

To produce the results presented in this report, the SOA Research Institute developed an Excel/VBA workbook. The workbook is available for download from the SOA's website. The workbook uses stochastic simulation to develop confidence intervals for mortality improvement rate estimates. 100,000 stochastic trials is sufficient to produce robust results, but you may select more trials if you wish.

To run a simulation, a user must first specify (1) the number of years of data to simulate, (2) the number of lives ("exposure") to be simulated in each year, (3) the "true" mortality rate in the first simulation year, and (4) the "true" rate of mortality improvement (which is assumed to be constant across the simulation period). The simulation will reflect the "true" values, but the randomness of individual mortality risks causes simulation results to deviate from the assumptions. The gap between the "true" MI rate and an MI rate estimated from simulated mortality data can be viewed as an estimation error. Using many stochastic trials, the workbook creates distributions of errors, which, in turn, provide confidence intervals for MI estimates.

Table 5 presents an example of the calculation process for one stochastic trial. In this example, six years of data are simulated using an exposure of 100,000 in each year. The initial "true" mortality rate is 1%, and the "true" rate of mortality improvement is 2% per year.

Table 5

MARGIN-OF-ERROR AT 90% CONFIDENCE FOR MORTALITY IMPROVEMENT RATE ESTIMATES

1	2	3	4	5	6	7	8
Year	True Mortality Rate	Expected Deaths	Stdev of Deaths	Random Number (0 to 1)	Simulated Deaths	Mort Rate Estimate	Logged Mort Rate Estimate
1	1.000%	1000	31.5	0.1009	959.8	0.960%	-4.646
2	0.980%	980	31.2	0.3118	964.7	0.965%	-4.641
3	0.960%	960	30.8	0.1352	926.4	0.926%	-4.682
4	0.941%	941	30.5	0.4539	937.7	0.938%	-4.670
5	0.922%	922	30.2	0.0574	874.7	0.875%	-4.739
6	0.904%	904	29.9	0.2022	879.0	0.879%	-4.734

A linear regression is performed using the logged results in column eight. The slope parameter produced by the regression, multiplied by negative one, is equal to the estimated MI rate. In this example, the estimated MI rate is 2.063%, which exceeds the "true" rate by 0.063%. Thus, for this stochastic trial, the MI estimation error is 0.063%.

Column two shows the expected progression of mortality rates given an initial mortality rate of 1% and an assumed rate-of-improvement of 2%. In each successive year, the assumed mortality rate declines by 2% relative to the prior year's rate. The expected deaths in column three are produced by multiplying the

mortality rates in column two by the assumed exposure of 100,000. If each individual life is considered an independent mortality risk, then actual deaths (as opposed to expected deaths) will follow a binomial distribution with a mean equal to column three, and a standard deviation equal to the value shown in column four (note that column four is approximately equal to the square root of column three).

If exposure is large, the binomial distribution converges to the normal distribution; therefore, the simulation workbook assumes that the number of deaths follows a normal distribution. For each simulation year, the workbook generates a random number between zero and one (shown in column five), and then feeds this value into an inverse cumulative normal function with a mean and standard deviation equal to the values shown in columns three and four, respectively. This produces a simulated number of deaths, shown in column six. Simulated deaths are divided by exposure, producing the mortality rate estimates shown in column seven. These values are logged, producing the results in column eight.

A linear regression is performed using the logged results in column eight. The slope parameter produced by the regression, multiplied by negative one, is equal to the estimated MI rate. In the example presented in Table 5, the estimated MI rate is 2.063%, which exceeds the "true" rate by 0.063%. Thus, for this stochastic trial, the MI estimation error is 0.063%. This calculation process is repeated across many stochastic trials, resulting in a distribution of MI estimation errors. The distribution, in turn, can be used to develop MI confidence intervals.

It is worthwhile to explain why the slope of a time series of logged mortality rates is equal to the MI rate (multiplied by negative one). If mortality rates are changing across time at a constant geometric rate, then their progression from one year to the next is as follows:

1. $q_{year+1} = q_{year} * C$, where $C = 1 - \text{MI rate}$
2. $\ln(q_{year+1} / q_{year}) = \ln(C)$ (take the natural log of both sides of the equation)
3. $\ln(q_{year+1}) - \ln(q_{year}) = \ln(C)$
4. $\ln(q_{year+1}) = \ln(q_{year}) + \ln(C)$
5. $\ln(q_{year+T}) = \ln(q_{year}) + T * \ln(C)$

Equation 5 reveals that if the geometric rate of mortality improvement is constant across time, then the logged mortality rates will change at a linear rate across time. Therefore, a linear regression can be applied to the time series of logged mortality rates, and the resulting slope term (multiplied by negative one) provides an estimate of the MI rate.

Acknowledgments

A working group was assembled to support the development of this report. The SOA Research Institute would like to thank the group members for their guidance, direction, and feedback throughout the project.

Group members:

Mark Rosa, ASA
Larry Stern, FSA
Haofeng Yu, FSA, PhD, MAAA, CERA

At the Society of Actuaries Research Institute:

Philip Adams, FSA, CERA, MAAA, Senior Experience Studies Actuary
Stuart Klugman, FSA, CERA, PhD, Senior Staff Fellow
Ronora Stryker, ASA, MAAA, Senior Practice Research Actuary
Patrick Wiese, ASA, Lead Modeling Researcher

About the Society of Actuaries Research Institute

Serving as the research arm of the Society of Actuaries (SOA), the SOA Research Institute provides objective, data-driven research bringing together tried and true practices and future-focused approaches to address societal challenges and your business needs. The Institute provides trusted knowledge, extensive experience and new technologies to help effectively identify, predict and manage risks.

Representing the thousands of actuaries who help conduct critical research, the SOA Research Institute provides clarity and solutions on risks and societal challenges. The Institute connects actuaries, academics, employers, the insurance industry, regulators, research partners, foundations and research institutions, sponsors and non-governmental organizations, building an effective network which provides support, knowledge and expertise regarding the management of risk to benefit the industry and the public.

Managed by experienced actuaries and research experts from a broad range of industries, the SOA Research Institute creates, funds, develops and distributes research to elevate actuaries as leaders in measuring and managing risk. These efforts include studies, essay collections, webcasts, research papers, survey reports, and original research on topics impacting society.

Harnessing its peer-reviewed research, leading-edge technologies, new data tools and innovative practices, the Institute seeks to understand the underlying causes of risk and the possible outcomes. The Institute develops objective research spanning a variety of topics with its strategic research programs: aging and retirement; actuarial innovation and technology; mortality and longevity; diversity, equity and inclusion; health care cost trends; and catastrophe and climate risk. The Institute has a large volume of topical research available, including an expanding collection of international and market-specific research, experience studies, models and timely research.

Society of Actuaries Research Institute
8770 W Bryn Mawr Ave, Suite 1000
Chicago Illinois 60631
www.SOA.org