

VM-22 Field Test: Preliminary Summary of Participant Results

February 5, 2025

Valuation Manual (VM)-22 (A) Subgroup

Academy and EY Collaboration: Aggregating the Field Test Results

- The VM-22 field test results have been independently aggregated, clarified, and aligned by the Academy and EY.
- EY contacted every submitter, gaining valuable insight.
- Today's results, as presented by EY, reflect the collaborative effort and EY's leadership in the final stages of analysis.
- This presentation represents the publicly discussable results.
- Regulator-only briefings can be scheduled, should that be desired.

Disclaimers

- All participant data received is treated confidentially.
- Participating companies noted varying levels of simplification used to produce field test results within the submission timeframe. Examples include using placeholder assumptions/margins, simplified asset portfolios, only running the Stochastic Reserve and not the Standard Projection Amount, and aggregating inconsistently with proposed VM-22 requirements. Best efforts have been made to analyze and aggregate data submitted by participants. The accuracy and reliability of the results are ultimately dependent on the quality of participant submissions.
- To maintain anonymity of participants per Academy standards, data and metrics for categories with fewer than 5 participants will not be shared publicly.

Overview and Status

Field Test Participant Results

The purpose of this presentation is to provide a preliminary summary of the VM-22 field test participant results.

This first presentation of results focuses on reserves, including overall impacts, sensitivities, and SERT results.

Where applicable, model office results are shown for comparison or to supplement the field test participant results.

VM-22 field test key objectives

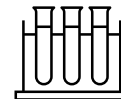


Measure the impact on actual business of the proposed reserve and capital frameworks relative to the current standards to ensure frameworks are working as intended.



Ensure pillars of framework are met

- Appropriate Reflection of Risk
- Comprehensive
- Consistency Across Products
- Practicality and Appropriateness

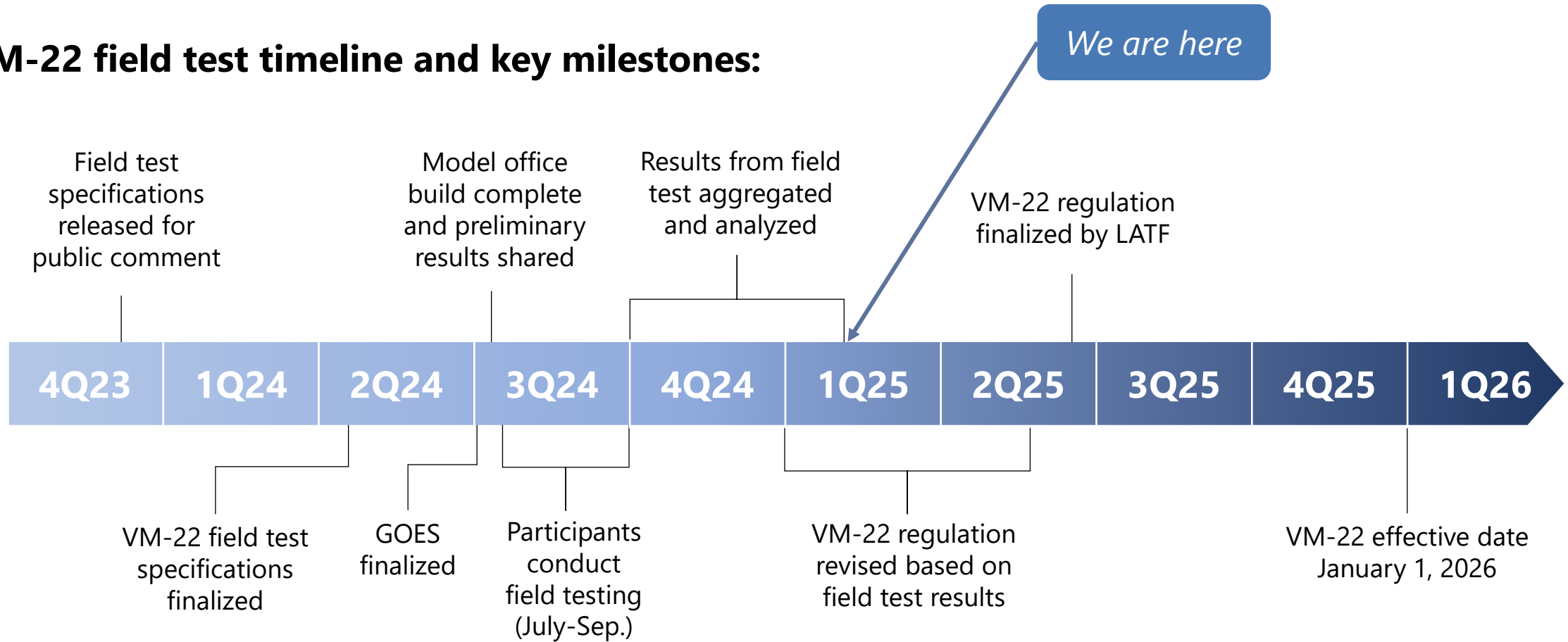


Test the impact of key open VM-22 design decisions

- Aggregation
- Reinvestment guardrail mix
- Stochastic Exclusion Ratio Test threshold
- Standard Projection Amount (SPA) assumptions

Timeline

VM-22 field test timeline and key milestones:



Field Test Results

VM-22 Participant Data Submitted

The tables below show the counts of companies which submitted results for different components of the field test. Note that cells shaded in gray represent data sections which failed to reach the 5-count threshold, resulting in limitations to the analysis presented in the following slides to uphold participant anonymity.

Product	Overall results	SPA results	Margin Sensitivities	Reinvestment Sensitivities
SPIA	8	5	2	3
PRT	6	4	1	2
SSC	5	4	1	2
FDA (no WB)	11	6	6	6
FDA (WB)	4	3	1	2
FIA (no WB)	12	7	6	6
FIA (WB)	12	6	5	5

Limitations in Participant Results

The accuracy and reliability of the field test results are ultimately dependent on the quality of participant submissions. There were a wide variety of limitations noted from participants which could result in materially different impacts of VM-22 once fully implemented. Below are some of the common limitations that were observed.

1. Assets

The Field Test is showing that assets are one of the key drivers of VM-22 results. Many participants used a simplified approach to allocate assets for the field test, which could have a significant impact on results in some cases. Before applying VM-22 in the future, we expect that companies will perform more analysis and refine their approach to determine the assets that will be used to back VM-22 business, potentially aligning both the asset types and duration matching to the prospective VM-22 business.

2. Standard Projection Amount

Some companies did not provide SPA results or provided SPA results on a different level of aggregation than the SR and therefore could not be analyzed on a product level. Because of this, the overall VM-22 impact from CARVM could be misestimated for those companies.

For the companies that did provide SPA results, there were some inconsistencies in the application of the prescribed assumptions. These were discussed throughout the field test Q&A process and have since been clarified in the requirements.

3. Assumptions and Margins

Many companies noted using placeholder assumptions and/or margins for the field test, and that they plan to do additional analysis to set PBR prudent estimate assumptions for VM-22.

4. PIMR

There was inconsistent treatment of PIMR across participant results. Some companies explicitly disclosed PIMR, some included it in the final reserve, some did not reflect PIMR at all. The summary of results is based on the final VM-22 reserve that participants provided.

5. Aggregation

There were some inconsistencies in the way companies aggregated results, for example including GLWB payout streams in the payout category rather than the accumulation category.

6. Business Included

The field test specification asked for at least 10 years of inforce. Some companies provided less than 10 years (e.g. if the product hasn't been sold for that long), and some companies provided significantly more than 10 years of inforce.

Introduction to the Overall VM-22 Results Slides

- Splits by product:
 - Payout Category: SPIA, PRT, and SSC
 - Accumulation Category: FDA (no WB), FIA (no WB), FIA (WB)
- Model office results for each product
- Total number of companies providing results
- Change in final VM-22 reserve compared to CARVM
 - Mean
 - Median
 - Standard deviation
 - Range

*Final VM-22 reserve = Stochastic Reserve + ASPA – PIMR (when provided). Some companies did not reflect PIMR in the results provided with the field test.

SPA vs. SR by Product

The tables below shows summary statistics comparing the ratio of SPA (including buffer) to the SR. A positive % indicates that the SPA is greater than the SR, while a negative % indicates that the SPA is less than the SR.

The SR is driving the final reserve more often than the SPA for most products, including those that could not be aggregated.

It is expected that most companies will refine the assumptions and margins used for the field test before adopting VM-22, which could have a significant impact on the results below.

Product Overview

Product	Model office impact
SPIA	-2.3%
FDA (no WB)	-1.0%
FIA (no WB)	1.6%
FIA (WB)	3.4%

Participant results – SPA vs. SR for VM-22

Total # of companies	# SPA =< SR	# SPA > SR	Mean	Median	Standard Deviation	Range
5	5	0	-2.4%	-2.7%	1.1%	3.1%
6	5	1	-0.8%	-1.0%	1.6%	5.0%
7	5	2	-3.1%	-3.0%	6.0%	18.6%
6	3	3	1.3%	0.0%	4.0%	12.6%

Observations

- The SPA is expected to highlight outliers, so it is not surprising to see the SR dominate for most products.
- Where SPA dominates, it is challenging to pinpoint what the driver is, and whether that is due to simplifications for the Field Test, or whether that is a legitimate outcome in the results. For the WB block, it is believed that the choice in lapse assumptions drove the results in the Field Test.

Overall VM-22 Results: Payout Category

The tables below shows summary statistics on the change from CARVM to the final VM-22 reserve* for field test participants, as compared to the model office results shared previously. Participant results have been normalized so there is equal weighting across companies.

Product Overview

Product	Model office impact
SPIA	-3.4%
PRT	-3.5%
SSC	-5.7%

Participant results – CARVM vs. VM-22

Total # of companies	Mean	Median	SD	Range
8	-3.3%	-0.9%	13.8%	44.8%
6	-0.4%	-1.0%	4.7%	13.2%
5	20.9%	9.7%	30.1%	83.1%

Observations

- Model office results show a decrease in VM-22 reserves compared to CARVM, largely driven by work done in the model office to optimize the assets backing the liabilities.
- Wide range of results seen by participants, with some showing an increase in reserves under VM-22.
- From discussions with participants, this is believed to largely be driven by the selection of assets as multiple companies noted they did not spend significant time selecting or optimizing the asset portfolio for the field test.
- PRT saw a tighter range overall, which is believed to be because PRT assets are usually optimized and allocated to specific PRT deals.
- The model office grouped PRT and SSC together, so they are not directly comparable to the participant results.

*Final VM-22 reserve = Stochastic Reserve + ASPA – PIMR (when provided). Some companies did not reflect PIMR in the results provided with the field test.

Overall VM-22 Results: Accumulation Category

The tables below shows summary statistics on the change from CARVM to the final VM-22 reserve* for field test participants, as compared to the model office results shared previously. Participant results have been normalized so there is equal weighting across companies.

Product Overview

Product	Model office impact
FDA (no WB)	0.3%
FIA (no WB)	4.6%
FIA (WB)	-16.7%

Participant results – CARVM vs. VM-22

Total # of companies	Mean	Median	SD	Range
11	2.6%	1.6%	4.6%	17.7%
12	6.3%	3.9%	7.9%	27.9%
12	-4.5%	-5.0%	8.4%	26.5%

Observations

- For FDA and FIA (no WB), most companies saw a modest increase while some saw modest decreases. From discussions with individual companies, the main driver appears to be how much effort participants put into asset optimization for the field test.
- As noted previously, the model office for FIA includes a modeling limitation related to the hedge costs and payoffs.
- Most companies saw a decrease compared to CARVM for FIA (WB). This was expected given the treatment for WB riders under CARVM.
- Some companies with FIA (WB) saw an increase, or more modest decrease. From some discussions with participants this may be explained by modeling simplifications and/or asset optimization.

*Final VM-22 reserve = Stochastic Reserve + ASPA – PIMR (when provided). Some companies did not reflect PIMR in the results provided with the field test.

Overall VM-22 Results: All Products

The tables below shows summary statistics on the change from CARVM to the final VM-22 reserve* for field test participants, as compared to the model office results shared previously. Participant results have been normalized so there is equal weighting across companies.

Product Overview

Product	Model office impact
SPIA	-3.4%
PRT	-3.5%
SSC	-5.7%
FDA (no WB)	0.3%
FIA (no WB)	4.6%
FIA (WB)	-16.7%

Participant results – CARVM vs. VM-22

Total # of companies	Mean	Median	SD	Range
8	-3.3%	-0.9%	13.8%	44.8%
6	-0.4%	-1.0%	4.7%	13.2%
5	20.9%	9.7%	30.1%	83.1%
11	2.6%	1.6%	4.6%	17.7%
12	6.3%	3.9%	7.9%	27.9%
12	-4.5%	-5.0%	8.4%	26.5%

*Final VM-22 reserve = Stochastic Reserve + ASPA – PIMR (when provided). Some companies did not reflect PIMR in the results provided with the field test.

High-level Observations Summary



Range of results: There was a wider range of results than was expected across all products. Every product had at least one company with an increase and one company with a decrease in reserves.



Selection of assets: The assets used in VM-22 modeling are a key driver of results for all products. Given the simplified approaches that many companies took for assets, results could change materially when asset portfolios are refined. Some participants noted that the reinvestment guardrail had a significant impact on results vs. modeling their company reinvestment strategy.



Dominant reserve: Where SPA results were provided, the SR is winning more often than the SPA for payouts and non-WB accumulation products. The SPA is winning more often on WB products. This is likely due to the SPA lapse assumption for WB products. 8 of the total 19 entities that participated in the field test did not provide SPA results.



Notable differences from model office results:

SSC – The model office included SSC as a subset of the PRT block but did not consider SSC as a standalone product so it's not directly comparable to participant results. SSC results also vary depending on the mix of business and inforce duration of the block, which for some participants was much longer than 10 years.

FIA – The model office results included a topside adjustment for the cost of FIA hedges due to a limitation in GGY Axis.

Sensitivity Results Summary

Introduction to the Sensitivities

- The Field Test Specs asked participants to set, and disclose with results, each sensitivity's impact from mortality, policyholder behavior, expenses, hedging, non-guaranteed elements (NGEs), withdrawals, and other assumptions as deemed necessary.
- Participants were also allowed to use some default margins as described in the Specs if they did not want to use their own margins.
- There was only enough information gathered for mortality, lapse rates, expenses, and the reinvestment guardrail; these are discussed on the following slides.
- Similar to the overall results, there are a number of limitations related to sensitivities, e.g., how companies stepped into and isolated each sensitivities impact.

Margin Sensitivities – Mortality

Background

- Field test participants were asked to remove each liability margin individually and provide sensitivity test results.
- The field test specifications included default margins that companies could choose to use in place of their own margins. For mortality, the default margin was +/- 10%.
- 4 out of the 7 companies included in the analysis below used the default margin. For those who used their own company margins, the margins were <10%.

Results and observations

- Many participants did not provide sensitivity results due to lack of time and resources for the field test.
- Mortality margins were more impactful on accumulation products with WB vs. those without WB, but generally not material for accumulation products overall. Results for the payout category could not be shared publicly, but for the companies that provided results they were largely in line with the WB product results.

Product	# of Companies	# of Products	# of Products > 0%	# of Products =< 0%	Mean	Median	Standard Deviation	Range
FA (no WB)	5	5	0	5	-0.03%	-0.01%	0.04%	0.11%
FIA (no WB)	5	5	0	5	-0.15%	-0.02%	0.21%	0.55%
FIA (WB)	5	5	1	4	-1.01%	-1.13%	0.92%	2.41%
FA + FIA (WB)	6	6	1	5	-0.97%	-0.96%	0.85%	2.41%

Margin Sensitivities – Lapse

Background

- Field test participants were asked to remove each liability margin individually and provide sensitivity test results.
- The field test specifications included default margins that companies could choose to use in place of their own margins. The margins provided were +/- 10% on base lapse and +/- 150% on dynamic lapse.
- 3 out of the 7 companies included in the analysis below used the default margin. For those who used their own company margins, 1 out of the 7 used margins >10% and 3 out of the 7 used margins <10%.

Results and observations

- Many participants did not provide sensitivity results due to lack of time and resources for the field test. For those that did provide results, we were able to aggregate the results of a base lapse sensitivity as shown below. Very few companies provided sensitivity testing on dynamic lapses and therefore results could not be aggregated.
- The base lapse margin sensitivity had an immaterial impact for most companies.

Product	# of Companies	# of Products	# of Products > 0%	# of Products =< 0%	Mean	Median	Standard Deviation	Range
FA (no WB)	6	6	0	6	-0.43%	-0.27%	0.49%	1.48%
FIA (no WB)	6	6	1	5	-0.62%	-0.03%	1.32%	3.57%
FIA (WB)	5	5	1	4	-0.64%	-0.05%	1.10%	2.85%
FA + FIA (WB)	6	6	1	5	-0.54%	-0.05%	1.03%	2.85%

Margin Sensitivities – Expenses

Background

- Field test participants were asked to remove each liability margin individually and provide sensitivity test results.
- The field test specifications included default margins that companies could choose to use in place of their own margins. For lapse, the default margin was +/- 5%.
- 3 out of the 5 companies included in the analysis below used the default margin. For those who disclosed their own company margins, the margins were <5%.

Results and observations

- Many participants did not provide sensitivity results due to lack of time and resources for the field test. The results below are aggregated across FA and FIA products without WB. We received limited results for other products that could not be aggregated, however the results were consistent across all products.
- The expense margin sensitivity had an immaterial impact for all participating companies.

Product	# of Companies	# of Products	# of Products > 0%	# of Products =< 0%	Mean	Median	Standard Deviation	Range
FA (no WB)	5	8	1	7	-0.01%	-0.02%	0.01%	0.04%

Reinvestment Guardrail Sensitivity

Background

Field test participants were asked to provide results for two reinvestment guardrail sensitivities:

- Baseline: 50% AA, 50% A
- Required Sensitivity: 5% Treasury, 15% AA, 40% A, 40% BBB
- Optional Sensitivity: 5% Treasury, 15% AA, 80% A

Results and observations

- Many participants did not provide sensitivity results due to lack of time and resources for the field test.
- Overall, the reinvestment guardrail sensitivities did not have a material impact on reserves for most companies. 5 of the 7 companies included in the below analysis had an impact of <1% for all products.
- The results below show the impact of the required sensitivity vs. baseline for products where we had a sufficient number of data points to aggregate results:

Product	# of Companies	# of Products	# of Products > 0%	# of Products =< 0%	Mean	Median	Standard Deviation	Range
FA (no WB)	6	6	2	4	-0.13%	-0.05%	0.18%	0.51%
FIA (no WB)	6	6	1	5	-0.10%	0.00%	0.30%	0.96%
FIA (WB)	5	5	1	4	-0.41%	-0.54%	0.46%	1.17%
FA + FIA (WB)	6	7	2	5	-0.29%	-0.34%	0.46%	1.21%

Stochastic Exclusion Ratio Test

Stochastic Exclusion Ratio Test

Background

- Field test participants were asked to perform the Stochastic Exclusion Ratio Test (SERT) as outlined in the proposed VM-22 requirements.

Results and observations

- Many participants chose not to provide SERT results due to several factors:
 - Lack of resources to produce results in time for the field test
 - Working assumption that their business would not pass the SERT and therefore they do not plan to run it
 - Do not plan to run the SERT because they want to calculate VM-22 stochastic reserves
- Several companies provided partial results but not enough information to calculate the final SERT ratio. If any participating companies have this information available but did not submit it already, please reach out.
- As a result, field test participant SERT results could not be aggregated and shared publicly.
- For the limited data points provided, the participant SERT results were consistent with the model office results.
- Out of the 11 companies that submitted at least partial results, 10 of them used a mortality margin of +/- 5%, while 1 of them opted to use a mortality margin of +/- 10%.
- The model office SERT results (presented previously) are included on the following slide for reference.

Stochastic Exclusion Ratio Test – Model Office

The table below summarizes the model office results of the stochastic exclusion ratio test for each product. The impact of applying a +/- 5% mortality margin did not materially impact the resulting ratio for all products.

Product	95% mortality factor	100% mortality factor	105% mortality factor
SPIA	3.6%	3.3%	3.1%
PRT	3.7%	3.4%	3.2%
FDA (no WB)	1.3%	1.3%	1.3%
FDA (WB)	2.1%	2.2%	2.3%
FIA (no WB)*	5.8%	5.8%	5.8%
FIA (WB)*	33.8%	33.7%	33.6%

* **Important disclaimer for the FIA model office results:** the cost of the FIA hedges is currently accounted for via a spreadsheet topside for each scenario. The model currently incorporates the payoffs of the hedges, but not the costs. We have included the costs via topside, estimated as $option\ budget \times AV / 12$ (since there are annual resets), which are reflected in the results above and throughout this presentation. A system enhancement is in progress from the vendor.

Next Steps for the Field Test

Next Steps for the VM-22 Field Test

- Discussions with Regulators to discuss results, submissions, data gaps, and any additional requests.
- Additional presentations as requested by Regulators or other interested parties.
- Resume the weekly VM-22 Working Group meetings and continue progress on the VM-22 draft.
- Participants to send any additional Field Test data that they have but have not already submitted.

Questions or Comments:

Amanda Barry-Moilanen
Policy Project Manager, Life
barrymoilanen@actuary.org

Steve Jackson
Director of Research
sjackson@actuary.org