

# Internal Model Validation Guidance

Lloyd's guidance on Solvency II Internal Model Validation  
December 2020

## Disclaimer

No responsibility or liability is accepted by the Society of Lloyd's, the Council, or any Committee of Board constituted by the Society of Lloyd's or the Council or any of their respective members, officers, or advisors for any loss occasioned to any person acting or refraining from action as a result of any statement, fact, figure or expression of belief contained in this document or communication.

Lloyd's is aware that this guidance may be helpful to non-Lloyd's firms. Lloyd's accepts no obligation to these third parties for their use and further, no responsibility or liability for any decision or loss arising from this guidance.

The views expressed in the paper are Lloyd's own. Lloyd's provides the material contained in this document for general information purposes only. Lloyd's shall not be liable for any loss which may arise from reliance upon the information provided.

## Contents

<b>Disclaimer</b>	<b>2</b>
<b>1 Introduction</b>	<b>4</b>
1.1 Purpose	4
1.2 Context	4
1.3 Reporting	4
1.4 The Role of the Board	5
<b>2 Qualitative Components of validation</b>	<b>6</b>
2.1 Overview	6
2.2 Independence	6
2.3 Risk profile	6
2.4 Risk coverage	6
2.5 Risk ranking	6
2.6 Validation cycle	7
2.6.1 Definition of Testing Programme and Targeted Validation	7
2.6.2 Tests that Lloyd's expects to be performed each year	7
2.6.3 Target review test type definitions	8
2.6.4 Application of tests	8
2.6.5 Analysis of validation test results	8
2.6.6 Escalation	9
2.6.7 Change Implementation	9
2.7 Reporting	9
2.7.1 Model Change Validation	9
2.8 Timing	9
2.9 Expert judgement	9
2.9.1 Expert judgements used in the internal model	9
2.9.2 Expert judgement used in validation	10
2.10 Limitations	10
<b>3 Validation Testing</b>	<b>11</b>
3.1 Overview	11
3.2 Sensitivity tests	11
3.2.1 Deterministically varying a set of assumptions ("ST-1")	11
3.2.2 Plausible alternative set of assumptions ("ST-2")	11
3.3 Stability tests	12
3.4 Stress and scenario tests	12
3.5 Reverse stress tests (RSTs)	13
3.6 Testing against experience	14
3.7 P&L attribution	15
3.8 Model Drift	16
3.9 Analysis of change	16
<b>4 Specific Areas of Validation</b>	<b>17</b>
4.1 One-year validation	17
4.2 Top-down and bottom-up validation	17
4.3 Dependencies and aggregation	17
4.4 Outwards reinsurance	18
4.5 External Models	19
4.5.1 Model completeness	19
4.6 Events not in data (ENIDs)	19
4.7 Special Purpose Arrangements (SPAs)	20
4.8 Post diversified contribution to capital	20
<b>Appendix A: Reverse stress tests: frequently asked questions</b>	<b>21</b>

# 1 Introduction

## 1.1 Purpose

This document provides guidance on internal model validation under Solvency II. It is intended to be used by managing agents to inform and support their validation processes.

This model validation guidance should be read in conjunction with the Lloyd's Minimum Standards, other Lloyd's communications on validation and relevant EIOPA and PRA documents.

This document builds on the minimum standards to provide more detailed requirements and guidance for managing agents on Lloyd's expectations of validation.

## 1.2 Context

The validation exercise is a requirement of Solvency II. The purpose of the validation exercise is to ensure that the internal model provides a realistic and robust assessment of all the material risks faced by the syndicate at the relevant point in time. It provides **the Board** with independent assurance that the modelling undertaken appropriately allows for the risks to the business and is compliant with Solvency II tests and standards.

The key context for validation is the risk profile of the syndicate. Validation should ensure that the up-to-date risk profile has been considered in the modelling and changes that do not reflect risk profile movements have not been made. Validation is an iterative process of identification of model limitations and the implementation of improvements. The most appropriate validation tools for achieving this task will vary by syndicate, depending on the size and complexity of its risks.

The principles of proportionality and materiality apply to validation as to other Solvency II requirements, where these have been applied to influence the scope of work this should be clearly highlighted in the reporting.

The validation exercise consists of many components which are covered in more detail in the following sections. The key high-level validation requirements include:

- independence in order to ensure objective challenge;
- confirmation that all material risks are covered in the model;
- a risk ranking exercise and a demonstration that the validation effort has been proportional to the materiality of the risk;
- validation tools and tests that are well defined and appropriate for the risks being validated;
- outcomes of tests that are clearly explained and justified, and the path from a "fail" outcome to escalation and model change (if appropriate) should be clearly mapped; and
- clear and appropriate documentation of the validation process and outcomes.

## 1.3 Reporting

A report on the validation should be produced for the Board. This report should focus on providing the Board with sufficient detail for their assurance that the validation exercise has been robust and comprehensive, and highlighting any findings from the exercise. It is expected that further detail underlying the work is also available and is not all included in the Board level report.

Throughout this guidance, the terminology 'validation report' is intended to specify the Board report, while 'validation pack' is the regulatory submission to Lloyd's (including both the validation report and the additional supporting technical documentation, deep dives and past testing which has been relied on, as well as the signposting template).

As part of the validation reporting, agents are required to document how feedback from Lloyd's (and other stakeholders) during the previous validation cycle has been considered as part of the development of the latest validation cycle.

In addition to the planned validation according to the 3-year cycle, agents should include in their validation report a section which outlines validation outcomes to Lloyd's thematic reviews, validator questions, and associated results. Thematic reviews will be set out annually in an additional

communication to this guidance. Validator questions will be specific to the syndicate's experience over the last year, or to the particular nature of the syndicate's business.

#### **1.4 The Role of the Board**

The Board are expected to play a key role in the validation of the Internal Model. In particular, syndicates should demonstrate that the Board uses validation as a means of better understanding the Internal Model, and its strengths and weaknesses. Specifically, Lloyd's expects to see evidence of:

- Board involvement in the design of the validation process;
- How any validation findings are reported to the Board;
- Board involvement in the process of remediating validation issues to resolution.

The Board is expected to be able to challenge the validation process and its results.

## 2 Qualitative Components of validation

### 2.1 Overview

The overall aim of the validation process is to support the following confirmation statements, from an appropriately skilled validator:

- Validation has been conducted in line with validation policy, with sufficient independence;
- the internal model reliably reflects the risk to the syndicate; and
- all material components and risks have been tested.

The validation exercise should have comprehensive associated reporting to ensure a record of the testing and action taking and support the conclusions drawn.

### 2.2 Independence

Validation requires robust, objective challenge. Agents should demonstrate that the individuals responsible for validation have sufficient skills and independence from the design, build, parameterisation and implementation of the model component being validated. Independence should extend to reporting lines and governance that facilitate the escalation of validation findings without any material conflict of interest.

Validation may make use of tests carried out by those responsible for the modelling, but not rely entirely on them. The level of work undertaken by the validation team may vary, with the overall aim being comprehensive testing that has been externally reviewed.

### 2.3 Risk profile

Regarding model changes, validators should consider the change in capital in the context of the change in risk profile, rather than being led by what the model reports as changes in capital. Evidence of any change in risk profile should be included in the validation report, along with any additional validation tests because of the specific changes.

Agents should include a section which outlines the specific thematic reviews and questions asked for each annual validation exercise, explaining the rationale for the specific questions. Agents should provide the conclusions on how comfort has been gained around the issue being reflected appropriately in the model.

### 2.4 Risk coverage

The objective of the validation process is to test that all material risks to the syndicate are adequately assessed in the internal model. The risk identification process should identify all sources of loss to which the syndicate could have non-trivial exposure. This process should not be restricted to insurance risks; it should, for example, include considerations such as the terms and conditions of the cover issued, data and operational systems, the current legal environment, recent market experience. It should also include the possibility of new sources of loss not experienced by the syndicate or market in the past, for example emerging risks or events not in data.

Many agents have maintained a risk register as a way of identifying risks faced by the syndicate and this could be mapped to model risk to assess coverage. Lloyd's requires that agents develop a comprehensive process for identifying potential risks to the business and assess their materiality on an ongoing basis. This should cover changes in the risk profile and provide affirmation that risks previously assessed as non-material remain so.

If material risks are not appropriately covered within the model, plans should be made for their inclusion. Mitigating measures, including additional capital to cover the associated uncertainty, should be put in place until this is complete.

### 2.5 Risk ranking

Risk ranking is an exercise to define the relative materiality of a defined set of risks. This could relate to ranking high level risks or classes of business within a risk type. This exercise has a dual purpose: to assess whether the ranking matches expectations and to direct an appropriate level of focus.

Agents should view risk ranking as a tool for making the validation process more efficient. It will enable them to allocate resources to areas which are quantitatively or qualitatively material to the internal model.

Lloyd's recommends that the validation policy includes a description of how the outputs of the risk ranking exercise are used within the validation process.

## 2.6 Validation cycle

An effective validation cycle comprises five steps:

- (1) definition of an appropriate programme of testing;
- (2) the application of the validation test or tool;
- (3) the analysis of test results;
- (4) the escalation of test results to appropriate individuals in the business; and
- (5) the implementation of any mitigating measures and/or changes necessitated by the validation test outcome.

### 2.6.1 Definition of Testing Programme and Targeted Validation

An appropriate testing programme should be defined, which is targeted at meeting the aims of the validation exercise. The frequency and range of testing should be proportionate to model risk materiality and changes in the risk profile, i.e. it is not necessarily required to run all tests in all validation exercises. Lloyd's requires that each component is reviewed in detail at least once every three years, this means that less material components do not require detailed validation every year (in the absence of change).

This can ensure that validation activity is directed at the areas which have been subject to change, merging risks, or existing stakeholder feedback. The aim of targeted validation is to encourage smarter validation, not less validation. Lloyd's expects references to be provided to any existing validation work that the validator has relied upon when deriving their overall conclusions (whether it forms part of the current validation report or not). They should also comment on why the existing validation work can continue to be relied upon. This should be part of complete validation reporting to ensure that the Board has comfort that there has been complete coverage of applicable validation testing.

Lloyd's requires that previous years' deep dives are submitted along with the validation report each year. This therefore requires more signposting, e.g. how each deep dive links back to the validation plan, and the work that has been done in the previous years. In moving to the targeted validation approach, agents will naturally place more reliance on previous validation work, be it performed internally or externally. However, the validation process should consider whether the previous validation performed is still appropriate. It is good practice to explain the reason(s) for not carrying out certain validation activity. Agents should outline a series of validation triggers in their validation plan, e.g. changes to methodology/risk profile/external environment, etc. which would mean that previous validation performed can no longer be relied upon and hence may need to be revisited.

Where syndicates have made material changes to an area of the modelling or there is a material risk profile update, Lloyd's expects syndicates to review their planned Testing Programme and consider whether changes are required. Lloyd's does not expect syndicates to rigidly stick to their plan, if changes mean that targeted validation is required in areas that they had not planned to validate for that submission.

### 2.6.2 Tests that Lloyd's expects to be performed each year

Top-down validation tests need to be performed every year and include:

- Assessing the overall movements in the SCRs, in particular with reference to the underlying movements in the risk profile and validation of **any** model changes. Note that the Analysis of Change should be a standalone document, which is required to be submitted with the LCR. Validators can draw on the analysis performed by the first line, but need to validate the changes independently.
- Risk ranking, and any movements in the ranking compared with previous SCRs
- Assessment of materiality of parameters (part of risk ranking) in order to assess the level of core validation necessary
- Backtesting historical events or near-misses

- P&L attribution
- Stress and Scenario Testing, including Reverse Stress Testing
- Stability/convergence testing
- Model output must be validated every year

Whilst these tests need to be performed every year, they should vary depending on the risk profile of the syndicate.

### 2.6.3 Target review test type definitions

To clarify the types of validation testing, in relation to targeted validation, the following are example definitions of different types of test:

- Core tests: these are validation tests which are run annually, regardless of any risk profile or model changes and should focus on material areas.
- Deep dives: these are considered as extensions of the core tests, which are only necessary to run once within a 3-year validation cycle, with the assumption that the relevant part of the model has not been subject to either model or risk profile changes.
- Additional tests: additional testing carried out in response to breaches of certain triggers. Triggers to be defined by the agents.
- Thematic areas of validation: Targeted validation to investigate certain issues/questions of interest to the Board which are not repeated on a regular cycle.

### 2.6.4 Application of tests

All validation tests should have explicit pass/fail criteria. It should be clear from these criteria what would cause a test to fail or result in a pass with limitations (or equivalent). In all cases it should be possible for a test to fail in order to ensure that it is providing a robust assessment of the modelling.

Specific pre-defined criteria, especially for quantitative tests, should be objective. The criteria selected should be adequately justified - for example by reference to materiality thresholds, a priori experience, expert judgement. The criteria should also be linked to the purpose of the validation test.

Pass/fail criteria are also required to be sufficiently robust. Setting criteria which are not robust may result in no or limited fails in the validation process. As a result, the process may miss out issues which should have been escalated.

In general, Lloyd's would expect initial failures (either outright "fails" or "passes with limitations") of some of the validation tests to occur as part of the validation process. A validation process where all tests are initially passed outright does not demonstrate a thorough level of objective challenge or a sound test design framework and may indicate that the pass/fail criteria are not robust enough. Such cases would result in a lack of credibility at Board level and the agent's validation process being subject to a greater degree of scrutiny by Lloyd's.

Validation processes are iterative and validation reports will include the final set of test results, which can often be all passes as a result of changes made following escalation of initial validation test results. It is good practice to summarise this process and include examples of some of the failed tests and resulting actions in order to evidence this key part of the validation process.

### 2.6.5 Analysis of validation test results

The results of validation tests should be justified with clear rationale linked to the pass/fail criteria discussed; this ensures the effectiveness of the test. A good way to provide adequate rationale is to support qualitative conclusions with quantitative analysis or vice versa. This rationale should include conclusions at aggregate level, for example, whether there is any aggregation of tests that do not pass within a particular risk type.

The validation policy should define the next step in the validation cycle, depending on the outcome of the analysis of results. In cases where the outcome is "pass, but with limitations", it is reasonable that for more material risks there should be some sensitivity testing of alternatives. A "fail" should lead to escalation, which is discussed below.



## 2.6.6 Escalation

Elements of a robust escalation process include:

- a well-defined trigger in terms of validation test outcomes and risk materiality;
- a description of the person(s) responsible for the escalation;
- a clear expectation of what is required for resolution of the validation escalation.

The process of escalation and resolution of issues during an annual validation cycle should be described and evidenced in the validation pack to ensure appropriate representation of the validation framework.

## 2.6.7 Change Implementation

The outcomes of the validation should inform any mitigating measures necessary to ensure the appropriateness of capital, such as management add-ons. They should also inform future model development to ensure that material validation issues do not apply to the modelling for extended periods of time.

## 2.7 Reporting

Validation reporting is required to support the Board's decision to agree appropriate capital numbers based on the current syndicate risk profile. Reporting is expected to clearly reflect this primary purpose. Given the volume of work undertaken to support the validation, it is expected that additional detailed reporting is available to the Board or to Lloyd's.

### 2.7.1 Model Change Validation

Any model changes should be validated as part of the validation exercise. Validation is required to accompany any major model change application to Lloyd's and is required to validate the change made as well as the resulting output.

## 2.8 Timing

Validation should be an integral part of model development and use. It may reduce resource constraint for validation to take place throughout a model cycle as many aspects can be validated before the final model inputs are agreed.

To this effect, we do not require that all validation is performed on the final SCR submission, but most parts can be carried out on various earlier model versions. For each part of the validation, the onus is on the validator to ensure that any appropriate updates are made and tests re-run if applicable.

## 2.9 Expert judgement

Expert judgement is relevant to many aspects of both internal models and the validation process. In terms of the validation process, agents need to both identify and validate the key expert judgements used in the internal model and describe how expert judgement is used in the validation process.

### 2.9.1 Expert judgements used in the internal model

All material expert judgements used in the internal model must be identified, documented and validated. The identification of these expert judgements is, in itself, a substantial task as many will be implicit in the design, implementation and operation of the internal model.

It is important that relevant expert judgements are documented. The documentation should include:

- a description of the expert judgement itself;
  - why expert judgement is required;
  - an assessment of the materiality of the expert judgement;
  - an identification of the individual(s) making the expert judgment, including the relevant qualification of that individual to make the judgement; and
- falsifiability criteria.

In addition, it is useful for agents to identify and quantify the subjectivity and sensitivity of expert judgements. Parameters determined using generally accepted actuarial/statistical methods with a high volume of data available may involve little expert judgement and subjectivity. However, estimating a

parameter where there is little data and no widely-used approaches might require high levels of expert judgement and be highly subjective. Expert judgements can also be linked to sensitivity tests to derive a view of how sensitive the internal model is to each expert judgement. A common approach to highlighting the most material expert judgements is to consider both subjectivity and sensitivity in combination.

All material expert judgements must be validated. Validation might include qualitative tools, but a description of the process or methodology is insufficient. It should be clear from the validation report how expert judgements have been validated and this might form part of an expert judgement log.

Whilst it is generally not practical to include each material expert judgement and how it has been validated in the validation report seen by the Board, it is suggested that key details on the most material expert judgements or those that have changed are included within the validation report. The remainder of the expert judgements, and the validation of these, could be described within the validation pack provided to Lloyd's.

### 2.9.2 Expert judgement used in validation

Many validation tests require the use of implicit expert judgement, particularly in the pass/fail criteria of individual tests. It is not necessary to document these expert judgements in the same way as those used in the internal model. However, a rationale for these expert judgements should be included in the validation pack.

A qualitative test might involve a review of a methodology. In that review, limitations of the existing methodology might be identified, and their materiality assessed. Alternative methodologies might be considered, and the advantages or disadvantages compared to the current methodology may be discussed. The expert judgement (in respect of a pass) would be that the methodology is appropriate, but this should be justified in terms of the qualitative assessment described above.

## 2.10 Limitations

Limitations can be grouped into two categories:

- limitations of the internal model itself; and
- limitations in the validation process.

It is important that the Board understands both limitations. Therefore, they should both be clearly identified and communicated in the validation report.

For all material limitations the validation report should identify:

- the nature of the limitation, including whether it applies to the internal model or the validation process;
- how it has been identified, for example via a validation test fail;
- the materiality of the limitation – assessed both quantitatively and qualitatively;
- any remediation steps planned, including timescales; and/or
- the rationale as to why the limitation has been accepted.

Agents should also consider whether accumulations of less material limitations may aggregate to a more material limitation and this should be made explicit in the validation report. An example of this would be where several low materiality limitations are identified, and their materiality provided (materiality could be determined by each limitation's approximate impact on capital) which, together, aggregate to a significant proportion of the SCR.

The validator is expected to identify (as a limitation) any instances where insufficient resources were being allocated to the validation process or the operation of the internal model itself.

## 3 Validation Testing

### 3.1 Overview

The purpose of this section is to provide guidance on some of the validation tests and tools that are expected to be used as part of model validation. The validation tests described are not a complete list of all validation tests: for example, the risk ranking referred to in Section 2 can be used to assess the appropriateness of the capital output. Benchmarking can be considered as part of (but not the sole test for) validation, but the basis for the appropriateness of any benchmarks should be clearly outlined.

The application of these tests is expected to vary by risk profile, modelling methodology and governance process of the syndicate. The tests may be conducted by the validation team or the first line team with oversight from the validation team. The key outcome is that the validation team are comfortable that an appropriate suite of tests has been run and analysed.

Tests should be clearly targeted at the issue that the validator wants to explore. For example, use of a JEP to provide testing of dependency between risk types should consider whether this test is measuring contribution of these risk types to capital and could be distorted by extreme tail observations.

### 3.2 Sensitivity tests

Sensitivity testing has a broad application throughout the model. Lloyd's considers two distinct approaches to sensitivity testing.

#### 3.2.1 Deterministically varying a set of assumptions ("ST-1")

One type involves deterministically varying a set of assumptions (such as loss ratio CoVs) by a given amount and measuring the effect on model outputs. This approach can be used to identify the relative materiality of different inputs. It can also test the mechanics of the model, for example, if outputs do not move in the expected direction, it could be the result of a coding error, broken link, etc.

*Example: A syndicate writes mainly property CAT and other short tailed lines. The SCR is mainly driven by underwriting risk. Sensitivity test shows that increasing CoV of reserving classes by 5% results in reduction in capital by 2%. This is a failure of the test and requires further analysis of the result/model.*

*On further investigation it emerges that lower percentiles of reserve risk contribute to the capital at the 99.5<sup>th</sup> percentile (as it is driven by underwriting risk at the tail). Increasing the volatility results in more skewed distribution for reserve risk and the lower percentiles contributing to the overall capital simulates a lower value (as the mean remains the same).*

The above is an example of where a sensitivity test identifies a possible error in the model. In this case the explanation provides reason for the failure and evidences that outputs from the model are likely to be reasonable.

#### 3.2.2 Plausible alternative set of assumptions ("ST-2")

The second application involves varying the inputs, but by using plausible alternative selections. The choice of plausible alternatives may be guided by a prior validation test, such as a test against experience. ST-2 will be less useful in determining the relative materiality of inputs, since the increase/decrease will vary for different inputs. However, ST-2 has the advantage of reflecting the uncertainty in different assumptions, and therefore in model outputs. As such, the results will be more informative to management.

Lloyd's prefers both approaches to be used. ST-1 has more relevance to risk ranking, whereas ST-2 is more useful for conveying the uncertainty in the internal model. ST-2 requires the additional effort of determining plausible alternatives, so it is more appropriate for material risks. It should be noted that neither approach provides validation of the absolute value of the input being tested. For this reason, sensitivity testing on its own will generally not be sufficient for validation.

### 3.3 Stability tests

This involves testing the stability of outputs using fixed inputs while varying the random seed or number of simulations.

With regards to stability testing, the principal challenges are:

- (1) the time and effort associated with re-running the model; and
- (2) defining appropriate criteria.

For regulatory submissions, internal models are often run on a higher number of simulations to ensure consistency and stability. However, validation tests may be run on lower number of simulations due to challenges around model run time. Results of these tests are thus susceptible to model convergence issues. Stability testing is a key validation tool that evidences possible variations in other validation test results and ensures a robust process. Agents will be expected to carry out this particular test in future validation cycles. Agents should ensure that a sufficient level of stability testing is performed on the SCRs submitted to Lloyd's, supplemented by period 'deep dives' where appropriate.

*Example: as part of the annual cycle for stability testing, a managing agent may consider the results of the model from a few different seeds. A 'deep dive' may further extend this by considering the results from a greater number of seeds (e.g. 10), along with increasing the number of simulations significantly (e.g. from 100,000 to 250,000) for the chosen seed. The validation scope should include performing the 'deep dive' on an ad-hoc basis, for example, following a significant change to the risk profile.*

With regards to (1), it may be possible to obtain an (approximate) indication of the minimum number of simulations required for convergence by running separate simulations for the longest tailed risk in the model.

*Example: if US quake is the main driver of a syndicate's capital and the internal model input is a RiskLink Event Loss Table (ELT), the agent could separately simulate from the ELT to test the number of simulations required to obtain a satisfactory level of convergence.*

Some outputs may be amenable to analytic calculation, such as the Occurrence Exceedance Probability (OEP) in some cases and comparing the simulated output with this accurate figure is a useful test.

It is the responsibility of agents to determine appropriate stability criteria; Lloyd's does however consider stability when assessing syndicate model results. EIOPA guidance also emphasises that models must be stable, and the validation should cover this point. The model should usually be run on a number of seeds and the results examined to ensure that the range is acceptable.

Lloyd's captures stability of modelled outputs within the LCR Supplementary Questionnaire which requires agents to consider a confidence interval around the 99.5<sup>th</sup> percentile value-at-risk (VaR). It is not sufficient for agents to simply complete the appropriate entry in the LCR Supplementary Questionnaire to demonstrate the validation of model stability.

### 3.4 Stress and scenario tests

Stress and scenario tests are valuable tools for validation of syndicates' internal models. As they rely heavily on expert judgement, they are particularly useful where data is limited, such as in the tail of a distribution. In addition, they have the advantage of being readily understood by individuals across the business. Furthermore, despite often being viewed as relatively unsophisticated, the outcomes of these tests can be viewed as coherent risk measures.

There are three key components of stress and scenario tests:

- selection and severity of the stress/scenario (i.e. what events could occur and what would be the financial impact);
- independent estimate of the probability of the event occurrence; and
- comparison with model output with application of a pass/fail criteria.

Stress and scenario tests must be based on realistic assumptions and extreme events in order to be credible. It is imperative that agents provide an explanation or narrative around their stress and scenario tests, otherwise the value of the exercise is diminished. It is also essential that a wide range of “stresses” are tested, including those that are sufficiently severe (for example greater than a 1-in-100-year return period).

The event severities and probabilities should be derived **independently** from the process used to derive the risk distributions in the model. In general, it will be more difficult to assess the probabilities than the severities. It may be helpful in this regard to also consider scenarios at lower return periods, such as 20 to 50 years (twice or once in a career) and then extrapolate or build up to higher return period type events. Scenario tests which are carried out only at lower return periods will in general be of limited use for validation purposes. It is good practice to consider a wide range of return periods in selecting which scenarios to use.

The evaluation of stress and scenario tests as validation tools requires the comparison of the return period from the test scenario with that output from the model. As with other validation tests, agents should clearly state the pass/fail criteria based on the model output.

*Example: if an appropriate expert, based on their experience, estimates that there is a 1-in-40-year return period for the outcome of reserve deterioration exceeding 20%, while the model indicates that this would be a 1-in-100-year outcome, it should be clear whether this result is a pass/fail/pass with limitations.*

One-sided scenario tests are frequently used for validation purposes, which test whether the modelled distribution is at least as volatile as predicted by the expert. An improvement to the test framework would be to introduce two-sided scenario tests. Modelled distributions, if too skewed, can result in unrealistic overall capital estimates. Two-sided scenario tests will ensure that this aspect of a loss distribution is validated.

Stress and scenario tests, if defined appropriately, help validate extreme losses and help assess different areas of the model (e.g. reinsurance) under extreme conditions.

Stress and scenario tests should be performed in the same way as you would a reverse stress test (RST) but where the scenarios target lower return periods than 1 in 200. The key difference between RSTs and stress and scenario tests (SSTs), are that SSTs are usually carried out at risk category or class of business level and usually at a lower return period. Reverse stress tests are expected to be carried out at syndicate level on the overall balance sheet distribution, at least. The objective in this case is to test if the outcome of the model is consistent with expectation, especially at the extreme percentiles.

### 3.5 Reverse stress tests (RSTs)

Reverse stress testing is an essential validation test of the syndicate SCR. Management should have a view on the risks to solvency as well as the opportunities for growth and profit. Reverse stress testing provides the best opportunity to evidence that understanding. The test also helps assess the reasonableness of the entire model and modelled outputs when considered in its entirety. In other words, this allows analysis of the model output, considering various assumptions (e.g. volatility, dependency) and compares it to expectation.

Reverse stress testing begins with consideration of the events or combinations of events that could threaten the viability of the business. ‘Viability’ here is the assumed outcome of business failure as a result of capital depletion and is not necessarily defined by exhaustion of capital. They should reflect the interaction of management (in)actions and external events.

Historically, breaches of the SCR/ECA (Economic Capital Assessment) have been driven primarily by items such as:

- reserve deteriorations on multiple years of account of a casualty portfolio; or
- large natural catastrophes, with a significant proportion of losses coming from perils not fully covered in the vendor models.

When formulating the reverse stresses that could threaten business viability or result in capital exhaustion, a possible approach could be:

- describe the cause of unviability;
- identify the largest drivers of risk to the syndicate;
- generate plausible events within these drivers which could cause significant losses;
- assess the return period and severity of the primary event similarly to scenario testing;
- consider the possible secondary impacts (across all risk categories), and likelihoods given occurrence of the primary event;
- offset any mean profit contributions (if not impacted); and/or
- combine these judgements to assess a combined return period, and loss split by risk category, for the reverse stress to be used for comparison to model output.

Lloyd's has discussed some examples of the challenges around reverse stress tests in the validation workshop held on 22<sup>nd</sup> May 2017, 2<sup>nd</sup> May 2019 (which includes an example RST testing template) and 12 May 2020. Slides presented at previous validation workshops can be found [here](#). A list of the frequently asked questions on reverse stress tests has been included as an appendix to this guidance.

The Reverse Stress Test is not required to have a return period of 1 in 200. These return periods are for the estimated capital requirement; that is, they are at the syndicate level. The stresses must therefore reflect the aggregation of risks across the syndicate, and not only the drivers of the insolvency.

*Example: a reverse stress could be based on the joint occurrence of a US windstorm and reserve deterioration on a casualty class, neither of which on its own would be sufficient to trigger an insolvency.*

The reverse stress test outcome should be based on a comparison between the balance sheet distribution and the reverse stresses. The distribution in the neighbourhood of the selected return period should include aggregations of outcomes that are consistent with the reverse stresses. In this sense, the reverse stress tests can be thought of as a "reality check" on the balance sheet distribution.

Reverse Stress Testing will be a more valuable validation tool if it is carried out at a sufficient level of granularity. As such, the test should define the scenario at risk category level as a minimum and ideally at a class of business level, for at least the most material classes. The test should then ensure that the model output is consistent with the test specification. For instance, if the RST is expected to impact a particular class of business, a return period should be calculated for the loss from that specific class, rather than for losses from all classes.

More generally, it is not appropriate to select a model output simulation that only has the overall loss expected and comment on the return period of this loss. Instead, it is important that simulation(s) are considered that have all of the characteristics of the RST scenario – that is, the loss split should be by class and risk category as defined in the RST scenario.

Syndicates should not only consider the losses from the risk categories and classes that they expect to be impacted by the RST scenario, but also those that they do not expect to be impacted. For instance, if a particular class of business is not expected to be affected by the RST scenario, the syndicate should exclude the profit or loss from that class when calculating the modelled loss for the RST. Syndicates should then check that the losses from the excluded class are random in the sims considered to represent the RST, as expected for an area not expected to be impacted by the RST scenario.

### 3.6 Testing against experience

Testing against experience includes goodness of fit tests and comparisons between model outputs and historical results ("backtests").

In general, the claims history may be too limited to provide conclusive evidence for the optimum assumption or method. By definition, the claims history will not initially include "Events Not in Data" (ENIDs). However, as time elapses, some ENIDs may 'emerge' within the data. Whilst this may be the case, agents must not rely on these 'emerged' ENIDs within the parameterisation process as by definition, ENIDs will not be present within the claims' history.

In many instances, the portfolio will have changed over time. In these cases, there may be good reasons for excluding some parts of the history from the tests against experience. However, such exclusions should be based on objective reasons relating to unique characteristics of the risks, not simply on underwriting results. Conversely, it should also be recognised that a limited history may not capture the

full tail risk of any portfolio; there may therefore be reasons, based on expert judgement, for model risk exceeding that indicated by the data. Lloyd's will not accept agents taking a one-sided approach of excluding unfavourable history as being irrelevant, while not making allowance for tail risk that may not be reflected in the experience.

Syndicates should ensure consistency between modelled output and historical data. For example, syndicates should ensure consistency between use of an ultimate or one-year basis and the inclusion or otherwise of Catastrophes.

Where there are areas of testing against experience that fail or pass with limitations, Lloyd's expects the syndicate to provide detailed justification for that result, backed up by evidence. Syndicates should not presume the reason for a fail or pass with limitations without thorough investigation.

Finally, tests against experience can assist in communication of model outputs to management by linking them with recent results.

Lloyd's recommends that comparisons to past experience are made wherever this is available. These comparisons should include clear explanations, based on expert judgement, of the relevance of this experience to the current risks. They should also describe how ENIDs have been taken into account (the [PRA Supervisory Statement SS5/14](#) gives more information on ENIDs).

*Example: A test which results in a test failure compares the modelled output to historical experience. As a result, the validator analyses the historical information over the past 20 years and adjusts data for rate changes, reinsurance changes and business mix. The validation test is run again, and the result is significantly different. The detailed analysis to explain the failure would have been sufficient, but the managing agent improves it further by adjusting the data and re-running the test.*

Lloyd's discussed some of the issues around Testing Against Experience in the Validation Briefing on 12 May 2020. One particular issue is the requirement to define pass/fail criteria which are robust enough to identify issues. Furthermore, in addition to considering objective quantitative pass/fail criteria syndicates should consider more subjective areas of the testing, such as patterns in the return periods of historical data. This should be considered, even in cases where the test has passed the formal objective criteria

### 3.7 P&L attribution

Article 240 of the Solvency II Directive requires that undertakings review the causes and sources of profits and losses for each major business unit on at least an annual basis. It also requires undertakings to demonstrate how the categorisation of risk chosen in the internal model explains the causes and sources of material profits and losses.

The scope of validation is understandably generally focussed around the 99.5<sup>th</sup> percentile of the risk distributions. Certain validation tests can be utilised to consider more frequent return periods of various distributions (e.g. testing against experience). However, an area which is often overlooked is the mean of the aggregate distribution.

The basic concept is to ensure that all material risk contributions are captured in the model as follows:

- review the actual vs. expected outcome over the period;
- investigate the material sources of profit and loss;
- ensure that the model is sufficiently granular to capture the sources of profit and loss; and
- reflect the results and conclusions in the model validation process.

P&L attribution should have clear and explicit pass/fail criteria and the results should be subject to the same escalation process as all other validation tests. A qualitative description of the results from the P&L attribution and how they compare to model outputs is essential in ensuring the effectiveness of the test.

Lloyd's would expect a "fail" if any material profit/loss drivers are not sufficiently captured in the model (for example, tsunami losses), or a "pass" if all material sources are explicitly captured (for example, if IBNER releases on large motor claims are explicitly modelled).

Lloyd's expects agents to:

- maintain a clear, concise set of policies and procedures setting out the definition of P&L, granularity of business units, categories/line items to be tested and whether these are modelled (or not) and governance of the P&L process;
- conduct, at least annually, a P&L attribution exercise for the previous calendar year. This will include identifying the outcomes to be tested against the equivalent modelled probability distribution function; and
- incorporate the outcomes of the P&L exercise in the model validation process and include results and commentaries in the model validation report.

In line with Article 240 of the SII Delegated Regulations, Lloyd's expects syndicates to allocate the profit or loss to both the risk categories used in the Internal Model and major business units. Note that it is a requirement to attribute the profit or loss to **both** of these levels of granularity, not one or the other.

Lloyd's requires P&L attribution to be extended to a Solvency II basis as outlined in the [EIOPA guidelines](#). At a granular level, P&L attribution on an economic basis may not be practical at present; however, agents should make pragmatic use of available accounting information to consider the economic basis at an aggregated level. For clarity, Lloyd's expects this would be based on the P&L for the latest complete calendar year of experience (for example, for the 2021 YoA validation exercise, the 2019 calendar year P&L would be used) but any more recent experience that gives insight into the adequacy of model risk coverage should be considered. The model version used for the test should be the model currently proposed for use with this exercise being targeted at ensuring the model captures all material risk drivers. Lloyd's recognises that there is an inconsistency between backtesting experience from e.g. 2019 (which should strictly be measured against the 2019 plan) against the model for 2021 YoA. If model changes (including data and risk profile changes) affect the profit and loss attribution outcome it is expected that this is clearly highlighted in the validation.

Where a bridging exercise is performed, from a GAAP basis to SII basis, the resulting SII balance sheet should be compared to model output, with results of this test included and commented upon in the validation report.

Further guidance on the P&L attribution exercise can be found in the [Lloyd's validation workshop slides \(9/10 May 2011\)](#).

In addition to the retrospective P&L attribution, the mean P&L output from the internal model should also be validated prospectively. Given that there are many contributing components to internal models; it is an important test to reconcile all material sources of future profit and loss at the mean, as well as reconciling the opening balance sheet.

### 3.8 Model Drift

Model drift refers to assessing the model basis against exposure measures to capture any high-level trends.

Comparison with Standard Formula is a useful check for modelled outputs (both on an ultimate and one-year risk horizon). The benefit of this exercise will depend on the detailed explanation of the differences in outcomes.

These tests should allow managing agents to monitor model drift. Details of approaches to monitoring model drift have been provided by the PRA ([CP22/16](#)) and Lloyd's expects syndicate validation to comment on these measures.

### 3.9 Analysis of change

This involves comparing the values of key inputs and outputs with those of the previous version of the model. Many agents have made such comparisons with the SCR and at high level risk categories. The analysis should include an investigation of why the values have or have not changed, and reasons as to why the changes are or are not appropriate. The benefit of this exercise will depend on how well the previous version of the model was validated. It is helpful to see a waterfall chart with itemised changes, supported by descriptions of the changes and impact of the itemised changes by risk category, not just overall capital.



## 4 Specific Areas of Validation

### 4.1 One-year validation

Agents are required to determine their SCR on both an ultimate and one-year basis, the latter being a regulatory requirement. The one-year SCR must be appropriately validated with visibility given to that validation within the validation report. The one-year SCR cannot be accepted by virtue of being a function of the ultimate SCR, either in absolute terms or as a function of model changes.

The one-year SCR can be validated using many of the same tools used to validate the ultimate SCR. Lloyd's requires validation tests to cover the methodology and assumptions/parameters used to calculate the one-year SCR as well as the one-year SCR figure itself. Lloyd's expects the validation report to make it clear what validation tools have been applied to the validation of the one-year SCR as opposed to the validation of the ultimate SCR. This could be achieved by either having a separate section on the validation of the one-year SCR or referencing the one-year SCR validation for each risk type.

Lloyd's also expects that any validation results which indicate inconsistencies between the one-year and ultimate SCR are highlighted. For example, a sensitivity test which moved the one-year and ultimate SCR in the opposite directions should be highlighted and explained within the validation report.

### 4.2 Top-down and bottom-up validation

Lloyd's expects a high level of importance and effort to be placed on performing top-down validation, not just bottom-up validation. Bottom-up validation begins with an assessment of the inputs to the model (e.g. an assessment of the drivers behind the dependency at risk category and class level). Top-down validation begins with assessment of the aggregated model output (e.g. analysis of joint exceedance probabilities).

Top-down validation allows senior management/Board and other model users to take a step back and think of scenarios which threaten the viability of the business (Reverse Stress Test) or would cause a great loss to the syndicate (Stress and Scenario Tests) and test if these scenarios are captured in the model output. The results of these then aid discussions within senior management/Board around appropriateness of the modelling methodology and parameterisation.

The table below summarises the main features of each, in the context of validating dependencies (see validation workshop 2<sup>nd</sup> May 2019 for an example of this in use – the slides can be found [here](#)):

Dependencies: two approaches to validation		
Validation component	Bottom-Up	Top-Down
Key validation tools	Sensitivity testing Backtesting	P&L attribution Stress and scenario testing Reverse stress testing
Sample metrics	Correlations Copula parameters	Contribution to capital

Both approaches have their advantages and limitations. The bottom-up approach explicitly addresses the appropriateness of the individual drivers of dependencies between pairs of risks; the limitation is that it can provide only indirect assurance on the aggregate result. Conversely, the top-down approach can provide comfort that the aggregate result is appropriate but will be of less use in providing assurance on the appropriateness of the individual drivers of dependency.

### 4.3 Dependencies and aggregation

Dependencies are frequently a very material driver of capital and the validation of them is challenging. The assumption of independence (either explicit or implicit) between variables is a strong assumption and should in itself be subject to validation.

In the early stages of embedding validation, agents relied primarily on bottom-up validation based on sensitivity testing of correlations. Whilst this approach can play an important role in the validation of dependencies, it has some significant drawbacks when used on its own:

- firstly, many agents relied on Gaussian copulas, which do not introduce tail dependence; therefore, varying the correlations had little effect on the overall 99.5<sup>th</sup> percentile. A more robust sensitivity test would consider alternative dependency structures;
- secondly, there was often only limited support (either qualitative or quantitative) provided to justify the choice of correlations or dependency structure; and
- thirdly, there will frequently be drivers of dependency not explicitly captured in the model; the correlation-based bottom-up approach may not capture these other drivers.

At a lower level, agents have made (for example) comparisons between historical YoA loss ratios aggregated across classes of business with the distribution of aggregated output from the model. Like the bottom-up validation, the top-down approach will not be sufficient on its own.

Consideration of alternative dependency structures is an effective way of validating dependency assumptions. The success of such a test would depend on the rationale provided for selecting one structure over alternatives. Validation tests of this kind that would require significant changes to the calculation kernel are not expected to be carried out annually.

Scenario tests are often used to validate the volatility/adequacy of stochastic distributions at a class of business or risk category level. Scenario tests that impact more than one class can be used to validate not only the distributions for those classes, but also the dependency assumptions between them.

It is challenging to use backtesting as a validation test for dependency for a number of reasons. Lack of data is one of the common issues, which should be supplemented by stress and scenario testing of any area where combinations of losses are expected to present a risk to the business.

Validation of the modelling methodology can be extended to include a comparison of input and output correlations. Depending on the dependency structure used, the modelling software might adjust the input correlation coefficients to ensure that the modelled correlation matrix is positive semi-definite. This can lead to lower output correlations. If such a comparison is used for validation, then reasons for dampening of correlation coefficient should be explained. In the event of significant differences, the test should explain the reason for the continued legitimacy of the input correlation selected/justified by experts.

#### 4.4 Outwards reinsurance

Outwards reinsurance can provide a material capital benefit. The validation report should highlight the extent to which the SCR benefits from outwards reinsurance and explain how this area has been validated.

A range of validation tests for outwards reinsurance that could be considered are:

- an analysis of the distribution of outwards reinsurance recoveries, including a comparison with maximum recoveries from the reinsurance programme;
- a comparison of the distribution of outwards reinsurance loss ratios versus the distribution of inwards loss ratio distributions;
- testing interaction of the benefit of reinsurance programme cover with reinsurance credit risk. If the addition of reinsurance provides a benefit to capital it is expected to contribute to an increase in RI credit risk;
- an analysis of reinsurance recoveries in individual simulations driving the more extreme return periods and a comparison with a manual calculation of recoveries for those losses performed by the outwards reinsurance team; and
- for areas where a net to gross ratio approach is used to estimate net losses for earned business, then the validation process should evidence the continued appropriateness of the selected ratio when compared to the existing reinsurance programmes.

For material and non-standard contracts (e.g. Stop Loss, Aggregate Catastrophe XoL, Adverse Development Cover (ADC) or Industry Loss Warranty (ILW)), detailed analysis of the impact and adequacy of the programme is expected to be considered as part of the validation process.

*Example: if a syndicate uses an ILW to mitigate against catastrophe losses, then comparison of ILW recoveries corresponding to gross catastrophe losses could provide useful comparison for validation purposes. Analysis showing low ILW recoveries against significant modelled catastrophe losses would show that the model allows for situations that capture unfavourable but possible extreme scenarios.*

In addition to usual validation tests, reasonableness checks of assumptions for reinsurance programmes are beneficial for stakeholder/regulator assessment.

*Example: a syndicate reinsuring extreme reserve deteriorations through an ADC could analyse certain metrics (probability attachment, exhaustion, mean recovery) at stressed layers. Different values of the metric should indicate the reasonableness of the modelling methodology. Additionally, the metrics (for different layers) should be compared to expectation and that should evidence reasonableness of the programme.*

## 4.5 External Models

If the internal model uses third party or external modelling to inform some risk types (most commonly ESGs and catastrophe models) these need to be validated as part of the Board's consideration of capital. A change to a third-party provider view of risk cannot be automatically accepted as representing an update to the syndicate's view of risk and this update must be validated in the context of the syndicate's risk profile.

Where external vendor models are used as an input to the internal model and a newer version is available than that used, the choice of version (whether moving to the newer version or not) should be justified in the validation report. Additionally, the validator would be expected to comment on the appropriateness of a particular vendor model in comparison with others. However, typically this would be performed as stand-alone piece of work periodically, and hence it would need to only be referenced.

One specific issue arising in relation to the choice of model version involves clustering of events as some agents use catastrophe models that do not include this. Lloyd's regards this as a material limitation which should be explicitly identified and assessed in the validation report.

Validation should also include an explanation to justify modelling of, or mitigation against, the uncertainty of loss estimates from external vendor models.

### 4.5.1 Model completeness

Syndicates must continue assessing whether they have exposure to sources of loss not covered (or only partially covered) by external vendor models. Internal models must include all material causes of risk so that the derived SCR is complete and accurate. The methods used will vary according to available information and materiality.

The validation report should make clear how the validator has gained comfort that Non-Modelled Risks are either not material or have been adequately captured in the internal model.

## 4.6 Events not in data (ENIDs)

Many agents use historic data to a greater or lesser extent either directly or indirectly in their internal models. This could be in the setting of volatility assumptions, the selection of statistical distributions to be used or choosing the most appropriate copula for establishing a dependency structure between two (or more) variables. In many cases, this historic data will be limited in volume and, by definition, will not include ENIDs.

As well as considering how appropriate allowance for ENIDs is made in their internal models, agents must also consider how ENIDs are incorporated into the validation process.

*Example: it is not uncommon in respect of backtesting to find portions of the history being excluded but no allowance being made for the fact that remaining experience may now not reflect sufficient tail risk.*

Lloyd's would generally expect ENIDs to be included in the model, or a credible explanation for not considering any ENIDs to be included. In any event, the validation process must address how the

validator has made themselves comfortable that the model contains sufficient allowance for ENIDs and how the validator has allowed for ENIDs in the validation tests performed (particularly backtesting). These are similar but not identical issues.

Consideration of ENIDs should include any emerging risks to the syndicate which have do not have a track record in the data. This includes consideration of climate change.

#### 4.7 Special Purpose Arrangements (SPAs)

Although SPAs at Lloyd's typically use the same internal model as the host syndicate, their SCRs are still required to be validated. It is a requirement of the Solvency II Directive that certain validation tests are carried out for each insurance or reinsurance undertaking at least annually - these include (but are not limited to) P&L attribution and reverse stress testing.

Due to the quota-share arrangement in place with the host, certain aspects of validation for the host syndicate will be directly applicable to the SPA. For example, in respect of the validation of a large loss distribution, reference can be made to the relevant validation of the host syndicate in the SPA's validation.

However, the quota share may only cover certain years of account, have different shares for different years, or may only cover certain classes. Therefore, the risk profile (not necessarily limited to insurance risk) for the SPA can be significantly different from the host, and hence a sufficient breadth of validation tests should still be performed for SPAs.

*Example: the exposure of a SPA comprises a 20% whole account quota share program from another syndicate. As such, the SCRs for the SPA are broadly scaled proportionally to the host syndicate. This leads to the operational risk for the SPA being £0.3m. Whilst the operational risk may appear immaterial and hence may not receive scrutiny during the validation process, validation must consider whether the contributions to the SCRs are appropriate. For example, if an operational failure were to occur (e.g. high staff turnover or loss of an individual of key dependence), then the mitigation of this failure could be the hiring of external consultants or contractors, where the cost may be much greater than modelled operational risk figure.*

It is common for agents to include the validation of the SPA in the same validation report as the host syndicate. Depending on the individual circumstances, this may be a reasonable approach that avoids duplication and indeed assists in senior management's understanding of the validation of the SPA. However, in all cases sufficient reference must be made to the SPA individually for the Board to satisfy itself that the validation of the SPA has been appropriate.

At a minimum, validation of a SPA should include:

- stress and scenario tests,
- an RST,
- a P&L attribution on a SII basis
- any relevant sensitivity tests,
- testing against experience

If any of the host's validation tests are used by the agent to validate their SPA(s), the agent must explain how the host's validation test is appropriate to the SPA and why.

#### 4.8 Post diversified contribution to capital

Agents are expected to validate their capital results on a post diversified basis, at a total level and risk category level. In general, the principle should be adhered to that additional risk should add to capital, and the natural extension of this is that no addition of risk to the syndicate profile should have a negative contribution to capital.

## Appendix A: Reverse stress tests: frequently asked questions

- 1) Q: A syndicate is set up such that a combination of events that leads to a 1:100 overall loss threatens unviability of the business (e.g. capital replenishment from the group or loss of capital provider support). What return period should a syndicate consider for reverse stress test?

A: The syndicate may, in this case, consider a return period lower than 1:200. However, the scenarios should be appropriate to test the various relationships within the model at syndicate level.

- 2) Q: A syndicate writes short tailed business and the capital is mainly driven by natural catastrophe. The syndicate provides a detailed narrative of the catastrophe that results in extreme losses for the syndicate. Is any further information required?

A: Information only on the main event driving the loss in the scenario is not sufficient. The reverse stress test should also include details of the expected results from other areas of the model (e.g. at class of business level or by risk category). Most notably, the knock-on (secondary) impacts should be assessed qualitatively and quantitatively.

- 3) Q: A syndicate carries out a reverse stress test for the ORSA report. A set of scenarios leading to extreme loss is constructed along with return period and risk mitigation methods explained. Are any additional steps required for validation purposes?

A: The validation process requires the additional step of comparing the set of scenarios to the model. The expert should comment on the return period, and a collar of simulations (defined by the validator/validation process) around the return period from the model should be analysed. It is expected that at least one simulation would exist that is consistent with the scenario, evidencing the reasonableness of the model; its assumptions, parameters and outputs.

- 4) Q: What level of granularity should the comparison between the reverse stresses and model output be performed, and how should this be considered in the pass/fail criteria?

A: We expect that the comparison is performed to at least the risk category level (and hence the reverse stresses should also be specified to this level of granularity). However, you may find it beneficial to extend this to a more granular basis e.g. class of business or catastrophe peril if deemed appropriate.

The pass/fail criteria should not only consider the absolute value of the loss but also the proportion contributed by each risk component. The objective of the comparison is to find a simulation which is representative of the reverse stress; this may be identified by considering a range of absolute values, or a range of % contributions to the total loss, for each of the risk components which are impacted by the reverse stress. This is a more robust set of criteria than the identification of a simulation which 'approximates' the reverse stress.

*Example: the test may be categorised as a pass if a simulation is found within a suitable collar of simulations around the return period with the below proportions:*

- 60%-80% catastrophe risk
- 15%-25% credit risk
- 5%-15% operational risk

- 5) Q: Lloyd's has requested that reverse stress tests are carried out annually, does this mean that we have to review our reverse stresses every year?

A: Not necessarily, Lloyd's only requires that a comparison is performed between the reverse stresses and the latest model output (used for the LCR submission). If the syndicate's risk profile is stable, then the same drivers of the reverse stresses would likely still be appropriate, hence only the comparison to the model output needs to be updated. In this case, the reasons to believe that the same scenarios represent the risk profile of the syndicate should be adequately explained.

- 6) Q: Due to the limited number of simulations within the collar around the extreme RST return period, it is not possible to find a suitable simulation within this interval. Do I need to revise my reverse stress scenarios?

A: A test failure does not necessarily imply that the internal model is invalid. If a reasonable explanation has been provided on how the validator has gained comfort given this result, then this may not require revision of the scenarios. It is acknowledged that at extreme percentiles, which have not yet been experienced by experts, desired accuracy may not always be achieved.