

Internal Model Validation Guidance

Lloyd's Guidance on Solvency II Internal Model Validation

January 2023

Contents

1	Introduction	3
1.1	Purpose	3
1.2	Context	3
1.3	Reporting	3
1.4	The Role of the Board	4
2	Qualitative Components of Validation	5
2.1	Overview	5
2.2	Validation Policy	5
2.3	Independence	6
2.4	Risk Profile	7
2.5	Risk Coverage	7
2.6	Risk Ranking	8
2.7	Frequency and Scope	8
2.8	Reporting	11
2.9	Timing	12
2.10	Expert Judgement	12
2.11	Limitations	13
3	Validation testing	15
3.1	Overview	15
3.2	Top-down and Bottom-up Validation	15
3.3	Sensitivity Tests	15
3.4	Stability Tests	16
3.5	Stress and Scenario Tests	17
3.6	Reverse Stress Tests	18
3.7	Testing against Experience	19
3.8	P&L Attribution	20
3.9	Model Drift	22
3.10	Analysis of Change	22
4	Specific Areas of Validation	23
4.1	One-year Validation	23
4.2	Dependencies and Aggregation	23
4.3	Outwards Reinsurance	24
4.4	External Models	24
4.5	Climate Change	25
4.6	Non-Natural Catastrophe Risk	25
4.7	Events Not In Data	26
4.8	Special Purpose Arrangements	26
4.9	Post-Diversified Contribution to Capital	27
4.10	Other Areas to Consider	27
5	Appendix A: List of Acronyms and Abbreviations	28
6	Appendix B: Changes from Previous Versions	30

1 Introduction

1.1 Purpose

The purpose of this document is to outline Lloyd's requirements with respect to internal model validation. This document is intended for use by and is specific to managing agents and syndicates to provide guidance on internal model validation under Solvency II.

This model validation guidance should be read in conjunction with the ['Principles for Doing Business'](#) at Lloyd's ("The Principles"), other Lloyd's communications on validation and relevant EIOPA and PRA documents.

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.

In this guidance the words "must" or "required" mean a specific mandatory requirement. In contrast, the guidance uses "should" or "expected" to indicate that, while the presumption is that syndicates / managing agents comply with the provision in question, it is recognised that there will be some circumstances in which syndicates / managing agents are able to justify non-compliance.

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, 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).

The primary purpose of the validation pack is to demonstrate that agents have completed sufficient work to gain confidence that the model is capturing all the risks to which syndicates are exposed, that these risks are modelled appropriately, and that key stakeholders have a thorough understanding of the materiality and sensitivities around

each element of risk. Validation is a fundamental part of acceptance of the internal model into the business and the validation pack should reflect this.

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, Agents 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 validation findings are reported to the board;
- board involvement in the process of remediating validation issues from escalation 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.

These confirmation statements should be included in the validation report. The validation process is required to be conducted and validation tools selected and applied in compliance with the validation policy. It must support conclusions about the appropriateness and reliability of the internal model. Agents must ensure that:

- the risk management function has responsibility for testing and validating the internal model;
- validation tests result in conclusions about the appropriateness of the area being validated;
- tests giving results outside pre-defined requirements are addressed in accordance with the validation policy and shall be subject to escalation and further validation, if appropriate; and
- the validation process covers all material limitations in the internal model and their impact on the appropriateness and reliability of the model.

Actuarial and statistical methods should be tested on current information and practice and a detailed understanding of the model. The validation of the actuarial and statistical methods should be based on:

- current information, including actuarial progress and the generally accepted market practice; and
- a detailed understanding of the theory and assumptions underlying the methods.

Agents are required to use current actuarial methods and techniques and relevant, up-to-date information in validating internal models. This is essential in ensuring that agents can continuously improve their models. This does not imply that internal models should always be based on the most commonly used methods; agents remain under a responsibility to seek out the best approach for reflecting the business and risk profile of their syndicates.

Agents should demonstrate their detailed understanding of the theory and assumptions underlying their models in the design and application of a specific, detailed and appropriate validation programme including the use of relevant tools. This understanding should be supported by the design and operation of the validation process, and validation failures should be used to improve understanding, as well as enhancing the model itself. This area forms an integral part of the validation testing undertaken by an agent to ensure that the actuarial approach is appropriate.

The validation process must be supported by comprehensive reporting. This is to ensure record of testing performed, actions taken and final conclusions drawn, including the process to get there.

2.2 Validation Policy

Managing agents are required to establish, implement and maintain a validation policy which sets out the way in which they will validate the internal model, and explains why their approach is appropriate. Agents must ensure that the validation policy specifies at least:

- the purpose and the scope of the validation process;
- the processes and methods used to validate the internal model;
- for each part of the internal model, the frequency of regular validations and the circumstances which trigger additional validation;
- the persons who are responsible for each validation task;
- the procedure to be followed in the event that the validation process identifies problems with the reliability of the internal model and the decision-making process to address those concerns;
- an assessment by the agent of the quality of the validation process; and
- an assessment of the independence of the validation process.

Agents should include any other material matters relevant to their approach to validating the internal model as satisfactory coverage of these elements alone does not guarantee compliance with the tests and standards. The scope of the validation will depend on the scope of the internal model itself. The validation policy must cover all quantitative and qualitative aspects of the internal model. If, exceptionally, there are parts of the internal model not covered by the policy, agents will be required to state this and detail why it is appropriate to exclude those parts from the validation process.

The validation policy must set out the validation tools that are used. The broad description of a validation tool is any approach designed to gain comfort that the internal model is appropriate and reliable. It is up to each agent to decide which tools will be used. Various validation processes and tools that an agent may choose to employ are set out in this document. Many of the validation processes and tools described are mandatory under SII level 2 guidance, including stress and scenario testing, sensitivity testing, and back testing. These tests should be applied to every major risk category in order to validate the model.

A key element of the validation process is the setting of pass/fail criteria for validation tests, since failures could indicate problems with the reliability of the internal model. The validation policy must set out how such cases are to be managed, including escalation as far as the board, if appropriate. Validation failures may also lead to the recognition of model limitations and management add-ons to capital, the escalation process for which should also be covered by the validation policy.

The validation policy must demonstrate that it sets out an independent validation process and must explain how independence is achieved.

2.3 Independence

Agents must ensure that:

- the validation process is independent from the development and operation of the internal model;
- the validation process and outcomes are subject to a level of objective review and challenge sufficient to justify the conclusions reached on the quality of the validation process;
- the validation process incorporates governance to ensure that validation results are escalated where necessary; and
- the escalation path for validation results should also be independent of the development and operation of the internal model.

The agent's definition of independence should be detailed and explained. This means generally that it is reviewed by individuals who were not involved in, and free from influence of, the model's design, development, parameterisation or operation, which could include external and internal staff. Agents have used external resources either within their company group or third-part consultancies to ensure that appropriate independent challenge is maintained within the validation.

Independence does not require that only those with no connection to the model should be involved in the validation, only that all aspects of validation should incorporate an independent view. This allows for the modelling team to conduct validation which can be used in the independent validation process, provided an adequate independent review element is incorporated. In particular, the setting of pass/fail criteria should incorporate adequate independence.

In addition, agents should ensure that independence extends to reporting lines. Where the risk management function responsible for validation and the modelling team report into the same person, steps should be taken to ensure governance is in place to facilitate the escalation of validation findings without any material conflicts of interest. Agents need to be able to demonstrate that independent review is maintained. Where reporting lines are not independent, the agent should demonstrate that measures are in place to ensure independence in the validation process. Some agents have achieved this by demonstrating that there is an independent intermediate step in the validation process, such as a review committee, which is independent of the executive to whom the modelling and validation teams report. Others have used periodic external validation or the involvement of another director. Where it is not clear that the validation owner and team performing the validation are sufficiently independent of the modelling team, some agents have provided a table mapping out validation responsibilities by risk category to demonstrate how independence is ensured.

Owing to the wide-ranging scope of the internal model and of the validation of the internal model, it may be that the risk management function is not always in the best position to perform the tasks set out in the validation policy. It

may be more practical in some cases for other parties to complete some of the tasks required in the validation process. The risk management function must remain responsible for co-ordinating the completion of the various tasks.

The independent review will require various skills and expertise and agents must be able to demonstrate that the process involves sufficient skills in all areas of the model to be validated. The agent should also explain the role of the internal audit function in the validation of the model where they are involved and show that it is not incompatible with the normal “third line of defence” role. The agent needs to confirm that validation has been conducted in line with the validation policy and that the appropriate level of independence has been maintained in the validation. If the internal audit function assesses the compliance and independence of the validation, then it needs to be clear that they are sufficiently independent of the validation to make this assessment. Third party reviewers should have terms of reference making clear the nature of their role.

Agents must also consider how independence is maintained over time. As an example, if model changes are implemented in response to an independent review, the review of the change by the same reviewer in future validation cycles may result in a decrease in independence over time. A proportionate approach to maintaining independence over time needs to be taken to ensure that it is manageable.

Assessment of independence should explicitly cover as a minimum:

- in relation to the internal validation process, the responsibilities and reporting structure of the persons involved in the process; and
- in relation to any external validation process, the remuneration structure of the persons, including where applicable their employees or other persons acting on their behalf, who are involved in the process and other mandates of these persons relating to the agent or syndicate.

The existence of objective challenge is crucial in maintaining an effective validation process. It is the responsibility of the risk management function to ensure that the procedures set out in the validation policy create and maintain independence, as noted above.

Agents must be able to demonstrate the existence and use of a clear escalation path setting out how the results are escalated within the governance structure of the internal model. The escalation procedure followed for failed tests must be free from constraints on independence and be undertaken in compliance with the validation policy.

2.4 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.

2.5 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. Agents must be able to show that all material, quantifiable risks are covered within the scope of the internal model, and hence within the capital calculation, and that these are adequately validated.

2.6 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. The risk ranking must be of sufficient quality that the agent can rely on it to make decisions about model use, validation and other business functions, such as risk management. Agents must document the risk ranking. There must be an escalation process, and a feedback loop to support continuous improvement of the ranking.

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.

2.7 Frequency and Scope

Agents should ensure:

- the validation process shall apply to all parts of the internal model and shall cover all requirements set out in Articles 101, 112, 120 to 123, 125 and 126 of Directive 2009/138/EC;
- they test the results and key assumptions of the internal model at least annually; and
- they test the results and key assumptions following events as identified in the model change policy.

The validation of the calculation kernel of the internal model should include premium risk (ex-cat), catastrophe risk, reserving risk, market risk, credit risk, operational risk, dependencies and diversification. There should be appropriate validation on both an ultimate and one-year basis.

The scope of validation is wide and covers the full scope of the internal model and not just the calculation kernel. For example, any IT systems defined as being included within the internal model should by default be considered for inclusion within the scope of the validation. Agents will need to consider carefully where to draw the line between policies within the scope of the internal model and policies and processes that are linked to the internal model but outside the scope of the approval process. All areas within the scope of the model must be covered by validation.

The SII level 2 guidance requires the validation process to at least cover the following areas (alongside each, examples of what validation might look like are given):

Area of validation	Example
Data	Accuracy, appropriateness and completeness checks, data reconciliations
Methods	Selected statistical distributions are supported by the data (or at least are not contradicted by past history)
Assumptions	Validated against actual experience or market data
Expert judgement	Actuarial judgement is challenged by underwriters' views and vice versa
Documentation	Evidence that it has been read and approved by relevant staff
Systems/IT	Appropriate systems controls, file back-ups and security in place
Governance	Evidence that validation output is reviewed by appropriate committees
Use	Evidence that the model is genuinely used and appropriate for the use. Show minutes, supporting analysis and any workflow history Show how use influences model design and subsequent changes

A full cycle of validation is expected to contain:

- definition of an appropriate programme of testing;

- the application of the validation test or tool;
- the analysis of test results;
- the escalation of test results to appropriate individuals in the business; and
- the implementation of any changes necessitated by the validation test outcome.

The full cycle of validation of the internal model must be at least an annual process, which does not necessarily mean every part of the validation occurs strictly every twelve months. Some parts may be more frequent or triggered by events such as changes in the business plan or external environment.

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 targeted to areas which have been subject to change, emerging risks or stakeholder feedback.

The timing of the validation cycle also needs to be capable of supporting calculation of capital requirements and model change for regulatory submissions.

2.7.1 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 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 the validation pack 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, syndicates must review their testing programme and consider whether changes are required. Lloyd's does not expect syndicates to rigidly stick to original planning. If changes occur that mean targeted validation is required in areas that had not been planned for, this is acceptable and should be documented in the validation pack.

2.7.2 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 needed
- Backtesting historical events or near-misses
- P&L attribution
- Stress and scenario testing, including reverse stress testing, the results of which should be monitored and assessed on an ongoing basis and updated annually

- Stability and 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. Agents are expected to be able to demonstrate use of a wide range of appropriate and effective validation tools, based on defined and documented selection criteria. They should be able to explain how the selected tools meet the validation objectives and why they are more appropriate than alternatives.

Agents should have a documented process in place to choose the appropriate suite of validation tools in order to ensure a robust validation process; this may be covered adequately in the validation policy in some cases. The following characteristics or issues should be considered when selecting the validation tools:

- Level of sophistication – validation tools range from simplified techniques to sophisticated methods;
- Nature – validation tools may be qualitative, quantitative or a combination of both;
- Knowledge required - extent of knowledge required by the persons doing the validation;
- Independence - level of independence required by the persons doing the validation;
- Information available – the extent of the types of information available for external versus internal validation; and
- Cycle of validation – validation will be performed at different stages of the internal model from development, to implementation, to operation.

In addition to agents developing their own suite of stress tests, Lloyd's may specify market-wide stress and scenario tests, including the existing Realistic Disaster Scenarios. As a required tool, stress and scenario tests will be reviewed at least annually as part of the validation process, but Lloyd's takes the view that agents should reconsider their scenarios whenever the circumstances or information on which they depend changes materially. Agents should document the circumstances under which scenarios will be reviewed.

2.7.3 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 of the model and risk profile.
- Deep dives: these are considered as an extension of core testing, 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 are to be defined by agents.
- Thematic areas of validation: targeted validation to investigate certain issues/questions of interest to the board which may not repeat on a regular basis.

2.7.4 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 or 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. Lloyd's expects agents to summarise this process and include examples of escalated test results and how these were remediated in order to evidence this key part of the validation process.

2.7.5 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.

2.7.6 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; and
- 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.

2.7.7 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.

2.8 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 and to Lloyd's.

The report (and supporting technical documentation in the validation pack) must:

- outline the validation performed for each syndicate under management;
- be submitted to Lloyd's at least annually in line with the LCR submission timetable; and
- include the validation results, and conclusions and consequences from the analysis of the validation results.

In addition:

- if parts of the validation have been delegated to internal or external parties, the data, tools and results should be documented in the validation pack;
- the validation pack should describe the validation tools used and the reasons for selecting them;
- the validation pack should include a description of known limitations in the internal model and the validation of the internal model, including the impact and materiality of the limitations. Accumulations of limitations, both of the internal model and the validation should also be considered and noted. Action taken to mitigate the effect of the limitations should be described; and
- the validation pack should describe the actions taken to ensure the adequacy of independence and objective challenge

- the validation pack should include examples of the escalation process in action i.e. examples of issues that have been raised from failed tests; how these were escalated and where to (e.g. risk and technical committees); and the outcome such as changes to the testing approach, changes to the model or implementation of a remediation plan.

As part of 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 process in the latest validation cycle.

In addition to the planned validation according to the 3-year cycle, agents should include in their validation pack a section which outlines validation outcomes to Lloyd's thematic reviews, validator questions and associated results. Thematic reviews, when they do happen, will be communicated separately by Lloyd's. Validator questions should be specific to the syndicate's experience over the last year, or to the particular nature of the syndicate's business.

The primary purpose of the validation pack is to demonstrate that agents have completed sufficient work to gain confidence that the model is capturing all the risks to which syndicates are exposed, that these risks are modelled appropriately, and that key stakeholders have a thorough understanding of the materiality and sensitivities around each element of risk. Validation is a fundamental part of acceptance of the internal model into the business and the validation pack should reflect this.

The validation report should be primarily an internal report that the agent uses to give its board comfort that the internal model and corresponding SCR calculation are appropriate for the business and meet the relevant regulatory requirements. For this reason, Lloyd's does not intend to mandate the exact format or content of the reporting. Agents should first and foremost produce a report that is appropriate for their business and internal model structure. This should be supported by and make reference to more detailed work that is included in the validation pack.

Agents can include the validation of more than one syndicate under management in the same validation report. Depending on the individual circumstances, this may be a reasonable approach that avoids duplication and assists in senior management's understanding of the validation of the syndicates. However, agents should first discuss this with Lloyd's and in all cases sufficient reference must be made to each individually for the board to satisfy itself that the validation of them has been appropriate. The report should also clearly outline where and why testing for one syndicate has been relied upon for another. In particular, as for SPAs which are covered later in the guidance, there should at least be SSTs, RSTs, P&L attribution, testing against experience and relevant sensitivity testing that is specific to each syndicate.

2.8.1 Model Change Validation

Any model changes should be validated as part of the model change process. 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. This may therefore require specific validation of major model changes outside of the usual annual validation exercise. If the major model change application accompanies the annual LCR submission some model changes (like e.g. parameterisation changes) might be covered by the usual annual validation exercise. However, validators should tailor the validation report to include specific validation of model changes when required. Please refer to the [Model Change Guidance](#) for further details.

2.9 Timing

Validation should be an integral part of model development and use. It may reduce resource constraints if validation takes 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 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 are re-run if necessary. The validation pack should include appropriate information to demonstrate that validation applied to earlier versions of the model remains sufficient to validate the current model, and provide bridges between versions where appropriate.

2.10 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.

2.10.1 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. It is useful to show the impact on capital of a range of plausible alternatives for the key expert judgements to show that the selected assumptions sit at an appropriate place within a reasonable range. This is discussed further in section [3.3.2](#). The remainder of the expert judgements, and the validation of these, could be described within the validation pack provided to Lloyd's.

2.10.2 Expert Judgement used in Validation

Many validation tests require the use of implicit expert judgement, particularly when applying pass/fail criteria to individual test outcomes. 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.

For example, 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. If limitations are potentially material, alternative methodologies should be considered, and the advantages or disadvantages compared to the current methodology should 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.

2.11 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 must 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 2.6 can be used to assess the appropriateness of the capital output. Benchmarking can be considered as part of validation, but the basis for the appropriateness of any benchmarks should be clearly outlined, and benchmarking should not be the only form of validation used.

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 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 (RST) or would cause a great loss to the syndicate (SST) 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 2nd May 2019 for an example of this in use – the slides can be found in the [Model Validation](#) section of Lloyds.com):

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

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.

3.3 Sensitivity Tests

The sensitivity of the internal model results to changes in key assumptions should be tested. Sensitivity testing has a broad application throughout the model.

Agents should assess the effect on the results of the internal model of changes in the key assumptions underlying the model. Where the results of the internal model are highly sensitive to an underlying assumption, agents should be able to explain the underlying reasons and how this is taken into account in their decision-making process.

There are multiple objectives that sensitivity tests can help achieve. At a basic level, they can be used to check the model is working as expected i.e. if a certain parameter is increased then an agent can check that the model output has moved in the expected direction. Sensitivity testing can also be used in validating parts of the internal model which place reliance on expert judgement, to assist in determining their materiality. In particular, testing the sensitivity of the internal model may also be useful to identify cases where a small difference in the input leads to significant changes in the output.

Lloyd's considers two distinct approaches to sensitivity testing.

3.3.1 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 99.5th 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 – i.e. the model is working as intended. However, consideration should then be given to whether this is reasonable or indicates that the modelled dependency is inappropriate.

3.3.2 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 expects 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.

3.4 Stability Tests

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

Stability testing may also include re-running the model with small changes to parameters.

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 should ensure that a sufficient level of stability testing is performed on the SCRs submitted to Lloyd's, supplemented by periodic '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 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 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 and appropriate figures are reported in the LCR.

Lloyd's captures one measure of stability of modelled outputs within the LCR, which requires agents to consider a confidence interval around the 99.5th percentile VaR. However, this does not test model stability in relation to varying model seeds. It is not sufficient for agents to simply complete the appropriate entry in the LCR to demonstrate the validation of model stability.

3.5 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 at the SCR level but may help validate other parts of the distribution. 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 RST but the scenarios may target different return periods than would threaten syndicate viability. The key differences between RSTs and SSTs, are that SSTs may be carried out for individual components of the model, such as 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.

Stress and scenario tests should include consideration of Realistic Disaster Scenarios (RDS) as reported to Lloyd's to ensure that the model incorporates these scenarios and captures them at a reasonable return period. This should include natural catastrophe and non-natural catastrophe RDS relevant to the syndicate.

3.6 Reverse Stress Tests

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. Thus, the RST might be based on a return period below 1:200. They should reflect the interaction of management (in)actions and external events.

Historically, breaches of the SCR/ECA 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.

The RST can be considered as a type of scenario test to check that the model adequately captures scenarios that the management consider could make the business unviable. This is a top-down validation which may be regarded as validating the model in two stages:

- (1) Does the model capture the nature of events that might threaten the business? This may be regarded as a qualitative test that the risks identified are considered within the model.
- (2) Is the model calibration consistent with independent experts' views of the risks? This is done by comparing the independently assessed return period of the RST with that observed in the model. In making this

quantitative comparison, care is required that the independently specified return period is clearly defined and that the model is interrogated in a manner that is consistent with this definition.

It is acknowledged that at extreme percentiles there is uncertainty over the independent return period for the RST. Therefore, an RST test failure does not necessarily imply that the internal model is invalid. However, it would indicate that further consideration should be given to the model calibration and the validation documentation should provide a reasonable explanation on how the validator has gained comfort given the test result.

Lloyd's has discussed some examples of the challenges around reverse stress tests in the validation workshop held on 22nd May 2017, 2nd May 2019 (which includes an example RST testing template) and 12 May 2020. Slides presented at previous validation workshops can be found [here](#).

The RST 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.

3.7 Testing against Experience

Agents should ensure that appropriate testing of the internal model against experience is undertaken and:

- test the results and the key assumptions of the internal model against experience and other appropriate data to the extent that data are reasonably available;
- apply the tests to aggregated results and appropriate single elements of the results. At a minimum Lloyd's expects this to be at a risk category level, and at a class of business level, where possible;
- identify the reason for any significant divergence between assumptions and results; and
- ensure actuarial and statistical methods are tested against current data and market practice based on a detailed understanding of the model.

The testing of results of the internal model against experience is used to assess the variances between forecasts made by the model and actual outcomes. Where actual outcomes may not be directly available, the model forecasts may be compared to those made on the basis of a comparable data set. Such datasets may include, for example, market level data, data from other parts of the group or data based on other business deemed to have similar characteristics.

This type of test against experience, or “backtesting”, can be used to find various kinds of discrepancies. The objective of the analysis is, for example, to determine whether differences come from omission of material risk factors from the model, whether they arise from errors from other aspects of the model specification such as the dependency structure including the assumptions of linearity, or whether the discrepancies are purely random and thus consistent with acceptable performance of the model.

One way to use backtesting is to test statistically the hypothesis that the observed frequency of exceptions equals the expected frequency, subject to the availability of adequate data. Agents should use backtesting widely, as it offers useful information about the model which is relatively easily understood by directors and other model users.

In general, 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 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.

Agents should justify why the chosen comparable data set is appropriate. The reliability of the test depends on the selection of data used and specific attention to the data selection will increase the comfort agents obtain from the test.

Agents should ensure consistency between modelled output and historical data. For example, agents 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, supported by appropriate evidence and actions taken to remediate the test result or model. Syndicates should not presume the reason for a fail or pass with limitations without having conducted a 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 validation briefings. One particular issue is the requirement to define pass/fail criteria which are detailed and 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.

3.8 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 99.5th 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 can be regarded as having two parts. Firstly, testing whether material drivers of risk are captured in the model and, secondly, whether the calibration of those drivers is appropriate. The second part is similar to testing against experience. P&L attribution can also be used to investigate dependencies between classes or between risk categories.

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;
- explain sources of profit and loss in respect of the major business units of the syndicate;
- explain profit and loss based on the categorisation of risks chosen in the internal model and show how this explains the results by business unit;
- ensure the P&L exercise is performed at least annually, for the previous calendar year;
- ensure it is used in the management of the business and demonstration of use test; and
- ensure it is used in the validation process, at a minimum to validate that the internal model appropriately reflects the risk profile of the syndicate.

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.

Generally, P&L attribution is done in comparison to the relevant model version for the accounting period being reviewed. This is particularly the case for the "testing against experience" element of the P&L attribution, so that testing compares actual business against the relevant plan. For example, it might be carried out at the time of reporting an annual result using the version of the model for which that result was the prospective underwriting year (i.e. the year-end 2022 result would be tested against the 2022 YoA SCR model). This would then be used to inform changes to be made to the latest version of the model. However, any more recent experience that gives insight into the adequacy of model risk coverage should be considered.

If model changes (including data and risk profile changes) affect the P&L 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.

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.

3.9 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.

3.10 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 for 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 with reference to any changes in risk profile. 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.

The Analysis of Change is a key part of validation. It may also help identify significant areas of change in the model or risk profile for which additional validation should be performed. It may also provide information on the extent to which previous validation can still be relied on for the current model.

Lloyd's requires an Analysis of Change document to be submitted with the LCR. Details of the requirements for this submission are given in the [Capital Guidance document](#).

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 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 99.5th 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 higher 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 but could be included in deep dive testing.

Validation of dependency structures should consider the risk profile of the syndicate and discuss, at least on a qualitative basis, what is expected to drive aggregate risk for the syndicate and how this is captured within the dependency modelling. This should include consideration of whether risks are expected to exhibit tail dependency and how this is reflected in the modelling methodology.

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.

Class of business dependency structures within models are often calibrated by allocating class pairs into bands such as background / low / medium / high dependency. Whilst limitations in historical data may make it difficult to validate

individual correlation assumptions, more statistically credible information may be available on the appropriate level for these bands. Sensitivity testing can also be used to assess to which parts of the dependency structure the model is most sensitive so that validation can be prioritise in these areas.

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 coefficients 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.

Dependency modelling may include both drivers and copula-based approaches. Validation should include testing that the structures used are behaving as intended. Where practical, this should include switching off drivers to demonstrate that they have the intended effect, which will also help assess the materiality of the modelled driver.

4.3 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, ADC or 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 of 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.

4.4 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. The validation should include ensuring that the calibration is appropriate to current economic and climate conditions and captures sufficiently

extreme outcomes. Calibration can depend on the source of data, allowance for ENIDs and expert judgements made in the parameterisation process (e.g. length of data window used, relevance to current conditions). Therefore, as with other areas of the modelling, validation should consider the appropriateness of expert judgements and sensitivity to realistic alternative assumptions. Where modelling is being done on an ultimate basis, the validation should also include ensuring that the long-term view is suitable, and features such as mean reversion have appropriate impact.

Agents need to demonstrate their understanding of external models, including any material limitations in these models. Validation should consider how external model limitations affect the internal model and discuss how these are mitigated. The validation should include rationale for why the external model and calibration are fit for the purpose of setting capital.

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.

Validation should also include an explanation to justify modelling of, or mitigation against, the uncertainty of loss estimates from external vendor models as well as consideration of adjustments made by the agent to align an external model to its 'own view of risk'.

4.4.1 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.

4.5 Climate Change

As outlined in Lloyd's ['Principles for Doing Business'](#), view of risk should be appropriate to current conditions, including climate. Lloyd's now requires a clear section in model validations (for climate related perils) which demonstrates that the syndicate's view of risk is appropriate for current climate conditions. Although this requirement arose from consideration of catastrophe models, it is also relevant to other areas of capital modelling where risks are potentially affected by climate conditions.

A forward-looking approach is key to ensuring that current and near-term future risk is being adequately represented. Syndicates should not simply be reconsidering their view of risk appropriateness to current climate in response to historical experience.

All of the risk categories may be impacted by climate change risk. Climate change risk is typically split into physical risk, transition risk and litigation risk. It may be useful to develop a framework to consider how each of these could affect the risk categories within the capital modelling to help validate that their impact is being captured appropriately. Consideration should also be given to how climate change can affect dependencies (e.g. natural catastrophe risk and market risk) and other risks (e.g. demand for metals such as lithium affecting inflation).

Although we expect a large part of validation can be independent of the syndicate's own business profile and, where external models are being used, be completed by the model vendors, both the model validation policy and model validation analysis documentation must make explicit reference to, or contain a section on, current climate. To limit repetition in the latter, this section could reference analysis in other sections of the document, but it must make statements attesting to the representation of current climate in the view of risk, including (where relevant) whether the current climate is suitably represented by model vendors.

Further details are given in Lloyd's report [Thematic Review: Catastrophe Modelling & Climate Change](#).

4.6 Non-Natural Catastrophe Risk

Non-natural catastrophe risks may include new or rapidly evolving risks, which adds to the uncertainty in parameterisation and the reliance on Expert Judgements. Documentation should explain how non-natural catastrophe risks are captured within the model. Where such risks are material, it is particularly important that thorough sensitivity testing and scenario testing is done to understand the impact on the capital modelling, with

rationale for why the selections are appropriate. The modelling should also be backtested against emerging experience and consideration given to how this is evolving over time, with feedback so that the calibration reflects the expected risk profile. The evolving nature of such risks may also mean that it is necessary to use relevant external information or benchmarks to help with parameterisation and validation.

4.7 Events Not In Data

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.

4.8 Special Purpose Arrangements

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

4.9 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.

4.10 Other Areas to Consider

The validation process should consider the outcomes of Lloyd's minimum tests (as outlined in the [Capital Guidance](#)) and in particular tests which have failed and whether the outcome and rationale provided by the capital modelling team is detailed and any planned remediation is appropriate.

Similarly, the validation process should also consider the output and rationale provided in the focus area return (and other relevant capital submission items where appropriate) and test that the capital modelling team have provided enough justification of the results.

5 Appendix A: List of Acronyms and Abbreviations

Acronym	Description
ADC	Adverse development cover
AoC	Analysis of Change
APS	Actuarial professional standards
BEL	Best estimate liabilities
Cat	Catastrophe
CoV	Coefficient of variation
ECA	Economic capital assessment
EIOPA	European Insurance and Occupational Pensions Authority
ELT	Event loss table
ENID	Event not in data
ESG	Economic Scenario Generator
FX	Foreign exchange
GAAP	Generally accepted accounting principles
IBNER	Incurred but not enough reported
ILW	Industry loss warranty
IT	Information technology
JEP	Joint exceedance probability
LCM	Lloyd's catastrophe model
LCR	Lloyd's capital return
LIM	Lloyd's Internal Model
MMC	Major model change
MRC	Market Reserving and Capital
OEP	Occurrence exceedance probability
ORSA	Own risk and solvency assessment
P&L	Profit and loss
PRA	Prudential Regulatory Authority
RDS	Realistic disaster scenario
RI	Reinsurance
RITC	Reinsurance to close
RST	Reverse stress testing
SCR	Solvency capital requirement
SCR1	One-year solvency capital requirement
SII	Solvency II
SPA	Special purpose arrangement

SST	Stress and scenario testing
ST-1	Type 1 sensitivity test
ST-2	Type 2 sensitivity test
ULO	Unincepted Legal Obligations
ULR	Ultimate loss ratio
uSCR	Ultimate solvency capital requirement
VaR	Value at risk
XOL	Excess of loss
YOA	Year of account

6 Appendix B: Changes from Previous Versions

The list below summarises changes in this guidance from the previous version, which was published in January 2022:

- Clarification has been given on the requirement for independence of the validation from the capital modelling team. This includes examples of how agents have demonstrated independence within their processes (section 2.3).
- Use of RDS as part of stress and scenario testing (section 3.5).
- Clarification on the model version to use in P&L attribution (section 3.8).
- Update on the specific areas of validation (section 4) to include Climate Change (section 4.5), Non-Natural Catastrophe Risk (section 4.6) and clarification that the calibration of external models should reflect current and long-term conditions appropriately (section 4.4).
- Updates to other sections where our reviews of validation submissions have indicated areas that may be helpful more widely (e.g. in section 2.10.1 Expert Judgements).