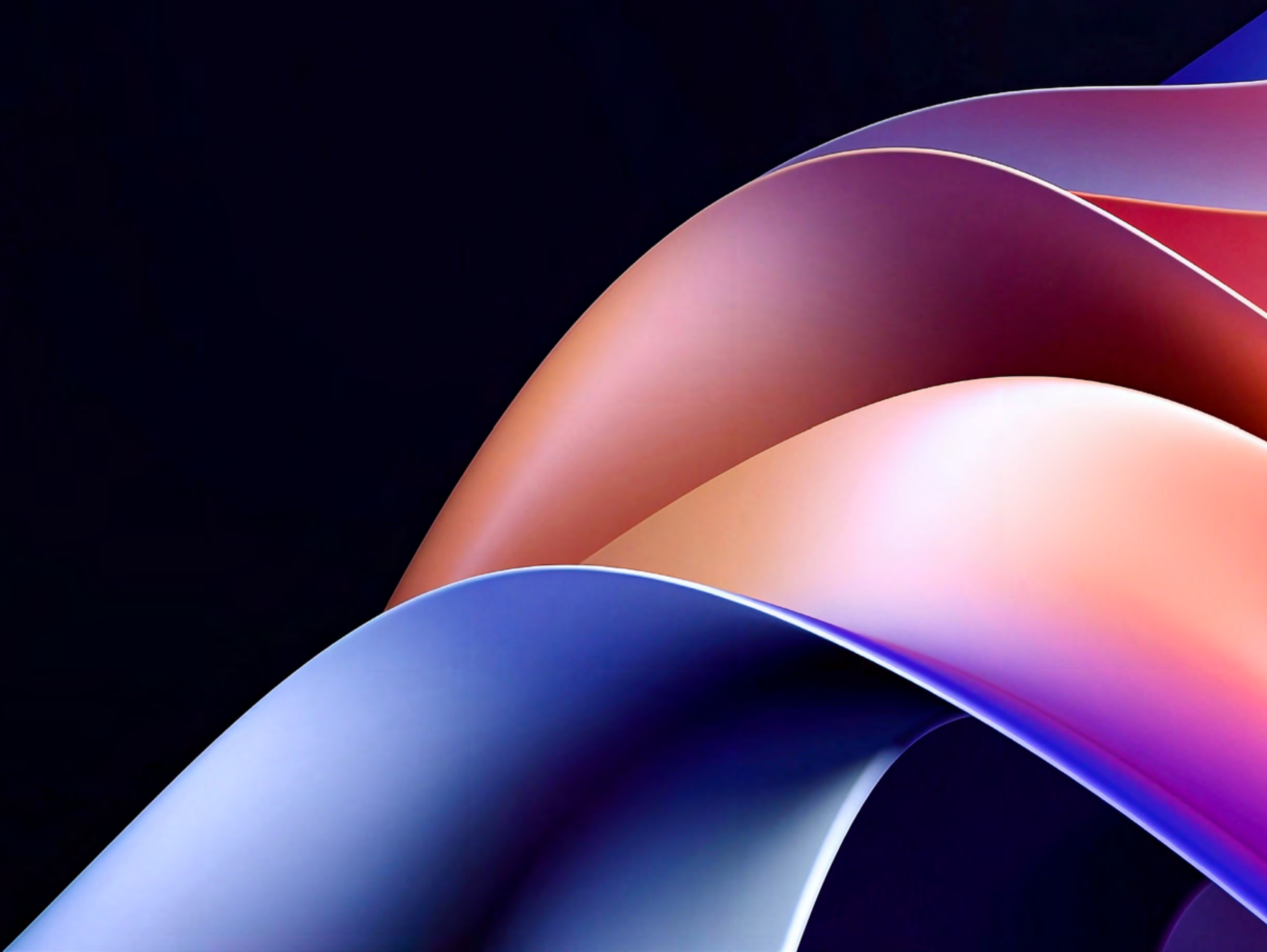
FairPlay's Model Validation Field Guide:

A Practical Handbook for High-Risk Model Review in the Age of AI and Alternative Data



Introduction

Federal guidelines for Model Risk Management (MRM) have always required financial models to meet basic standards for quality and control. But in recent years, regulators and courts have paid extra attention to models that use machine learning, alternative data, or other new methods. These tools can come with unpredictable risks, which means they need to be especially transparent and easy to explain in order to earn trust from regulators and other reviewers.

One of the most important parts of MRM is model validation. Validating a model involves looking at the model from all angles; reviewing how it was built, how it works, and how it is expected to perform. A strong validation process shows that your company is actively checking for problems and fixing them before they cause harm.

This guide—based, in part, on materials from the FDIC's and the OCC's respective Supervisory Guidance on Model Risk Management — is designed to make model validation easier. It is a practical resource for compliance, legal, and data science teams. It includes checklists for each part of the validation process and why each step matters. It can also help teams communicate more clearly and manage stakeholders effectively as they work through the model validation process.

What Is Model Validation and Why Does It Matter?

Model validation is a key part of Model Risk Management. It ensures that a model works the way that it is supposed to, supports the company's goals, and delivers useful results. A strong validation process checks whether the model is reliable, uncovers any hidden assumptions or weaknesses, and looks at how those might affect its performance.

Like any compliance review process, model validation work should be completed by people who not only understand models well but also have the authority and incentives to speak up if they see problems.

Validation should cover every part of the model—from the inputs it uses, to how it processes information, to the results it produces. This applies whether the model was built in-house or provided by a third party. The more important, complex, or wide-reaching the model is, the deeper the validation needs to be.

Lenders should review each model at least once a year, or more often if something changes – for example, if the business strategy shifts, market conditions evolve, regulations are updated, new data sources are introduced, user behavior changes significantly, or early signs of performance degradation begin to appear. That review may confirm that the model is operating as intended, reveal parts of the validation that need updating, or show that a deeper review is needed.
Any substantive changes to a model—like updates to the data, logic, or how the model is used—should prompt a fresh round of validation. Ideally, models should go through a full validation process at regular, scheduled times, and each step should be clearly documented.
Over time, models can lose accuracy or stop performing as expected. Validation helps catch these issues early and sets clear limits for acceptable errors. If a model consistently falls outside those limits, it may be time to rebuild it.

In short, model validation reduces risk. It helps catch mistakes, improve performance, and ensure models are being used properly. It also tests how much you can trust a model by reviewing its logic, methods, and underlying assumptions.

The checklists below provide a guide to assessing the core components of a robust model validation program. These include:

1. Evaluating the model's conceptual soundness and design choices;
2. Reviewing the appropriateness, accuracy, and fairness of the data used in development;
3. Verifying the integrity of the model's implementation, code, and computational logic;
4. Analyzing model performance through outcomes testing, sensitivity analysis, stress testing, benchmarking, and back-testing;
5. Establishing a framework for ongoing monitoring to ensure continued reliability in production; and
6. Assessing governance practices, including policies, change controls, and defined roles and responsibilities.

1 CONCEPTUAL SOUNDNESS OF THE MODEL

Reviewing a model for conceptual soundness is like doing a gut check; it's where you ask, "Does this model make sense, given what we're trying to achieve?" This step ensures that the model's design and development are truly aligned with the company's goals and the purpose behind building the model in the first place.

This part of validation focuses on the big-picture thinking behind the model; its overall design, structure, and logic. It involves reviewing the documentation and the evidence that supports the model's methods and choice of variables.

Strong documentation should clearly explain the model's assumptions and limitations. The validation team's job is to make sure the decisions made during development were thoughtful, research-based, and consistent with best practices in the industry.

Before launching a model—or after making major changes—it's essential to go back and re-check this foundational information. This ensures the model is not only technically sound but also still fits its intended purpose.

To help with this review, it's useful to answer a few core questions. Writing down these answers gives the validation team a solid starting point for evaluating whether the model was designed and built the right way.

Model Design and Construction

When reviewing how the model was designed and built, the validation team should take a close, critical look at every step of its development. Key questions to consider include:

Purpose, Methods, and Judgment

- **What is the model meant to do?**
  Is it clearly designed to serve a specific business goal—and does it actually do that?
- **How was the model built?**
  What methods were used, and why were they chosen?
- **How were variables chosen?**
  Was the selection process thoughtful and backed by evidence?
- **Who developed the model?**
  Were there checks and balances in place during development to reduce bias and mistakes?
- **Is there solid documentation?**
  Does the model have clear, high-quality, regulator-ready documentation and evidence to support its design and the choice of variables?
- **Was human judgment used wisely?**
  What qualitative decisions were made when building the model? Were they done thoughtfully and in a structured way?
- **Is the model based on good theory and business sense?**
  Does it reflect sound thinking and align with industry best practices?
- **How was the model tested?**
  What kinds of training and testing were done before using the model?
- **Were other approaches considered?**
  Did the team look at different modeling methods or theories, and do they explain why they
- **Has anything changed?**
  Are there new research findings or industry practices that suggest it's time to re-evaluate

Model Assumptions, Limitations, and Controls

When reviewing a model, it's important to understand what it assumes, where it might fall short, and what controls are in place to ensure it works reliably. Key questions to ask include:

- **What are the model's main assumptions and variables?**
  What does the model rely on to work properly?

- **How do those assumptions affect its limitations?**
  Could they lead to weaknesses or blind spots in how the model performs?

- **What are the known limitations of the model?**
  Where does it perform less reliably, and in what situations might it not work well?

- **What's being done to manage those limitations?**
  Are there fixes or safeguards in place to reduce the impact of those weaknesses?

- **Are there limits on how the model should be used?**
  Are there clear boundaries on when and how the model can be trusted?

- **Does the model perform well across different scenarios?**
  Is it stable and reliable, even when inputs or assumptions vary?

- **What testing and controls are in place?**
  Has the model been trained and tested well, and are there systems to make sure it continues to work over time?

- **Who developed the model?**
  Were there checks and balances in place during development to reduce bias and mistakes?

- **Are there any signs of bias in the data or results?**
  If so, what steps have been taken to detect and reduce those biases?

- **Does the model need stress testing or sensitivity analysis?**
  If yes, has that testing been done properly to check how the model reacts under pressure or unusual conditions?

2 DATA USED TO BUILD THE MODEL

A model is only as good as the data it's trained on. If the data are flawed—because of bias, mistakes, or poor quality—the model will learn and repeat those same shortcomings.

That's why it's so important to carefully review the data used to build the model. The data should be:

- **Accurate and complete**
- **Free from bias and errors**
- **Reflective of the real world—** your customers, their behavior, and current market conditions

This review is especially important when the model uses third-party data or is being applied to new products or different use cases. Poor data at the start can lead to bad decisions later on. The following questions can guide the validation team in assessing the suitability and completeness of the data used in model development:

Appropriateness and suitability of the Data

- What internal and external data sources were used to build the model?
- Why are these data consistent with the model purpose and design?
- Are these data reasonably representative of your portfolio and market conditions?
- Were the data tested for bias and consistency with intended model subjects?
- How were missing values treated?

- Are any model features proxies for membership in a protected class?
- What kinds of quality checks were run on the data?
- Why is the timeframe of data used for development appropriate?
- Are any data transformations or adjustments used in the model? If so, what are they?
- Were any proxy data used to build the model? If so, which?

Accuracy and Completeness of the Data

- How were internal and external data reviewed for any potential errors?
- How were internal and external data inspected for gaps or missing information?
- How were the data reconciled between source systems and the model?
- Is the development data set replicable by an independent party? If so, was that done?
- Did you conduct any data transformations? If so, how did you assure yourself they were performed correctly?

3 PROCESS VERIFICATION

Process verification is a core pillar of model validation. While conceptual soundness ensures a model is well-designed and outcome analysis assesses how it performs, process verification focuses on how faithfully the model has been built, implemented, and controlled in practice.

This part of the validation ensures that the model's code, computational logic, and implementation accurately reflect its intended design and theoretical foundations. It also confirms that the model produces reliable outputs from given inputs and that it is being used in the intended manner under controlled conditions.

Validation teams must independently verify not only that the model was coded and implemented correctly, but also that appropriate safeguards are in place to prevent unauthorized changes, misuse, or errors in execution. For highly automated or complex models, process verification provides critical assurance that the model works as expected not just in theory—but in practice, at scale, and over time.

The process verification review includes three key areas:

- Development and implementation of model code
- Testing of the model's computational and mathematical integrity
- Controls over model deployment and use

Each question is designed to help ensure that the model has been accurately translated from concept to code, that it is functioning properly in production environments, and that it is being managed with the necessary controls to support its continued reliability and regulatory compliance.

How was the code to develop and implement the model reviewed to ensure that it is correct?

- Were the reviewers independent from the model developers?

What kind of review of the model's computational engine and mathematical applications was completed?

- How did you verify that mathematical theories or numerical techniques were performed correctly?
- Was the model independently replicated to ensure that it can be recreated?
- How did you verify that the model's processing components successfully transform inputs into appropriate outputs?

What kind of controls are in place to govern the model's implementation and use?

- How are you assured the model is appropriately implemented?
- Are there controls to ensure the model is being used as intended?
- How do you verify that all model components are functioning as designed?
- How do you ensure that code cannot be changed without approval?
- How are you tracking which changes were made, when, and by whom?

4 OUTCOMES ANALYSIS

Outcomes analysis is the process of comparing a model's predictions to what actually happened. It helps determine whether the model is accurate and stable.
This step is essential for catching problems early—before they turn into business, legal, or regulatory risks. If issues are found, outcomes analysis can guide updates to the model, the addition of new controls, or other changes to improve reliability.
The type of testing used depends on the model's purpose. For example, you might check:

- How accurately the model forecasts results
- How well it ranks outcomes
- How stable it is when inputs change

The goal is to understand if the model is working and, if not, why not. This may involve statistical tools, but expert judgment also plays a role.

Outcomes analysis should be tailored to the type of model and how it's used. The following questions can help a validation team evaluate whether this testing has been done well:

Assessing model outputs and reporting

- How do you evaluate model outputs and determine whether they are reasonable?
- How do you assess model outputs for accuracy and completeness?

Model Performance Testing

- **Sensitivity Analysis:** Did you do any sensitivity analysis, for example by varying inputs one-by-one or simultaneously? If so, did outputs fall within expected ranges? Could deviations from expected results be reasonably explained?

- **Stress Testing:** Did you do any stress-testing, for example by checking model performance over a wide range (including extreme values) of input and parameter values? If so, did you identify any boundaries for the acceptable range of inputs? Did you identify any conditions under which the model may become unstable or inaccurate? Could deviations from expected results be reasonably explained?

- **Benchmarking:** What kind of benchmarking was performed during model development? Did you build any challenger/alternative models using alternative approaches? Did you compare your outputs to peer institutions, historical experience, or prior model versions? Could deviations from benchmark models be reasonably explained?

- **Back-Testing:** What kind of back-testing was performed during model development? Did you compare model forecasts to actual outcomes? Are models with long forecast horizons back-tested and supplemented by evaluation over shorter periods? If back-testing outcomes fall outside performance thresholds, how do you analyze the discrepancies and investigate the causes that are significant in terms of magnitude and frequency to determine the source of the difference? How do you incorporate back-testing results into the model development, use, and risk management? How are back-testing results communicated and models recalibrated based on back-testing results?

- **Fair Lending Testing:** Did you do fair lending testing? Which decisions were tested for fairness? Did you discover disparities for any protected groups? If so, what variables drove disparities? Were disparities, if any, reasonably related to a legitimate business necessity? How was that analysis documented?

- **Additional Testing:** Did you perform testing based on the model's limitations and assumptions? Did you perform any additional quantitative and qualitative tests or analytical techniques based on the model's methodology, complexity, data availability and the magnitude of potential model risk to your institution? How did you analyze the impact of key assumptions and choice of variables on model outputs?

5 MONITORING

Ongoing monitoring is a vital part of model validation. It helps ensure that the model is working properly in production and that its outputs remain reliable over time. Monitoring also confirms that performance targets, thresholds, and other controls are effectively managing risk.

Monitoring isn't a one-time task; it requires regular checks to confirm that the model still fits your products, customers, data, and market conditions.

If any of these elements change—like launching a new product, serving a different customer segment, or responding to shifting market trends—you may need to update, retrain, or even replace the model. From day one, there should be a clear plan for tracking performance continuously. That includes setting benchmarks, running process checks, and setting up alerts when things drift. The frequency of monitoring should reflect how risky the model is and how quickly your data environment changes.

The following questions can help assess whether your monitoring approach is thorough, well-documented, and responsive to change:

- What is the plan for ongoing monitoring of the model?
- Do you monitor the population subject to the model for consistency with development data?
- Do you monitor outcomes for consistency with model predictions?
- What policies and procedures assure your monitoring plan is followed?
- What are your monitoring activities and associated monitoring thresholds? Did you consider alternative monitoring measures/metrics? If so, why did you disregard them?

6 GOVERNANCE

Governance doesn't just happen at the moment of model creation; it must span the full model lifecycle.

Strong model governance is the foundation for ensuring that models are developed, used, and maintained in a responsible, compliant, and well-controlled manner. Governance ensures that the right people, policies, and processes are in place to oversee how models are developed, deployed, maintained, and updated. It also ensures that changes to models are made responsibly and that key decisions are well-documented, traceable, and aligned with business and regulatory expectations.

This includes setting clear roles and responsibilities, managing access; implementing and documenting change control processes, maintaining written policies and procedures, and planning for future model updates or redevelopment. A well-governed model governance program supports accountability, reduces operational risk, and reinforces trust in model-driven decisionmaking.

The questions below are designed to help assess whether appropriate governance practices are in place for high-risk models, especially those using advanced techniques like machine learning or alternative data.

Compliance policies and procedures

- How do you assure appropriate and adequate model governance and oversight?
- Do you have policies and procedures for operating, maintaining, and updating the model?

Evaluate access and change controls

- Which parties have access to the model?
- Which parties can make changes to the model?
- What are the procedural steps to change the model?
- Are changes to the model logged and auditable?

Plans for model management

- What is the model's usage horizon? Why is that horizon appropriate?
- Are there established thresholds or periods for redevelopment or revalidation?
- Do you have plans for future updates to the model?
- Are the roles and responsibilities for staff involved in model management defined?
- How do you decide who will use, maintain, and update the model?

Putting It All Together

Model validation is a cornerstone of effective Model Risk Management. For high-risk models—particularly those powered by machine learning or alternative data—validation provides a critical check for whether models are robust, reliable, and aligned with both business goals and regulatory expectations.

The checklists in this Field Guide offer a practical, structured approach to assessing key aspects of model risk: conceptual soundness, data quality, process integrity, performance outcomes, monitoring, and governance. Taken together, they form a comprehensive framework to help your institution identify weaknesses, document strengths, and build defensible practices. But validation isn't just about following a list; it's about applying judgment, ensuring accountability, and fostering a culture of rigor and transparency.

Importantly, model validation is not a one-size-fits-all exercise. The depth and scope of your review should reflect the model's complexity, novelty, and risk. Effective validation must also evolve as business strategies shift, new data sources are adopted, or regulatory standards change. When done correctly, validation becomes more than a regulatory obligation; it becomes a competitive advantage. It builds internal confidence, supports faster and safer innovation, and safeguards customers by ensuring that models treat people fairly and work as intended.

By embedding these practices into your workflows, you create a sustainable model validation process that scales with your institution's growth and complexity.

FairPlay can help. Our platform automates many of the statistical, fairness, and documentation tasks described in this Field Guide. By streamlining these steps, FairPlay empowers compliance, legal, and data science teams to validate more models in less time—without compromising on quality or rigor. If you're looking to strengthen your validation program, improve transparency, and reduce manual effort, request a demo see how FairPlay can support your goals.

Model validation is just one piece of the broader Model Risk Management ecosystem, which also includes responsible development, thoughtful implementation, and strong governance. For additional regulatory guidance, consult the FDIC's Supervisory Guidance on Model Risk Management and the OCC's Comptroller's Handbook. With the right tools and practices in place, your organization can turn model validation from a compliance exercise into a driver of long-term value, trust, and impact.

fairplay

To learn more or get in touch, email us at **info@fairplay.ai**