Navigating fairness in insurance

Toward responsible AI adoption

The better the question. The better the answer.
The better the world works.

EY
Shape the future
with confidence

RGA

Table of contents

1. Background

Advancements in modeling techniques within insurance underwriting and pricing have led to increased use of external consumer data and information sources (ECDIS), machine learning (ML) techniques and artificial intelligence systems (AIS). The use of ECDIS and AIS can provide benefits both to companies and consumers, simplifying insurance underwriting and pricing processes, leading to (potentially) more accurate pricing of risk. However, the use of nontraditional data sources and complex methodologies in predictive modeling can codify unfair bias. To mitigate this risk while enabling business strategy and customer value, organizations need a well-defined process for model bias testing to enable the responsible adoption of AI.

As of June 2025, the Colorado Division of Insurance, the National Association of Insurance Commissioners (NAIC) and the New York Department of Financial Services (NYDFS) have all proposed or posted guidance on the governance and risk of ECDIS and AIS. This paper utilizes the NYDFS definition of ECDIS as "lifestyle indicators" to supplement or proxy traditional underwriting factors (e.g., credit scores, biometric data, insurance risk scores), while an AIS is any machine-based system of reasoning, learning or self-improvement that is used fully or partially to supplement traditional health, life, property or casualty underwriting or pricing. While industry practices in the insurance sector are evolving in response, decades of established norms around credit and employment offer a useful indication of the likely trajectory moving forward.

RGA has long had a model governance and model bias testing framework in place. However, as data models have evolved, so, too, has the guidance provided by regulators around risk management and model governance as it pertains to insurers' use of AI and big data. To provide greater insight and transparency into these regulations and their implications, RGA collaborated with Ernst & Young LLP to review and update our model bias testing and governance framework. Through a comprehensive consultative process, the joint effort produced the RGA Model Bias Testing Playbook (the Playbook).

Informed by industry practices and regulatory guidance, the Playbook serves as a framework that can be applied to test for bias in any model for use in both individual and group markets. RGA's Model Risk Management team has adopted the Playbook to guide testing of RGA's models and is delighted to share it with clients to help responsibly advance the industry's use of AI and big data.

The remainder of this paper outlines the key elements contemplated in the Playbook, from establishing a fairness framework through testing, monitoring and reporting considerations.

2. Establishing a fairness framework

Effective risk management starts with strong governance, including clear roles and responsibilities across owners and users of models, developers, independent testing, and legal and compliance. Scalability requires defining risk assessment criteria and corresponding control expectations. In the context of bias testing, organizations should consider defining criteria for the risk tiering of models and AI for bias to focus efforts where inherent risk is the highest and outline expectations to measure and monitor risks accordingly.

In model lifecycle roles, four key considerations are:

1. **Model user:** defining guidelines for appropriate model usage, including business justification and effective review and challenge
2. **Model developer:** assessing fairness tests when selecting a model methodology or outcome
3. **Independent tester:** establishing a review process inclusive of conducting a fairness prioritization assessment, qualitative review and quantitative testing; establishing ongoing monitoring for emerging biases
4. **Legal/compliance officer:** developing a risk governance framework to adhere to guidance and assist in defining and identifying the use of ECDIS and AIS in predictive modeling; establishing transparent reporting protocols

2.1 Fairness prioritization

By implementing a fairness prioritization assessment, organizations can evaluate whether underwriting models with traditional variables pose greater risk than marketing models with ECDIS, and whether automated decisioning models without ECDIS should take priority over those with ECDIS and human oversight. This process allows organizations to define variable relationships and key risk factors, supporting a standardized procedure for determining testing scope, priority and rigor.

Organizations often deploy numerous predictive models, each with varying levels of risk and impact. For some models, a qualitative review can be used to assess and justify variables within the model specification based on their relationship to demographics. In other cases, quantitative testing may be needed in addition to applying fairness tests to determine whether a model is deemed biased per an established threshold and testing objective. Testing limitations, such as whether outcomes are rankable, meaning there are objectively advantageous vs. disadvantageous outcomes, can also indicate whether a quantitative test is meaningful. A clear prioritization framework is crucial to enable, rather than strangle, innovation.

Three key risk considerations are:

1. **Process risk:** The impact of how the model is used can lead to increased risk from customer vulnerability or underwriting or pricing decision-making processes.
2. **Model risk:** The complexity of the model can impact the clarity of variables that contribute to a pricing or underwriting outcome.
3. **Variable risk:** The relationship between input variables and outcomes can lead to inherent risk included within the model.

3. Performing bias testing

3.1 Applying proxy methodologies

In many contexts, including insurance underwriting and pricing, sensitive race and ethnicity information is not collected. Thus, methods to proxy race and ethnicity are essential to test for bias across these attributes. Proxy methods referenced among regulators include Bayesian Improved Surname and Geocoding (BISG), which is a methodology developed by RAND that can help US organizations produce accurate, cost-effective estimates of racial and ethnic disparities within data sets — and illuminate areas for improvement. The implementation of these methods requires careful consideration of the data landscape and the limitations of geolocation proxy. When applying the BISG method, organizations should consider how to use the proxy in bias testing and at what granularity to represent proxies based on model design, availability of data and model use.

3.2 Creating a fairness metric decision framework

Some regulators, such as NYDFS, require an assessment to determine whether the use of ECDIS or AIS produces disproportionate adverse effects in underwriting or pricing for similarly situated insureds across demographics. However, there is not one universal definition of how to calculate fairness, and further challenges arise as some metrics used to quantify fairness may provide contradictory conclusions.

A metric decision framework selects a metric that is appropriate for the given model and provides rationale as to why a particular test was used within the testing objectives. Structured quantitative testing procedures should be developed based on data availability, model outcomes and fairness objectives.

1. **Data availability** can affect statistical requirements for accurate calculation.
2. **Model outcomes** can determine whether a categorical or continuous (e.g., scoring models) fairness metric is more suitable.
3. **Fairness objectives,** such as whether a fairness metric draws conclusions from practical or statistical significance, can determine whether a hypothesis test or measure of effect size is utilized. To illustrate the dimensionality of metric selection between model outcomes and objectives, example choices could contemplate a Z-test for statistical significance but vary the assessment of practical significance between a standardized mean difference for continuous models or an adverse impact ratio for categorical outputs.

3.3 Less discriminatory alternatives

If a model is found to be biased, clear procedures and documentation are required to investigate less discriminatory alternatives (LDAs) and assess the trade-off between decreased business performance and reduced bias. LDAs utilize alternative modeling approaches to reduce model bias below the metric threshold. Often, when using a drivers of disparity analysis, this can mean removing variables or variable combinations that could be influencing observed disparity. The use of an effective search for LDAs allows for the continuation of business-as-usual practices with decreased risk of bias using a clear trigger, procedure and stopping point.

3.3.1 What triggers a search for LDAs?

To efficiently search for alternatives within the highest risk areas, clear procedures should identify when a search for LDAs is necessary for a particular model (e.g., if meaningful disparities are observed between protected and control groups). If candidates are not produced automatically during model development, the process may be triggered by a breach, leading to an alternative methodology model selection process to find a suitable alternative.

3.3.2 How to conduct a search for LDAs

When conducting a search, a definitive methodology to search for LDAs promotes consistency in the derivation of candidate models, can help streamline processes, and defines the roles of fairness testers and model developers to limit risk during this process. This can incorporate a drivers of disparity analysis in which variables influencing observed bias are removed. Using this method, the same model development process can be used with a reduced variable pool. Alternatively, methods such as adversarial de-biasing utilize race and ethnicity information to guide modifications to the original model until it can no longer be used to predict race.

3.3.3 Defining a reasonable search for alternatives

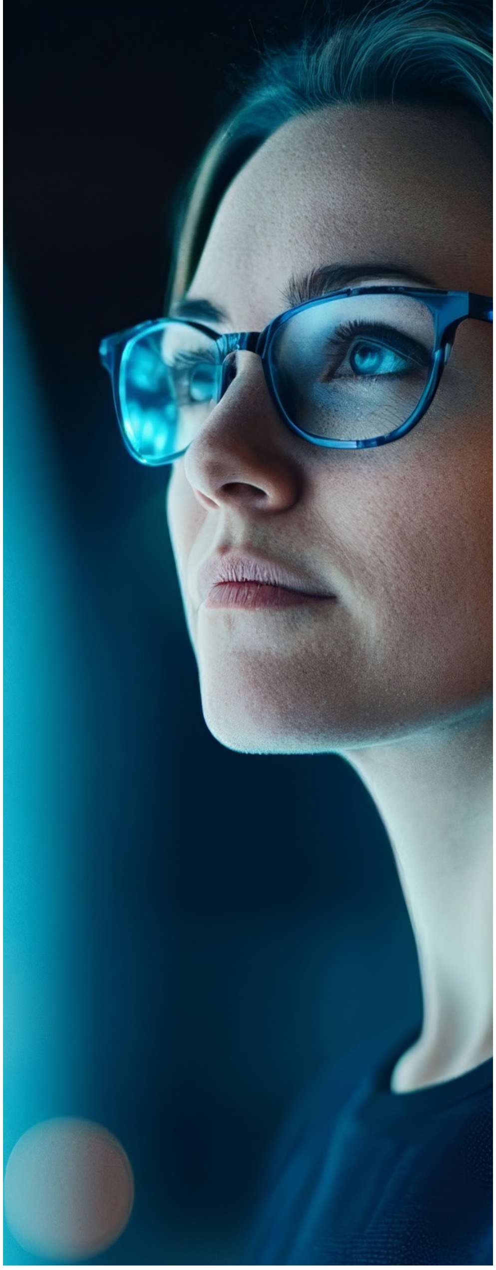
Based on the chosen method, clear procedures should identify how to efficiently end the search for LDAs to enable a timely process. Quantitative approaches can consider a set number of candidates, while outcome-driven approaches are only complete when a fair model has been identified.

3.4 Identifying drivers of disparity

Performing a drivers of disparity analysis leverages additional approaches (regression based and outcome based) in the creation of LDAs through incorporation of racial indicators to identify potential associations between key input variables and demographics. However, this approach necessitates a thoughtful evaluation of the use of sensitive racial data and ethical ramifications, verifying that their use does not inadvertently reinforce stereotypes or biases.

3.5 Business justification

In instances where a model exhibits measurable bias or reduced performance, its continued use may be justified when it serves a legitimate business purpose that cannot be met through less biased or more accurate alternatives. Such justification must be supported by a clear demonstration that the impact on demographics has been thoroughly assessed, that any contributing variables are essential to achieving the intended business outcome, and that no reasonably effective, fairer alternative exists. This decision is made transparently, with documented trade-offs between fairness and performance, and is accompanied by a robust monitoring plan to facilitate ongoing compliance and mitigation of potential disparate impacts over time.



3.6 Monitoring and reporting

As data updates occur, shifts in the relationships between variables and the outcomes associated with demographics may occur. This necessitates a robust monitoring framework so that models previously assessed for bias maintain their fair performance over time. Regular retesting is important to identify any degradation in fairness that may arise from changes in either production or development data.

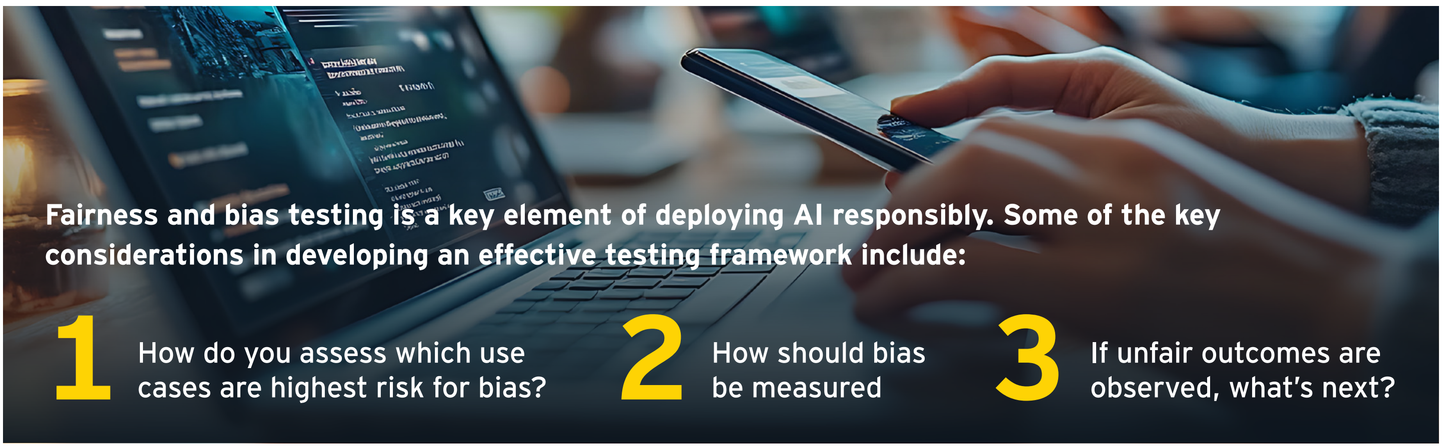
Key considerations to determine when to retest models for bias include:

1. **Model updates:** Any modifications to the model structure, calibration data or model specification warrant a comprehensive retesting to evaluate the impact on fairness conclusions.
2. **Data drifts in ongoing model use:** If the population from which the model was trained and tested differs from the current population on which the model is applied, new biases may be introduced from ongoing model use. In light of this, annual testing may be needed to continually assess whether the model remains fair in the event of structural changes. This aligns with the annual requirements for outcome-based evaluations, reinforcing the commitment to fairness in model performance.

To enable effective governance across the organization, results of these tests must be effectively communicated to business and risk governance forums with accountability over business unit and enterprise-wide risk management. Additionally, high-risk use cases and the use of ECDIS variables may be required in certain jurisdictions.

Comprehensive documentation is a critical component of the monitoring and reporting process. For reporting purposes, it is essential to assess whether the model contains ECDIS variables and to determine what specific information should be inventoried for comprehensive reporting. This facilitates a clearer understanding of the model's performance and its implications for fairness.

4. Key takeaways



The views reflected in this article are the views of the authors and do not necessarily reflect the views of RGA, Ernst & Young LLP or other members of the global EY organization.

About RGA

Reinsurance Group of America, Incorporated (NYSE: RGA) is a global industry leader specializing in life and health reinsurance and financial solutions that help clients effectively manage risk and optimize capital. Founded in 1973, RGA is one of the world's largest and most respected reinsurers and remains guided by a powerful purpose: to make financial protection accessible to all. As a global capabilities and solutions leader, RGA empowers partners through bold innovation, relentless execution, and dedicated client focus — all directed toward creating sustainable long-term value. RGA has approximately $4.1 trillion of life reinsurance in force and assets of $133.5 billion as of June 30, 2025. To learn more about RGA and its businesses, please visit rgare.com or follow RGA on LinkedIn and Facebook. Investors can learn more at investor.rgare.com.

Authors

Brian Clark
Managing Director
Ernst & Young LLP

Benjamin Pickett
Senior Manager
Ernst & Young LLP

Cole Stevenson
Senior Consultant
Ernst & Young LLP

EY | Building a better working world

EY is building a better working world by creating new value for clients, people, society and the planet, while building trust in capital markets.

Enabled by data, AI and advanced technology, EY teams help clients shape the future with confidence and develop answers for the most pressing issues of today and tomorrow.

EY teams work across a full spectrum of services in assurance, consulting, tax, strategy and transactions. Fueled by sector insights, a globally connected, multidisciplinary network and diverse ecosystem partners, EY teams can provide services in more than 150 countries and territories.

All in to shape the future with confidence.

SCORE no. 28935-251US
2508-12044-CS
ED None

ey.com