Unequal Odds: Addressing and Mitigating Bias in Life Insurance

kpmg.com

The integration of Artificial Intelligence (AI) and Big Data within the life, health, and annuity (Life) insurance industries marked a pivotal shift towards data-driven decision-making processes in underwriting, claims processing, pricing strategies, and risk assessments. This evolution, while heralding increased operational efficiencies and enhanced predictive accuracies, has concurrently spotlighted the critical issue of model bias. With the actuarial profession at the forefront of adopting innovative modeling approaches, there emerges an increased responsibility to navigate the ethical considerations inherent in their deployment.

Additionally, recent regulatory developments in the United States, and globally, have signaled a growing focus on addressing and mitigating model bias within the life insurance industry. This increased regulatory scrutiny aims to ensure that the deployment of algorithmic models and the use of big data and AI do not lead to unfair discrimination against various population segments.

This white paper seeks to dissect the multifaceted nature of model bias within the Life insurance domain, underscoring the imperative for actuaries to champion ethical standards in the development and application of these advanced analytical models.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

01 Introduction to Model Bias in Life Insurance: A Deep Dive into Fundamental Mechanisms

Model bias occurs when predictive algorithms yield outcomes that systematically favor or discriminate against certain groups based on flawed machine learning processes, erroneous assumptions, or biased data. The ramifications within the Life insurance sectors are profound, influencing key decisions around policy pricing, underwriting and risk evaluation, and eligibility with potential discriminatory impacts. This segment delves into the theoretical underpinnings of model bias, categorizing its various manifestations and illustrating its occurrence through sector-specific scenarios. By dissecting instances where biases have inadvertently been encoded into predictive models, we aim to illuminate the pathways through which such biases perpetuate societal disparities, thereby challenging the actuarial profession to critically evaluate and refine their analytical methodologies.

Approaches to assessing and mitigating model bias: Methodologies and frameworks

The detection and rectification of model bias necessitate a comprehensive evaluation encompassing scrutiny of input data, algorithmic fairness analysis, and outcome assessment. This section delineates advanced methodologies involved in identifying disparate impacts including sensitivity analysis for determining the impact of input variable variations on model outputs; the deployment of fairness metrics (e.g., demographic parity, equalized odds) for scrutinizing the equity of model decisions; and methods for summarizing the overall impact like odds ratios and impact ratios. We further explore innovative mitigation strategies, including the utilization of fairness-centric machine learning algorithms, augmenting training datasets for greater inclusivity, and instituting rigorous auditing protocols to ensure models are continuously aligned with fairness objectives. Through a technical examination of these methodologies, we aim to furnish actuaries with a toolkit for enhancing the fairness quotient of their predictive models.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

Technical methods for assessing model bias

Fairness Metrics: As described above, a strong model bias governance structure requires a series of interrelated statistical tools to help mitigate disparate impacts to key stakeholders. The first step is typically the implementation of fairness statistics. These statistics provide insight into the presence of bias for individual variables. Further, these statistics provide insights into different fairness principles and should be used in accordance with a Company's fairness philosophy. For example, a Company with the goal of equalizing opportunity would rely on different fairness statistics than one focused on equalizing outcomes. Examples include:

- **Demographic Parity:** Requires that the decision rate (e.g., acceptance for insurance) should be the same across groups. It's assessed using difference in means tests.
- **Equal Opportunity/Equality of Odds:** Requires that the true positive rate (or false positive rate) is the same across groups. This can be evaluated using logistic regression analysis, where the protected attribute and the prediction outcome are modeled, and the interaction term's significance is analyzed.

Predictive Parity: Requires that predictive values (positive and negative) are equal across groups. Confusion matrices for each group are compared to assess this metric. These statistics can be applied in isolation to evaluate the fairness of individual variables, as well as in combination with other statistics and methodologies to determine whether bias exists under a variety of perspectives and scenarios.

Sensitivity Analysis: Once fairness statistics and philosophy are settled, techniques like Monte Carlo simulations and bootstrapping are employed to understand how changes in model inputs affect outputs. This can reveal input variables and data elements that disproportionately influence the model predictions for certain groups, suggesting potential sources of bias. This may reveal situations where a model is unbiased in many scenarios but produces disparate impacts in tail events like medical access and outcomes during the Covid 19 pandemic.

Summary Statistics: Once models are adjusted for fairness statistics and sensitivity analysis, it can be helpful to summarize the overall results for a given model in a single statistic that summarizes any disparate impact from a given model. These approaches typically rely on traditional statistical techniques (e.g. confusion matrices) with tailoring to evaluate disparate impact. Examples include:

- **Impact Ratio Calculation:** this is the ratio of favorable outcomes for a protected group to that of the reference group. For instance, if 80% of male applicants are hired and 60% of female applicants are hired, the impact ratio for women is 60% / 80% = 75%.
- **Confusion Matrix Statistics:** In the context of model bias, we adjust the confusion matrix statistics to evaluate whether a risk from the protected group was treated fairly in relation to the reference group. Examples include:
  - ***Accuracy:*** provides a sense of how often the model is correct by calculating the ratio of correctly predicted outcomes (both true positives and true negatives) to the total number of instances for both the protected and reference groups.
  - ***Precision:*** provides a sense of well the model predicts positive outcomes by calculating the ratio of true positive predictions to all positive predictions (true and false positives) for both the protected and reference groups.
  - ***Specificity:*** provides a sense of well the model predicts negative outcomes by calculating ratio of true negative predictions to all actual negative instances for both the protected and reference groups.

Splitting the above confusion matrix statistics into protected and reference group statistics allows us to evaluate if the model is performing equally well for protected groups and the general population. The importance given to these statistics will depend on the insurance product and process step. For example, specificity may be the key ratio for underwriting decisions, while precision and accuracy will be more revealing in pricing scenarios.

[1] Mheidly, Nour et al. Emerging Health Disparities during the COVID-19 Pandemic. Avicenna Journal of Medicine, 2022, https://pmc.ncbi.nlm.nih.gov/articles/PMC10038746/#:~:text=Health%20inequity%20increased%20the%20risk,medical%20conditions%20among%20underserved%20populations

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

Technical methods for mitigating model bias

Pre-processing Techniques: Once the approach for assessing model bias is selected, the next step is to remediate bias throughout the modeling lifecycle. Pre-processing techniques like re-weighting the training data to reduce dependence on protected attributes or using adversarial de-biasing where a model is trained simultaneously with an adversary model that tries to predict the protected attribute from the main model's predictions, thereby minimizing the information about the protected attribute in the predictions.

Another advanced pre-processing adjustment is application of variational autoencoders (VAEs) to reduce bias in underlying data sets. VAEs accomplish this by learning representations of the data that are invariant to the protected attributes, ensuring that downstream tasks using these representations do not perpetuate or amplify existing biases.

In-processing Methods: The next step in the modeling lifecycle is model building. Incorporating fairness constraints directly into the model training process helps prevent biased results upfront. For instance, constrained optimization techniques can be used where traditional loss functions are modified to include terms that penalize unfairness measures, forcing the model to learn fair representations.

In addition, developing models that provide insights into how predictions are made can help identify and correct bias. Techniques like SHAP (Shapley Additive explanations) and LIME (Local Interpretable Model-agnostic Explanations) can elucidate how different features contribute to model predictions, highlighting potential sources of bias.

Post-processing Approaches: While preventing bias during the modeling process is ideal, many completed models will require remediation as model bias regulations go into effect. This is accomplished by adjusting model outputs to ensure fairness criteria are met. One method is equalized odds post-processing, where model thresholds are adjusted for different groups to equalize true positive and false positive rates across groups.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

02 Recent Regulatory Developments: Navigating the Evolving Legal Landscape

In response to the expanding use of AI and Big Data in insurance, regulatory entities worldwide have begun formulating directives aimed at safeguarding against unfair discrimination. In the United States, state regulators such as New York, California, and Colorado are leading the charge with regulatory guidance focused on model bias in insurance. The commonalities of these regulatory requirements touch on implications of model transparency, fairness, and accountability in automated decision-making processes within the insurance sector. Additionally, these regulations have direct impact on actuarial and data science practices, highlighting strategies for compliance and the integration of legal considerations into model development and management protocols.

Regulatory developments in the US

The past few years have seen significant transformations in U.S. model bias regulations. As advancements in big data, modeling techniques, and artificial intelligence continue to evolve, state regulators are progressively shifting the onus onto corporations to safeguard fairness and equity in their models. We at KPMG are actively tracking the continual evolution and development of State-specific model bias regulatory requirements, a subset of these laws and regulations is outlined below.

California:

- **California Consumer Privacy Act (CCPA):** CCPA provides consumers with rights to access, delete, and opt-out of the sale of their personal data, which impacts how life insurers use personal data in models. Algorithms used by insurers must comply with these privacy regulations, which indirectly influence model bias.
- **Proposed Algorithmic Accountability Regulations:** California is considering regulations that may require insurers to demonstrate that their models do not lead to unfair discrimination based on race, gender, and other protected classes.

New York:

- **New York Department of Financial Services (NYDFS):** NYDFS Circular Letter No.1 requires life insurers to avoid using external data sources and models in underwriting unless they can demonstrate that the models do not disproportionately impact protected groups. Insurers must provide transparency and ensure fairness when using algorithms.
- **Annual Reporting Requirements:** NYDFS requires insurers to report how they use external consumer data and algorithms, promoting accountability and minimizing the risk of biased outcomes.

Illinois:

- **Illinois Biometric Information Privacy Act (BIPA):** BIPA regulates the use of biometric data, such as facial recognition and fingerprinting, which some life insurers use in underwriting. Insurers must ensure that the use of this data does not result in biased decisions against any individual or group.
- **Anti-Discrimination Laws:** Illinois prohibits discrimination based on protected categories, such as race, gender, and national origin, ensuring that algorithms used by life insurers do not reinforce these biases.

Colorado:

- **Insurance Fairness Laws:** Colorado prohibits life insurers from discriminating based on protected classes, similar to other states. This extends to the use of data-driven models in pricing and underwriting.
- **Consumer Protection in Interaction with AI Bill:** A bill that prohibits algorithmic discrimination against individuals based on their actual or perceived age, skin color, disability, ethnicity, genetic information, sex and other classifications.
- **Data Privacy Law (CPA):** Colorado's data privacy law impacts life insurance companies by requiring that personal data be handled responsibly. This law indirectly influences how insurers design and monitor algorithms for fairness, since biased models could violate privacy laws.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

03 Examples of Model Bias in Insurance and Actuarial Models: Case Studies

In discussing model bias, it is important to translate theoretical concerns into tangible consequences by presenting a series of case studies and empirical research findings that underscore the real-world implications of model bias within the insurance industry. Intuitive examples include predictive models in health insurance that may disproportionately underpredict healthcare costs for minority populations, leading to inadequate coverage plans. Alternatively, life insurance algorithms may inadvertently impose different premium rates based on non-risk-related personal characteristics. Other examples extend to actuarial models employed in assessing risk and determining annuity payouts, highlighting instances where biases could skew assessments, disproportionately affecting certain demographic groups. Through an in-depth analysis of these scenarios actuaries and data scientists can develop a nuanced understanding of the complexities involved in modeling and the pivotal role of ongoing research and regulatory development in mitigating bias.

Product-specific examples of model bias

Model bias in Life insurance models can significantly affect the fairness and equity of insurance products. The following are product-specific examples illustrating how model bias can manifest and potentially impact policyholders:

1. Life Insurance Underwriting Models:

Gender Bias in Premium Calculation Life insurance models might inadvertently incorporate gender bias, leading to unfair premium rates for women or men. Historically, female policyholders have had longer life expectancies, which could be used to justify lower premiums. However, such a straightforward application can overlook the nuances of modern lifestyle changes, medical advancements, and the narrowing gender gap in life expectancy. An overreliance on outdated actuarial tables without adjusting for these trends could unfairly disadvantage one gender.

2. Health Insurance Pricing Models:

Socioeconomic Status and Zip Code Bias Health insurers use models to set premiums or decide on coverage levels. If these models heavily weigh the zip code or socioeconomic status, it could lead to higher premiums or reduced coverage for individuals from poorer or historically marginalized communities, who might already be at a disadvantage in terms of accessing healthcare services. This form of bias exacerbates health inequalities instead of mitigating them.

3. Long-term Care Insurance:

Bias Against Pre-existing Conditions Models used for underwriting long-term care insurance might unduly penalize applicants with certain pre-existing conditions by assuming a direct and high correlation with future long-term care needs. This can manifest as significantly higher premiums or denial of coverage. If the model does not accurately account for the variability in disease progression, treatment effects, and individual health practices, it could unjustly impact those with manageable conditions.

4. Annuity Products:

Socioeconomic Bias in Mortality Projections Annuities are often priced based on mortality projections. If these models inaccurately project longer life spans for individuals of higher socioeconomic status due to access to better healthcare, nutrition, and living conditions, it could result in less favorable annuity rates for lower-income individuals. This could inadvertently reinforce socioeconomic disparities.

For each of these examples, it's crucial for actuaries and model developers to employ rigorous bias detection and mitigation strategies, such as regular model audits, sensitivity analyses, and the incorporation of fairness-aware algorithms. Moreover, engaging with a diverse range of stakeholders can provide valuable insights into the real-world impacts of these models, guiding more equitable model development and application.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

04 How to Get Started: Implementing a Bias Mitigation Framework

Embarking on the journey to address model bias requires a structured approach, beginning with the formation of a cross-disciplinary team that brings together expertise from actuarial science, data science, ethics, legal and compliance. Structuring a roadmap for institutions seeking to initiate or enhance their efforts in combating model bias is not an easy task. Life insurers can benefit from lessons learned in other industries in developing their roadmaps, such as property & casualty insurance and banking. These roadmaps should be developed using a risk-based approach and emphasize the importance of conducting thorough audits of existing models to identify and quantify biases, implementing a robust framework for continuous assessment of fairness and accuracy, and ensuring transparency in model operations to all stakeholders. Strategies for fostering an organizational culture that prioritizes ethical considerations in model development alongside recommendations for leveraging industry standards and engaging in collaborative initiatives to advance fairness in actuarial practices should be discussed broadly with the model bias steering committees and senior leadership within the organization.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

Lessons learned from other industries

The life insurance industry can garner significant insights from the property and casualty (P&C) insurance and banking industries regarding assessing and mitigating model bias. These sectors have historically contended with similar challenges and, over time, have developed robust frameworks and innovative approaches to manage bias effectively.

From Property and Casualty Insurance:

1. Granular Data Utilization: P&C insurance has long harnessed granular data for risk assessment, pricing, and fraud detection. Life insurance can adopt similar strategies to enhance model accuracy and fairness. By incorporating more detailed data points, models can better capture the nuances of risk, potentially reducing bias associated with broader categorical variables.

2. Geographic Risk Assessment Techniques: P&C insurers often use geographic information systems (GIS) and spatial analysis for risk assessment (e.g., flood or earthquake risk). Applying similar spatial analysis in life insurance, for example, assessing health risks, can uncover and adjust for geographic biases that broader models might miss.

3. Use of External Data Sources: The P&C sector frequently integrates external data sources to improve risk modeling accuracy. Life insurers could similarly benefit from leveraging a wider array of data sources, including socioeconomic and environmental data, while carefully assessing these for inherent biases.

From Banking Industry:

1. Adverse Action Notices: In banking, particularly in credit scoring, lenders are required to provide reasons for adverse actions (e.g., loan denial) to applicants. Life insurers could adopt a similar approach, offering transparency around underwriting decisions. This could encourage the development of more interpretable models, as insurers would need to understand and communicate how decisions are made, potentially reducing the reliance on "black-box" models.

2. Fair Lending Practices and Regulation Compliance: The banking industry is subject to stringent regulations designed to prevent discrimination, such as the Fair Housing Act and Equal Credit Opportunity Act. Life insurance companies can learn from how banks have navigated compliance with these regulations, including the implementation of fairness analyses and bias mitigation strategies in their models.

3. Advanced Analytical Techniques for Bias Detection: Banks have employed advanced analytical techniques to detect and mitigate bias in lending models. Techniques like counterfactual fairness, which assesses a model's decisions by simulating alternatives where sensitive attributes are varied, can provide a robust framework for life insurers to assess model fairness.

Cross-Industry Strategies for Mitigating Model Bias:

1. Regular Model Audits: Both P&C insurance and banking perform regular audits of their models to check for accuracy and fairness. Life insurers could adopt continuous monitoring and regular external audits of their models to identify and address bias promptly.

2. Fairness-Aware Modeling: Both sectors are exploring fairness-aware algorithms that adjust for bias during the model training process. Life insurers can similarly incorporate fairness constraints or objectives into their modeling processes to ensure more equitable outcomes.

3. Stakeholder Engagement: Engaging with diverse stakeholders, including consumers, regulators, and advocacy groups, can provide valuable perspectives on fairness and bias. This practice, common in banking, can help life insurers identify blind spots in their models and develop more inclusive products.

In summary, by learning from the experiences and strategies of the P&C insurance and banking industries, the life insurance sector can enhance its approach to assessing and mitigating model bias, leading to fairer and more equitable outcomes for policyholders.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

05 In Summary: A Call to Action

Life insurers are at the beginning of their model bias journey. Companies should take proactive stances towards integrating considerations of fairness into the actuarial and risk model development lifecycle. Initial milestones on the model bias roadmap include establishing of a comprehensive governance framework that encompass rigorous bias assessment and mitigation practices, development of consensus-based standards for model fairness, and promotion of open dialogue among leadership and modelers with industry peers, regulators, and technology providers. Much research is underway focusing on further exploration in this domain, highlighting areas such as the development of advanced fairness metrics, exploration of new mitigation techniques, and the study of the long-term societal impacts of model bias. Through a concerted industry-wide effort, we envision a path forward wherein actuaries not only comply with regulatory mandates but also uphold the highest ethical standards, ensuring equitable treatment for all policyholders and sustaining public trust in insurance practices.

06 KPMG. Make the Difference.

At KPMG, our approach to addressing model bias is designed to provide you with robust, actionable insights and enhancements. Firstly, we will review your current modeling processes and bias controls, collaborating with your team to evaluate overall process performance and identify areas for improvement. Our goal is to recommend practical enhancements that effectively mitigate bias. Beyond this review, we will conduct independent testing of your models. This rigorous evaluation aims to identify any existing biases and reinforce your processes with enhanced controls, providing you greater assurance regarding the fairness of your models. If unfair discrimination is identified, our team can employ various techniques, such as adjusting algorithmic decision thresholds, rebalancing or resampling training data, or applying fairness-aware algorithms to develop unbiased predictions.

Upon completing our assessment, KPMG will provide a thorough Model Bias Assessment Report detailing our findings and recommendations. This report will cover the appropriateness of current processes, identified biases and remediation actions, and suggestions for process and control enhancements. Additionally, we will deliver detailed analysis results from the model re-execution tests, including training data modifications, evaluated model bias metrics, and remediation results. By leveraging our expertise, you can benefit from a meticulous assessment and enhancement of your model bias controls, ensuring fairness, compliance, and improved decision-making processes.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

07 Appendix: Model Bias Remediation Use Cases

In the appendix, we detail three of the most common areas of model bias in the Life insurance industry. This includes a discussion of the source of bias, its impact on consumers and companies, techniques for remediation, and lessons learned for each use case:

Socioeconomic Status in Health Insurance Pricing Models

Description	Health insurance companies often rely on data about applicants' health and lifestyle behaviors, such as smoking habits, exercise, and diet to price new policies. However, socioeconomic factors beyond applicants' control can significantly affect access to healthcare, diet, and overall health outcomes, leading to biases in risk prediction models.
Practical Impact	In some cases, individuals from lower socioeconomic backgrounds may be categorized as higher risk due to health conditions that are exacerbated by poverty (e.g. obesity, hypertension, and diabetes). These individuals may have to pay higher premiums, even though the health conditions may be less related to personal choices and more related to environmental and economic factors. Conversely, individuals with higher socioeconomic status may be underpriced due to access to better healthcare and preventive services.
Technical Detail	**Data Dependencies:** Actuarial models often use socioeconomic status (SES) as a proxy for health risks, because SES affects access to healthcare and the likelihood of chronic conditions. Common indicators of SES include **income, credit score, education level, occupation,** and **living conditions.** **Bias in Risk Assessment:** SES data can be biased in several ways: • **Underrepresentation:** Individuals with low SES may not access healthcare, leading to underreported conditions (e.g., undiagnosed hypertension or diabetes). • **Overgeneralization:** Using broad SES categories as risk proxies can misclassify healthier individuals (with low SES) as higher risk.
Mitigating Bias	**Adjusting for SES:** Actuaries must carefully weigh SES factors based on empirical research. For example, using propensity score matching (PSM) techniques to match individuals with similar health risks but differing SES. **Statistical Models:** Instead of SES as a direct factor, actuaries may use variables like hospital admissions, medical tests, and lifestyle surveys to refine health risk predictions.
Statistical Techniques	**Multilevel Regression:** This technique can account for hierarchical data (e.g., individuals nested within neighborhoods or social groups) to adjust for the broader socioeconomic environment. **Instrumental Variables (IV):** If direct SES data is biased (e.g., due to underreporting), IV methods can be used to find uncorrelated proxies that are more representative of an individual's true risk. **Machine Learning: Random Forests or Gradient Boosting Machines (GBM)** can help identify interactions between SES and other risk factors, uncovering non-linear relationships between SES and health outcomes.
Lessons Learned	Pricing actuaries must ensure that models do not disproportionately penalize individuals based on factors beyond their control. Alternative data sources, such as more precise health metrics or lifestyle tracking, may be needed to refine models and prevent the use of proxies that unfairly penalize individuals from disadvantaged backgrounds.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

Smoking Status and Life Insurance Data Bias

Description	Life insurance models often use smoking status to classify risk, with smokers typically facing higher premiums due to the increased risk of chronic diseases. However, biases in the way smoking data is reported can distort model accuracy.
Practical Impact	Individuals in certain demographic groups (e.g., low-income communities or specific cultural groups) may underreport their smoking status due to social stigma, fear of discrimination, or lack of awareness about the importance of reporting smoking habits. This underreporting can lead to inaccurate risk assessments. Some policyholders may be misclassified as non-smokers, resulting in an underestimation of their actual health risks.
Technical Detail	**Self-reporting Bias:** Smoking status is often self-reported in life insurance applications, and this data is subject to underreporting or inaccuracies. Individuals may be reluctant to disclose their smoking habits due to stigma. **Health Impact of Smoking:** The impact of smoking on health risk is well-documented, with smokers having a higher probability of developing diseases like cancer, cardiovascular diseases, and respiratory disorders. However, misclassification due to inaccurate self-reporting can lead to inaccurate pricing and risk assessments.
Mitigating Bias	**Verification with Biomarkers:** Insurers are increasingly exploring the use of biomarkers or health monitoring devices (e.g., nicotine tests, wearable devices tracking health metrics) to verify smoking status more accurately. **Predictive Modeling:** Machine learning models can use longitudinal health data (e.g., from electronic health records) to more accurately model the risk associated with smoking, adjusting for other risk factors like BMI, exercise, and family history
Statistical Techniques	**Latent Class Analysis (LCA):** This technique can be used to identify subgroups of individuals who may be underreporting smoking but share similar health risk profiles. **Bayesian Inference:** Bayesian models can be used to update beliefs about an individual's smoking status based on prior knowledge (e.g., national smoking rates) and evidence (e.g., health metrics or biomarkers).
Lessons Learned	Insurers need to implement more accurate data collection methods and may consider using alternative data sources, such as health monitoring devices or smoking cessation programs, to better assess risk without relying solely on self-reported data.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. USCS025485-1A

Algorithmic Bias in AI/ML Automated Underwriting Models

Description	The advent of machine learning (ML) and artificial intelligence (AI) in underwriting has introduced new challenges in model fairness. These algorithms are often trained on historical data, which may contain inherent biases. If the training data reflects societal biases (e.g., health disparities or discrimination), the resulting models may perpetuate or even amplify these biases.
Practical Impact	A case of algorithmic bias occurred when AI-based underwriting models were found to disproportionately deny coverage or charge higher premiums to minority groups, even when controlling for traditional risk factors such as health, age, and lifestyle. For instance, models might implicitly learn to associate certain zip codes (often correlated with race or ethnicity) with higher risk, leading to biased decision-making.
Technical Detail	**Training Data:** Machine learning algorithms in automated underwriting often train on historical claims data, which may encode biases from previous decisions. For example, if the training data overrepresents certain demographics in claims or underwriting decisions, the model will likely perpetuate those biases. **Bias Amplification:** Models like **neural networks** or **random forests** can "learn" patterns that may inadvertently favor one group over another. If the training data is historically skewed (e.g., underrepresentation of certain demographic groups), the model may **overfit** to these imbalances.
Mitigating Bias	**Re-sampling techniques** (like SMOTE—Synthetic Minority Over-sampling Technique) can help balance underrepresented groups in the training data. **Fairness Metrics:** Implement fairness metrics like equalized odds (ensuring equal false positive rates for all demographic groups) and predictive parity (ensuring similar predictive accuracy across groups). **Explainability:** Using explainable AI (XAI) methods, such as LIME (Local Interpretable Model-agnostic Explanations), allows insurers to audit decisions made by automated underwriting models, ensuring transparency and identifying any potential biases.
Statistical Techniques	**Fairness Regularization:** In machine learning models, adding fairness constraints directly to the objective function (e.g., using **adversarial debiasing**) can reduce bias in the predictions by penalizing decisions that deviate from fairness. **Gradient Boosting with Fairness Constraints:** Popular algorithms like **XGBoost** or **LightGBM** can be modified to account for fairness by incorporating constraints that adjust for bias during training.
Lessons Learned	It's crucial for insurers using AI and machine learning to apply fairness audits and mitigation techniques such as fairness constraints, re-weighting, or bias correction to their models. Transparent, explainable AI models are also essential to ensure that biases can be identified and corrected before they affect customers.

Some or all of the services described herein may not be permissible for KPMG audit clients and their affiliates or related entities.

Learn about us: **kpmg.com**

The information contained herein is of a general nature and is not intended to address the circumstances of any particular individual or entity. Although we endeavor to provide accurate and timely information, there can be no guarantee that such information is accurate as of the date it is received or that it will continue to be accurate in the future. No one should act upon such information without appropriate professional advice after a thorough examination of the particular situation.

© 2025 KPMG LLP, a Delaware limited liability partnership and a member firm of the KPMG global organization of independent member firms affiliated with KPMG International Limited, a private English company limited by guarantee. All rights reserved. The KPMG name and logo are trademarks used under license by the independent member firms of the KPMG global organization. USCS025485-1A

Contact us

Alex Zaidlin
Principal, Actuarial
E: azaidlin@kpmg.com

Nate Loughin
Managing Director, Actuarial
E: nloughin@kpmg.com