Journal of Organizational Behavior Research
Vol.: 10, Is.: 4, Year: 2025, Pages: 71-84
https://doi.org/10.51847/2DgUHKt7Sk
2528-9705
Available online
https://odad.org/

Calibrated Turnover Risk Models for Capacity Constrained Retention Triage

Sarawut Laksawut[1], Warawut Narkbunnum[2]*

[1]Program in Business Administration and Digital Innovation, Mahasarakham Business School, Mahasarakham University, Thailand.
[2]Department of Mahasarakham Business School, Mahasarakham University, Thailand.

***Corresponding Author**
E-mail: warawut.n@acc.msu.ac.th

ABSTRACT

Employee turnover remains a critical challenge for organizations because it erodes human and social capital and generates substantial replacement costs. While the literature highlights job attitudes, social exchange, and organizational justice as key antecedents of turnover intention, HR departments still lack calibrated, decision-oriented tools for targeting scarce retention resources. This study develops and evaluates calibrated turnover risk models to support capacity-constrained retention triage. Using an anonymised HR dataset of employees in a large service organization, we estimate logistic regression and random forest models and then apply probability calibration, Brier score decomposition, and expected calibration error to assess probability reliability. We combine discrimination and calibration results with decision curve analysis under realistic capacity constraints and translate alternative operating points into workload and outcome metrics, such as alerts per day and true-positive and false-positive burdens. The findings show that a calibrated logistic model preserves high discrimination while substantially improving probability calibration and delivering higher net benefit than uncalibrated alternatives across plausible capacity ranges. Scenario analysis demonstrates how HR managers can select thresholds that align with available staff time and acceptable error trade-offs. The study contributes a calibration-first pipeline for turnover risk modelling and provides actionable guidance on integrating HR analytics into retention decisions in a transparent, capacity-sensitive manner.

Keywords: Employee turnover, HR analytics, Probability calibration, Decision curve analysis, Retention management, Algorithmic HRM/fairness.

Introduction

Employee turnover remains a key concern in organizational behavior because it incurs high costs through lost human capital, replacement expenses, and disruption of work groups. Evidence from service-oriented sectors shows that turnover intentions are strongly influenced by job satisfaction, commitment, job stress, and motivational states, especially in difficult-to-replace customer-facing roles. Recent research highlights how satisfaction, motivation, and stress together predict intentions to leave and organizational commitment, underscoring the need for proactive, perception-based retention strategies (Yakupoğlu, 2025). The Job Demands-Resources framework connects these factors to outcomes by describing jobs in terms of demands that deplete energy and resources, and those that promote motivation, engagement, and performance. Empirical studies demonstrate that job stress and strain often mediate the relationships among job characteristics, satisfaction, and turnover intentions, and that careful management of demands and resources can help reduce quitting intentions (Yakupoğlu, 2025).

Meanwhile, the rapid growth of HR analytics and machine learning has opened new opportunities to turn these insights into predictive tools: large-scale HR information systems now support models that estimate individual turnover risk based on historical patterns in attitudes, performance, workload, compensation, and contextual factors. Systematic

Recieved: 01.08.2025 –Accepted: 07.11.2025 –Published: 15.12.2025

reviews document the use of over 20 supervised learning techniques, with ensemble methods often achieving the highest accuracy. Key predictors include job satisfaction, overtime, pay conditions, as well as competencies, commitment, trust, and cultural values, which emerge as critical features for case-specific risk profiles and employee segments (Avrahami *et al*., 2022; Akasheh *et al*., 2024; Shafie *et al*., 2024; Talebi *et al*., 2025).

Despite this progress, significant gaps remain at the intersection of HR analytics and organizational behavior. Reviews indicate that turnover prediction studies are highly focused on accuracy, emphasizing metrics such as accuracy, F1 score, and area under the ROC curve, while considering probability calibration or decision analytic value (Akasheh *et al*., 2024; Talebi *et al*., 2025). Poorly calibrated models, which systematically overestimate or underestimate actual turnover rates, can distort managers' interpretation of high-risk employees and misallocate scarce retention resources. Current work also rarely explains how predicted probabilities should be converted into actionable thresholds when organizations cannot intervene with all at-risk employees, even though retention strategies like coaching, job redesign, training, and targeted pay adjustments are costly and must be prioritized within budgetary and staffing constraints. Evidence from other applied fields shows that good discrimination is necessary but not sufficient; accurate probabilities and explicit consideration of consequences are equally important. Decision Curve Analysis (DCA) provides a simple way to evaluate models based on net benefit by balancing true positives and false positives across thresholds without requiring detailed cost data. Nonetheless, DCA and related decision analytic tools are still rarely used in HR analytics and organizational behavior research.

This study addresses these gaps by integrating JD-R theory and modern HR analytics within a calibration-first, decision-focused framework for employee retention triage. Conceptually, predicted turnover risk is viewed as a summary indicator of underlying demands, resources, and exchange relationships, consistent with JD-R and attitudinal models of withdrawal (Yakupoğlu, 2025). Methodologically, we develop and compare supervised learning models for voluntary turnover, explicitly assess both discrimination and calibration, and apply post hoc calibration when needed to enhance probability reliability. We then incorporate these calibrated models into a decision-analytic evaluation using DCA and workload-based translations to quantify net benefit relative to treat-all and treat-none strategies under realistic capacity constraints.

The Study Explores Three Research Questions.

1) Among commonly used machine learning models for turnover prediction, which configurations deliver the best combined performance in discrimination and calibration when applied to historical HR data?
2) Given capacity constraints on retention strategies, which operating thresholds on calibrated turnover risk maximize decision analytic net benefit, and how do these thresholds translate into understandable workload metrics such as alerts per day and the balance between true positive and false positive cases?
3) To what degree does a calibration-first, decision-focused HR analytics pipeline support transparent and theory-consistent retention policies when viewed through the JD-R and attitudinal frameworks on turnover?

By answering these questions, the study makes three contributions to organizational behavior research and practice. First, it strengthens the methodological foundations of HR analytics by emphasizing calibration and decision analytic evaluation that complement accuracy-based metrics with reliability and managerial usefulness. Second, it offers a structured approach for translating predicted turnover probabilities into capacity-constrained retention policies, using DCA and workload metrics understandable to HR practitioners. Third, by integrating model development and evaluation within JD-R and recent JOBR evidence on satisfaction, motivation, and stress (Yakupoğlu, 2025), the study connects advanced analytics to core organizational behavior constructs and supports the use of predictive tools that are both theoretically grounded and practically actionable.

Literature Review

Employee Turnover, Job Attitudes, and Organizational Behavior Perspectives

Employee turnover is a central concern in organisational behavior because it erodes productivity, service quality, and the stock of human and social capital. Classic turnover research identifies job satisfaction, affective commitment, and perceived organisational support as proximal predictors of turnover intention and resignations, consistent with Social

Exchange Theory, which argues that employees reciprocate favourable treatment with loyalty and reduced withdrawal (Meyer & Allen, 1991). Recent studies across business and healthcare contexts show that work engagement, organisational support, and commitment jointly reduce turnover intentions, while job satisfaction and work motivation strengthen commitment and weaken quitting intentions (Sartori *et al*., 2023; Zhu *et al*., 2023; Palma-Moreira *et al*., 2024; Poku *et al*., 2025; Yakupoğlu, 2025). Evidence on ethical leadership and organisational justice further indicates that fair and ethical climates lower turnover intention, partly through enhanced work motivation (Özkan, 2023; Rajopadhye *et al*., 2023; Moon *et al*., 2024; Zhao *et al*., 2024). Meta analytic work grounded in the Job Demands Resources model converges on the conclusion that engagement and access to resources are key protective factors against turnover intention (Mazzetti *et al*., 2023), yet this literature still relies largely on cross sectional survey designs and sample level averages, offering limited guidance on how HR departments can translate engagement, support and justice into individualised, risk based decision tools for allocating scarce retention resources.

Job Demands-Resources, Human Capital, and Retention Risk

The Job Demands-Resources (JD-R) framework explains how working conditions influence burnout, engagement, and turnover intentions by suggesting that high demands like workload, time pressure, and emotional strain drain energy. Conversely, resources such as autonomy, feedback, supervisor support, and development opportunities boost motivation and help buffer these effects (Bakker *et al*., 2023). Research across higher education, workplaces during the pandemic, and public hospitals shows that workload, emotional demands, role overload, and emotional labor increase turnover intentions. Meanwhile, participation in decision-making, recovery opportunities, supportive leadership, professional development, and team cohesion reduce these risks (Nicolls *et al*., 2022; Wong *et al*., 2024; Rehman, 2025). From a human capital perspective, turnover results in the loss of valuable, often organization-specific knowledge and skills, reinforcing the need for evidence-based retention strategies. Recent work on artificial intelligence and strategic HRM shows that organizations increasingly depend on data-driven decision-making to protect and utilize human capital, although reviews of algorithmic HRM highlight ongoing concerns about fairness, transparency, and accountability, calling for responsible, human-centered approaches (Hunkenschroer & Luetge, 2022; Bujold *et al*., 2023; Fenwick *et al*., 2023; Alabdali *et al*., 2024; Dima *et al*., 2024). However, most JD-R and human capital studies still treat turnover intention as a variable to be statistically explained, rather than as a probabilistic measure that can inform risk-based, capacity-limited retention decisions. This underscores the motivation for the present study to focus on predictive modeling and decision-oriented analytics for employee turnover risk.

HR Analytics and Machine Learning for Employee Attrition Prediction

HR analytics and machine learning (ML) are now widely used to forecast employee attrition and support proactive retention strategies. Studies on corporate HR datasets show that tree-based ensembles such as random forests and gradient boosting frequently outperform logistic regression in terms of accuracy and area under the ROC curve (AUC) (Raza *et al*., 2022; García *et al*., 2023; Gazi *et al*., 2024; Iparraguirre-Villanueva *et al*., 2024). Work on the widely used IBM HR Analytics dataset similarly finds that random forest and support vector machine models outperform single decision trees, reinforcing the role of ML for early identification of at-risk employees (Mansor *et al*., 2021; Guerranti & Dimitri, 2022; Sugimori *et al*., 2022).

Systematic and narrative reviews synthesising this literature document extensive experimentation with algorithms ranging from logistic regression and k-nearest neighbours to gradient boosted trees, deep neural networks, and ensembles, typically applied to cross-sectional HR datasets (Akasheh *et al*., 2024; Alqahtani *et al*., 2024; Talebi *et al*., 2025). These assessments note that evaluation practices remain dominated by discrimination metrics such as accuracy, F1 score, and AUC, often on benchmark datasets like the IBM attrition corpus and other public HR repositories with limited organisational context (Alqahtani *et al*., 2024; De Vos *et al*., 2024). Applied work in service and public sectors follows the same pattern: ML models trained on HR data from community mental health organisations and other knowledge-intensive settings identify employees at elevated risk and emphasise algorithm choice, feature selection, and variable importance profiles (Akasheh *et al*., 2024; Alqahtani *et al*., 2024; Wu & Fukui, 2024).

Taken together, these studies show that ML-based attrition models are feasible and often highly discriminative, yet evidence syntheses highlight two critical gaps. First, model evaluation is overwhelmingly accuracy-centric, with relatively few studies examining whether predicted probabilities are well-calibrated to observed attrition, despite the importance of calibration for trustworthy risk communication (Van Calster *et al*., 2019; Akasheh *et al*., 2024). Second, downstream managerial questions receive limited attention, including how many alerts a model would generate under explicit capacity constraints and how thresholds should be chosen to maximise net benefit relative to treat-all or treat-none strategies, issues that can be addressed through decision curve analysis (De Vos *et al*., 2024). Consequently, the link between predictive models and implementable HR triage policies remains underdeveloped.

Probability Calibration and Decision-Analytic Evaluation

Outside HR, probability calibration is recognised as a crucial but often neglected dimension of predictive performance, because poorly calibrated models produce probabilities that deviate from observed event frequencies and can mislead decision makers even when discrimination is acceptable (Van Calster *et al*., 2019). Standard practice, therefore, combines calibration plots with indices such as the Brier score, and post hoc methods including Platt scaling, isotonic regression, and temperature scaling are frequently used to improve probability estimates from complex classifiers. Evidence from clinical and financial domains shows that well-calibrated models support more trustworthy decisions, whereas miscalibrated risks can drive overuse or underuse of interventions (Van Calster *et al*., 2019). In contrast, reviews of machine-learning-based turnover prediction report heavy reliance on accuracy, F1 score, and AUC, with explicit calibration diagnostics rarely reported (Akasheh *et al*., 2024; De Vos *et al*., 2024; Talebi *et al*., 2025). Decision Curve Analysis (DCA) complements calibration by evaluating models in terms of net benefit across threshold probabilities, comparing them with treat-all and treat-none strategies, and linking thresholds to implicit trade-offs between false positives and false negatives (Van Calster *et al*., 2019; Piovani *et al*., 2023). Yet DCA is seldom used in HR analytics, where studies typically do not translate model performance into net benefit curves, benchmark against default policies, or express operating points in workload terms such as alerts per period and the implied true and false positive caseload under capacity constraints (Singh *et al*., 2023; Akasheh *et al*., 2024; Cárdenas López & Tabares Betancur, 2024; Shlash Mohammad *et al*., 2025), leaving limited guidance on how calibrated probabilities and DCA can jointly inform operational retention policies.

Algorithmic HRM, Fairness, and Governance

The rapid diffusion of predictive tools in HR has intensified concerns about fairness, transparency, and responsibility. Algorithmic HRM is framed as a form of digital HR where data-driven systems inform or automate core decisions, creating efficiency gains but also risks of discrimination, opacity, and reduced autonomy (Meijerink *et al*., 2021; Sienkiewicz, 2024; Strohmeier *et al*., 2026). Empirical work shows that these tools can strengthen strategic HR decisions when embedded in clear governance, yet many organisations lack policies for bias audits, documentation, and employee involvement (Hunkenschroer & Luetge, 2022; Bujold *et al*., 2023; Alabdali *et al*., 2024). Studies on employee reactions indicate that perceptions of justice, transparency, and opportunities for human review strongly shape acceptance of algorithmic HR and influence satisfaction, perceived support, and commitment (Moritz *et al*., 2023; Köchling *et al*., 2024; Jabagi *et al*., 2025). Meta-analytic and review evidence suggest that algorithmic decisions are often viewed as less fair and less respectful of autonomy than human decisions, which motivates governance frameworks that emphasise explainability, contestability, and continuous monitoring (Malik *et al*., 2023; Moritz *et al*., 2023).

Summary and Research Gap

Across organisational behavior and HRM, prior studies consistently link job attitudes, JD R mechanisms, and turnover intentions, showing that high demands with few resources elevate burnout and quitting intentions, whereas enriched resources, meaningful work and supportive climates foster engagement, satisfaction and retention (Van Heerden *et al*., 2022; Chen *et al*., 2023; Tang *et al*., 2024). Research, including recent work in service and public sectors, highlights satisfaction, motivation, and engagement as key levers for reducing intention to quit (Vinh, 2023; Chinyamurindi & Mashavira, 2024; Pham, 2024; Yakupoğlu, 2025). In parallel, HR analytics and ML studies demonstrate that attrition can be predicted with high discrimination, yet reviews show that most work remains accuracy-centric and pays limited attention to probability calibration, net benefit, or capacity-aware threshold selection (Kumar *et al*., 2023; Gazi *et al*., 2024). Methodological contributions on calibration and Decision Curve Analysis underline that reliable probabilities and decision analytic evaluation are essential for translating risk scores into decisions that create value (Van Calster *et al*., 2019), while the algorithmic HRM literature warns that poorly governed systems can undermine fairness and trust (Alabdali *et al*., 2024; Bujold *et al*., 2024). Few studies integrate these strands in a single framework that grounds attrition prediction in JD R, human capital, and social exchange, adopts a calibration-first modelling strategy, uses DCA under explicit capacity constraints, and links operating point selection to fairness-oriented governance. The present study addresses this gap by developing calibrated, decision-analytic HR analytics for retention triage.

Materials and Methods

Data, Outcome, and Preprocessing

The analysis uses an internal HR dataset of 1,480 employees from a large service organization, with a binary outcome indicating whether each employee left within the following year (an attrition rate of 16.1 percent, resulting in a moderately imbalanced class distribution). Predictors include demographics (age, marital status, education), employment status (job role, job level, department, contract type), working conditions (monthly income, overtime, distance from home, business travel), and attitudinal indicators such as job satisfaction and work–life balance, allowing job demands, resources, and human capital to be interpreted within JD–R, social exchange, and human capital frameworks. Personally identifiable information was removed, and identifiers were pseudonymized; categorical variables were one-hot encoded, with rare categories merged for stability; continuous variables were inspected for implausible values and winsorized at the 1st and 99th percentiles; and the few missing values were imputed using medians for numerical variables and mode categories for categorical variables. All steps were executed within a fixed preprocessing pipeline, applied consistently to training, validation, and test sets, and designed for straightforward reuse during deployment.

Data Splitting, Models, and Calibration

To obtain unbiased performance estimates and support threshold selection, the data were partitioned into a training set, a validation set, and a held-out test set. Stratified sampling preserved the overall attrition rate across splits. The training and validation data were used for model fitting, hyperparameter tuning, and calibration, whereas the test set was reserved for final performance reporting, decision-curve analysis, and fairness audits. Two baseline classifiers were estimated. A regularised logistic regression model (L2 penalty, class-balanced weights) provides a transparent linear baseline with well-defined coefficients. A random forest model (approximately 400 trees with class-balanced weights) captures non-linear interactions and higher-order effects at the cost of reduced interpretability. Both models output estimated probabilities of attrition. Because raw outputs from complex classifiers are often miscalibrated, we applied a calibration-first strategy: each model was first fitted on the training data, then post-hoc calibration was learned on the validation data using isotonic regression, and the calibrated models were finally evaluated on the test set.

Evaluation Metrics and Decision-Analytic Framework

Model performance was evaluated from three complementary angles. First, discrimination was assessed using the area under the receiver operating characteristic curve (AUC) and average precision (AP). Second, probability calibration was evaluated through the Brier score, expected calibration error, and calibration intercept and slope derived from bin-level reliability diagrams. These metrics quantify how closely predicted probabilities align with observed attrition rates, which is critical when risk estimates trigger interventions. Third, we conducted a decision-analytic evaluation using Decision Curve Analysis. For a range of threshold probabilities, we computed the net benefit of each model relative to two default strategies: intervening on all employees and intervening on none. Thresholds were translated into operating points that specify true positives, false positives, and the implied workload. We report decision curves with bootstrap percentile bands to reflect sampling uncertainty.

Workload, Monetary Evaluation, Fairness, and Governance

Workload, monetary evaluation, fairness, and governance were assessed together. Threshold probabilities were converted into expected alerts per day based on a notional volume of 1,000 employees screened and a capacity limit of 450 cases that can reasonably receive retention support; for each model and threshold combination, we calculated utilization as the ratio of alerts to capacity and focused on operating points where utilization does not exceed one. Monetary outcomes were evaluated using a net monetary benefit framework that assigns benefits to true positives and costs to false positives and to intervention activities, yielding a daily net monetary benefit and a benefit-to-cost ratio. Sensitivity analyses were conducted under alternative assumptions regarding the intervention, opportunity costs, and retention benefits. A subgroup fairness audit on the held-out test set examined discrimination, calibration, and confusion matrix metrics across age, gender, salary band, business travel, overtime, and job role, highlighting groups with low AUC, high calibration error, or extreme error rates at the chosen policy threshold. The complete workflow, including data transformations, model versions, calibration procedures, and governance triggers for recalibration and incident review, is packaged in a reproducible bundle for ongoing monitoring.

Results and Discussion

Attrition Patterns in the Sample

Overall attrition in the sample was moderate, but exit rates varied significantly across roles and conditions. **Figure 1** highlights the categories with the largest deviations from the mean: sales representatives, employees aged 18–25, and those with overtime contracts or frequent business travel had attrition rates close to 0.30–0.40, while managers, research directors, and higher salary bands showed much lower exit rates. Higher attrition was also seen among single employees and staff with technical degrees or employed in human resources. These patterns support the Job Demands-Resources perspective, which suggests that intense demands and limited resources tend to cluster within specific groups, thereby justifying targeted predictive triage **(Figure 1)**.

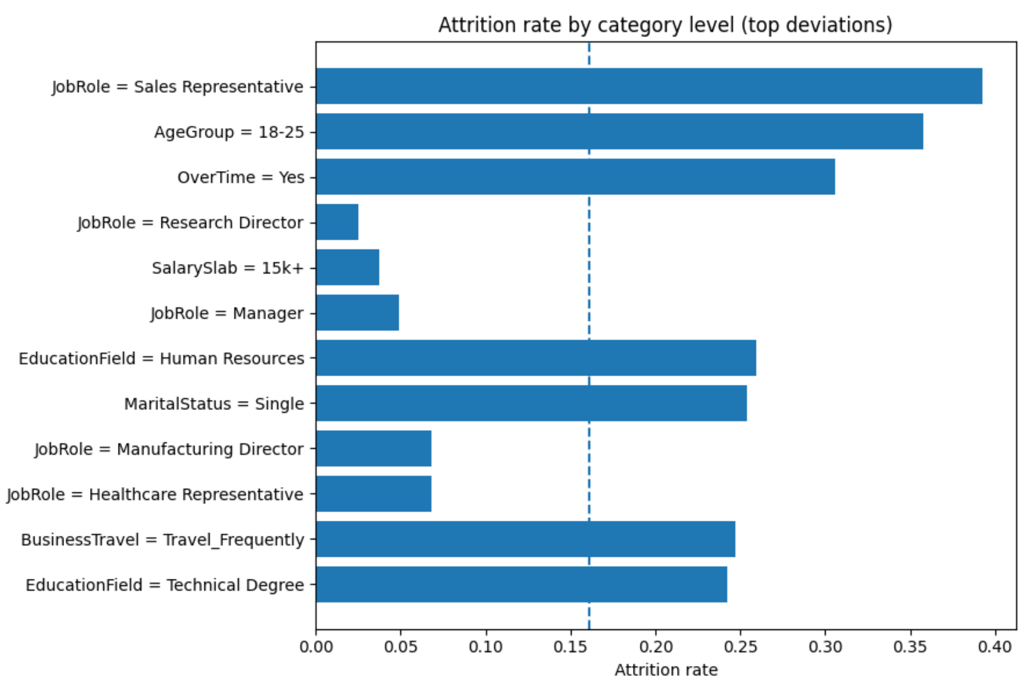

Figure 1. Attrition rate by key categories

Model Performance and Probability Calibration

Table 1 presents discrimination and calibration metrics for the four candidate models on the held-out test set. The uncalibrated logistic regression achieved the highest discrimination (AUC 0.885, AP 0.706), followed by the random forest (AUC 0.846, AP 0.586). However, both models were poorly calibrated, with Brier scores of 0.138 and 0.102, ECE values of 0.179 and 0.071, and substantial calibration bias.

Table 1. Model performance

model	logit	logit_calibrated	rf	rf_calibrated
AUC	0.885	0.881	0.846	0.829
AP	0.706	0.636	0.586	0.533
Brier	0.138	0.081	0.102	0.098
ECE	0.179	0.033	0.071	0.026
reliability	0.057	0.002	0.007	0.001
resolution	0.055	0.052	0.038	0.038
uncertainty	0.136	0.136	0.136	0.136
Calibration slope	0.954	1.104	1.581	0.183
Calibration in the large	-1.655	-0.063	0.966	-1.251

Post hoc isotonic calibration significantly improved probability accuracy for both models. For logistic regression, the Brier score decreased from 0.138 to 0.081, and ECE decreased from 0.179 to 0.033, while calibration in the large moved from -1.655 to -0.063, and the slope moved closer to 1. For the random forest, the Brier score dropped from 0.102 to 0.098, and the ECE from 0.071 to 0.026, reducing extreme overconfidence in the original slope. These improvements came with only slight reductions in discrimination (AUCs of 0.881 for calibrated logistic regression and 0.829 for calibrated random forests). The reliability plot in **Figure 2** shows that, after calibration, observed attrition rates closely follow mean predicted probabilities across the risk spectrum, supporting their use in capacity-constrained decision analysis **(Figure 2)**.

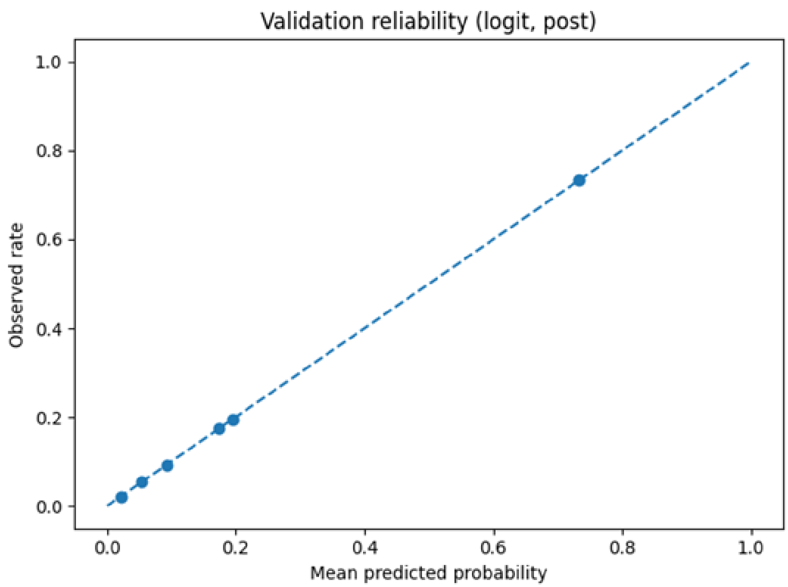

Figure 2. Reliability diagram for the deployed model on the validation set (logit, post-calibration)

Decision-Curve Analysis

Decision-curve analysis was used to assess whether deploying calibrated models would confer a net benefit compared with strategies that either intervene on all employees or on none. **Figure 3** presents decision curves for the calibrated logistic regression and random forest across the threshold range from 0.05 to 0.15, which corresponds to plausible intervention policies in the organisation. Throughout this window, both calibrated models achieved higher net benefits than the treat-all and treat-none strategies, indicating that model-based prioritisation improves the balance between correctly retaining high-risk employees and unnecessary interventions on low-risk staff **(Figure 3)**.

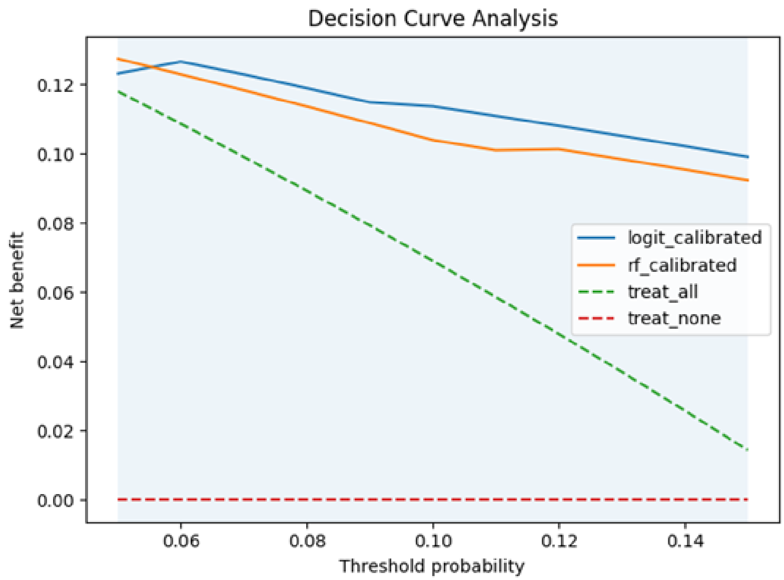

Figure 3. Decision-curve analysis for calibrated models versus treat-all and treat-non.

Table 2 reports the net benefit at selected thresholds relevant to managers. For the calibrated logistic regression, net benefit peaked at lower thresholds of 0.05-0.07 but remained favourable at 0.10 and higher. For example, at a threshold of 0.10, the logistic model attained an average net benefit of 0.114 with a conservative lower confidence band of 0.075, while the calibrated random forest achieved a slightly lower net benefit of 0.104 with a lower band of 0.064. These results suggest that a calibrated logistic model operating at thresholds around 0.10 can offer a robust improvement over a non-model baseline.

Table 2. Policy thresholds summary

model	threshold	NB_mean	NB_lower	alerts_per_day	utilisation	NMB_per_day	ROI	Pass_all
logit_ calibrated	0.05	0.123	0.083	920.27	2.05	9456.1	2.541	FALSE
	0.1	0.114	0.075	433.78	0.96	12614.9	5.362	TRUE
	0.12	0.108	0.069	433.78	0.96	12614.9	5.362	TRUE
	0.15	0.099	0.06	433.78	0.96	12614.9	5.362	TRUE
rf_ calibrated	0.05	0.127	0.088	660.81	1.47	11422.3	3.593	FALSE
	0.1	0.104	0.064	660.81	1.47	11422.3	3.593	FALSE
	0.12	0.101	0.063	425.68	0.95	11858.1	5.179	TRUE
	0.15	0.092	0.053	425.68	0.95	11858.1	5.179	TRUE

Workload and Monetary Implications under Capacity Constraints

To translate decision-curve results into operational terms, we computed the expected daily workload and monetary returns for each policy threshold, assuming a capacity of 450 cases per day. **Table 2** combines net benefit estimates with workload translation and cost parameters. At a low threshold of 0.05, the calibrated logistic model would generate approximately 920 alerts per day, exceeding available capacity by more than twice, and yield moderate monetary gains. In contrast, thresholds between 0.10 and 0.15 greatly reduce the workload while preserving attractive financial returns.

At a threshold of 0.10 for the calibrated logistic regression, the system would flag about 434 employees per day, which corresponds to a utilisation of 0.96 relative to the processing capacity. This policy is expected to correctly prioritise roughly 166 high-risk employees and misclassify approximately 268 low-risk employees per day. Under the assumed intervention costs and benefits, this operating point yields an estimated net monetary benefit of 12,615 currency units per day and a return on investment of about 5.4, while satisfying both the uncertainty and capacity criteria in the policy window. The calibrated random forest reaches comparable returns only at higher thresholds; at a threshold of 0.12, for example, it produces about 426 alerts per day with a utilisation of 0.95 and an estimated return on investment of approximately 5.2. These findings support selecting a calibrated logistic model with a threshold near 0.10 as the preferred operating point, balancing decision value, workload feasibility, and monetary efficiency.

Subgroup Calibration and Fairness at the Chosen Operating Point

At the chosen operating threshold, we examined calibration and error dispersion across employee subgroups. **Figure 4** ranks the 12 worst groups by their subgroup-expected calibration error (ECE). While overall performance remains satisfactory, several groups exhibit noticeably higher calibration errors than the full sample, including lower salary bands, laboratory technicians, manufacturing directors, employees with marketing-related education, and staff with overtime contracts or frequent business travel. Some smaller groups also show elevated false-positive or false-negative rates **(Figure 4).**

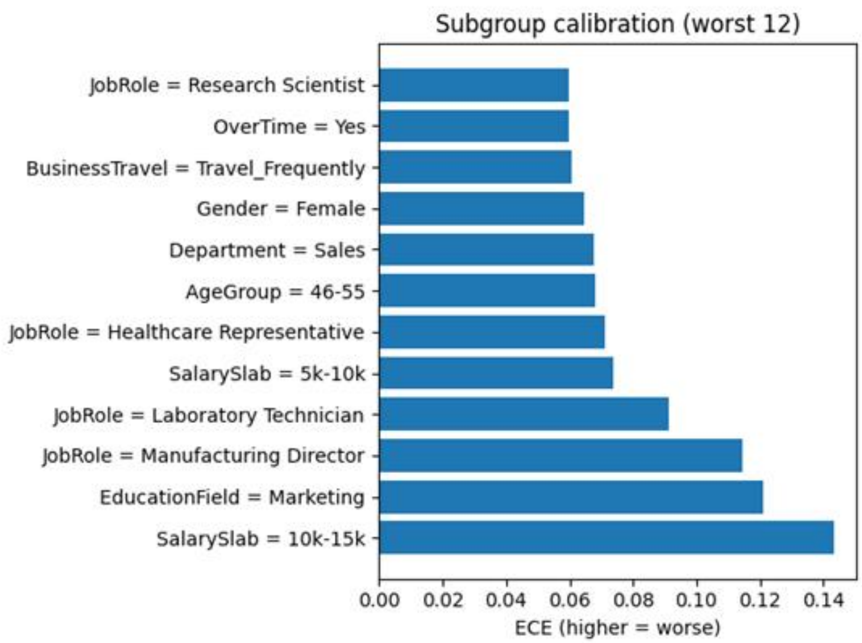

Figure 4. Subgroup calibration, the worst twelve groups ranked by ECE

These patterns show that the calibrated operating point works fairly well for most employees, but targeted monitoring is needed for specific roles and pay bands with higher miscalibration. In practice, this requires governance arrangements that regularly audit subgroup performance and include human review for decisions affecting groups most at risk of misclassification.

Overview of Main Findings

This study evaluated whether calibrated machine learning models can assist with capacity-constrained retention interventions. Using a comprehensive HR dataset, we found that attrition primarily occurs among specific roles and work conditions, notably sales representatives, younger employees, staff with overtime contracts, and frequent travelers. This aligns with evidence that high demands combined with limited resources increase exit risk. Comparing logistic regression and random forest models, we found that a calibration-first pipeline with isotonic recalibration significantly improved Brier scores and expected calibration error, with only slight reductions in AUC, yielding probability estimates that closely match observed attrition. Decision curve analysis also showed that a calibrated logistic model operating at a threshold around 0.10 provides higher net benefit than treat-all or treat-none strategies, while maintaining alert volume within staffing capacity and generating positive financial returns. However, subgroup analyses revealed pockets of miscalibration and varying error rates across roles and salary levels, suggesting inconsistent model performance across employee groups.

Theoretical Implications

The findings expand the Job Demands-Resources perspective by demonstrating how combinations of demands and resources can be converted into individual risk scores that inform targeted interventions rather than merely providing explanations. From a human capital perspective, calibrated probabilities enable explicit expected value calculations that align retention strategies with the likelihood and cost of losing specific employees, advancing the strategic allocation of limited human capital resources. The results also relate to social exchange theory: when development opportunities or flexible arrangements are based on transparent and reliable risk estimates, they can serve as signals of organizational support, although this potential benefit depends on avoiding systematic disadvantages for certain groups.

Managerial Implications for Retention Triage

For practitioners, the study provides a blueprint for designing retention analytics within capacity constraints. First, managers should look beyond accuracy and AUC, and regularly calibrate models so that predicted risks align with observed attrition, thereby reducing the over-treatment of low-risk employees and the under-treatment of high-risk ones. Second, decision curve analysis and workload translation offer an accessible basis for selecting thresholds that yield a positive net benefit while respecting daily processing limits; in this context, a threshold around 0.10 for a calibrated logistic model balances these factors. Third, presenting results in terms of net monetary benefit and return on investment helps facilitate discussions with finance and senior leadership about when algorithmic retention programs are economically justified compared to other initiatives.

Fairness, Governance, and Responsible Use of Retention Analytics

The subgroup results underline that even well-calibrated models can generate uneven benefits and burdens across employee groups, reinforcing concerns in the algorithmic HRM literature about fairness, transparency, and worker autonomy. HR departments that adopt retention analytics should therefore embed governance mechanisms that include routine monitoring of subgroup calibration and error rates at the chosen threshold, documentation of threshold selection and model changes, and opportunities for employee representation in oversight processes. Where disparities are detected, organisations may need to adjust thresholds, apply targeted recalibration, or introduce human review for sensitive cases. Clear communication about the purpose, benefits, and safeguards of the system is essential to maintain trust and to ensure that retention analytics support rather than undermine the social exchange relationships that underpin long-term employee commitment.

Conclusion

This study examined whether calibrated machine learning models can support retention decisions under capacity constraints. Using employee-level data on roles, work conditions, and rewards, we compared logistic regression and random forest models and evaluated them based on discrimination, probability calibration, decision curve net benefit, and workload. The results indicate that attrition risk is unevenly distributed and that a calibration-first pipeline, combined with decision curve analysis, provides a transparent basis for prioritizing scarce retention resources.

A calibrated logistic regression model provided the best balance between AUC and probability reliability. Isotonic calibration enhanced Brier scores and expected calibration error with only minor reductions in discrimination, and decision curve analysis identified an operating threshold around 0.10 where alerts stay within staffing capacity while delivering significant net benefit and a positive return on investment. Subgroup analyses showed areas of miscalibration and uneven error rates, indicating that algorithmic triage is not inherently fair and must be guided by governance, including monitoring, documentation, and human oversight.

For researchers, the study shows how Job Demands-Resources, human capital, and social exchange perspectives can be applied within a probabilistic triage framework, evaluated on decision value rather than accuracy alone. For practitioners, it provides a practical template for using calibrated models, decision curve analysis, and routine subgroup audits to create retention policies that are capacity-aware and fairness-sensitive. While replication across other organizations and sectors is necessary, the results suggest that carefully calibrated, well-governed retention analytics can help target limited resources toward employees who are both most at risk and most valuable to retain.

Limitations and Directions for Future Research

This study has several limitations that highlight areas for future research. The analysis relies on a single organizational dataset and cross-sectional data, so the observed patterns may not generalize across industries or establish causal links between predictors and attrition. Net monetary benefit estimates are based on simplified cost and benefit assumptions, and fairness analyses are limited to observable subgroups, leaving unexamined potential biases related to work allocation, performance evaluation, or informal networks. Future research could involve multiple employers and longitudinal data to study how calibration and decision value change across various contexts and over time, compare different calibration techniques and fairness-constrained optimization methods, and include qualitative research with managers and employees to understand how algorithmic triage tools are viewed and how human oversight and explanation should be designed in practice.

Acknowledgments: None

Conflict of Interest: None

Financial Support: This paper was financially supported by Mahasarakham Business School, Mahasarakham University, Thailand.

Ethics Statement: This study used publicly available, de-identified secondary data. According to institutional policy, it does not constitute human-subjects research and does not require IRB approval or informed consent.

References

Akasheh, M. Al, Malik, E. F., Hujran, O., & Zaki, N. (2024). A decade of research on machine learning techniques for predicting employee turnover: a systematic literature review. *Expert Systems with Applications, 238*, 121794. doi:10.1016/j.eswa.2023.121794

Alabdali, M. A., Khan, S. A., Yaqub, M. Z., Alshahrani, M. A., & Bolívar, R. (2024). Harnessing the power of algorithmic human resource management and human resource strategic decision-making for achieving organizational success: an empirical analysis. *Sustainability, 16*(11), 4854. doi:10.3390/su16114854

Alqahtani, H., Almagrabi, H., & Alharbi, A. (2024). Employee attrition prediction using machine learning models: a review paper. *International Journal of Artificial Intelligence and Applications (IJAIA), 15*(2). doi:10.5121/ijaia.2024.1520223

Avrahami, D., Pessach, D., Singer, G., & Chalutz Ben-Gal, H. (2022). A human resources analytics and machine-learning examination of turnover: implications for theory and practice. *International Journal of Manpower, 43*(6), 1405–1424. doi:10.1108/IJM-12-2020-0548

Bakker, A. B., Demerouti, E., & Sanz-Vergel, A. (2023). Job demands–resources theory: ten years later. *Annual Review of Organizational Psychology and Organizational Behavior, 10*, 25–53. doi:10.1146/annurev-orgpsych-120920-053933

Bujold, A., Roberge-Maltais, I., Parent-Rocheleau, X., Boasen, J., Sénécal, S., & Léger, P. M. (2024). Responsible artificial intelligence in human resources management: a review of the empirical literature. *AI and Ethics, 4*(4), 1185–1200. doi:10.1007/s43681-023-00325-1

Cárdenas López, P. A., & Tabares Betancur, M. S. (2024). Discovering key aspects to reduce employee turnover using a predictive model. *Lecture Notes in Computer Science*, 380–395. doi:10.1007/978-3-031-47372-2_30

Chen, X., Al Mamun, A., Hussain, W. M. H. W., Jingzu, G., Yang, Q., & Al Shami, S. S. A. (2023). Envisaging the job satisfaction and turnover intention among the young workforce: evidence from an emerging economy. *PLOS ONE, 18*(6), e0287284. doi:10.1371/journal.pone.0287284

Chinyamurindi, W. T., & Mashavira, N. (2024). Job satisfaction and turnover: the role of creativity, engagement, and decent work amongst employees. *SA Journal of Human Resource Management, 22*, Article 2713. doi:10.4102/sajhrm.v22i0.2713

De Vos, S., Bockel-Rickermann, C., Van Belle, J., & Verbeke, W. (2024). Predicting employee turnover: scoping and benchmarking the state-of-the-art. *Business & Information Systems Engineering, 67*(5), 733–752. doi:10.1007/s12599-024-00898-z

Dima, J., Gilbert, M. H., Dextras-Gauthier, J., & Giraud, L. (2024). The effects of artificial intelligence on human resource activities and the roles of the human resource triad: opportunities and challenges. *Frontiers in Psychology, 15*, 1360401. doi:10.3389/fpsyg.2024.1360401

Fenwick, A., Molnar, G., & Frangos, P. (2023). Revisiting the role of HR in the age of AI: bringing humans and machines closer together in the workplace. *Frontiers in Artificial Intelligence, 6*, 1272823. doi:10.3389/frai.2023.1272823

García, E., & Jaramillo, S. (2023). Telescopic retention in prosthodontics: a digital approach for enhanced patient outcomes. *Asian Journal of Periodontics and Orthodontics, 3*(1), 25–29.

Gazi, M. S., Nasiruddin, M., Dutta, S., Sikder, R., Huda, C. B., & Islam, M. Z. (2024). Employee attrition prediction in the USA: a machine learning approach for HR analytics and talent retention strategies. *Journal of Business and Management Studies, 6*(3), 47–59. doi:10.32996/jbms.2024.6.3.6

Guerranti, F., & Dimitri, G. M. (2022). A comparison of machine learning approaches for predicting employee attrition. *Applied Sciences, 13*(1), 267. doi:10.3390/app13010267

Hunkenschroer, A. L., & Luetge, C. (2022). Ethics of AI-enabled recruiting and selection: a review and research agenda. *Journal of Business Ethics, 178*(4), 977–1007. doi:10.1007/s10551-022-05049-6

Iparraguirre-Villanueva, O., Chauca-Huete, L., Prieto-Chavez, R., & Paulino-Moreno, C. (2024). Employee attrition prediction using machine learning models. In *Proceedings of the 22nd LACCEI International Multi-Conference for Engineering, Education and Technology (LACCEI 2024)*.

Jabagi, N., Croteau, A. M., Audebrand, L. K., & Marsan, J. (2025). Do algorithms play fair? Analysing the perceived fairness of HR decisions made by algorithms and their impacts on gig-workers. *International Journal of Human Resource Management, 36*(2), 235–274. doi:10.1080/09585192.2024.2441448

Köchling, A., Wehner, M. C., & Ruhle, S. A. (2024). This (AI) isn't fair? Employee reactions to artificial intelligence (AI) in career development systems. *Review of Managerial Science, 19*(4), 1195–1228. doi:10.1007/s11846-024-00789-3

Kumar, P., Gaikwad, S. B., Ramya, S. T., Tiwari, T., Tiwari, M., & Kumar, B. (2023). Predicting employee turnover: A systematic machine learning approach for resource conservation and workforce stability. *Engineering Proceedings, 59*(1), 117. doi:10.3390/engproc2023059117

Malik, A., Budhwar, P., & Kazmi, B. A. (2023). Artificial intelligence (AI)-assisted HRM: towards an extended strategic framework. *Human Resource Management Review, 33*(1), 100940. doi:10.1016/j.hrmr.2022.100940

Mansor, N., Sani, N. S., & Aliff, M. (2021). Machine learning for predicting employee attrition. *International Journal of Advanced Computer Science and Applications, 12*(11), 435–445. doi:10.14569/ijacsa.2021.0121149

Mazzetti, G., Robledo, E., Vignoli, M., Topa, G., Guglielmi, D., & Schaufeli, W. B. (2023). Work engagement: a meta-analysis using the job demands–resources model. *Psychological Reports, 126*(3), 1069–1107. doi:10.1177/00332941211051988

Meijerink, J., Boons, M., Keegan, A., & Marler, J. (2021). Algorithmic human resource management: synthesizing developments and cross-disciplinary insights on digital HRM. *International Journal of Human Resource Management, 32*(12), 2545–2562. doi:10.1080/09585192.2021.1925326

Meyer, J. P., & Allen, N. J. (1991). A three-component conceptualization of organizational commitment. *Human Resource Management Review, 1*(1), 61–89. doi:10.1016/1053-4822(91)90011-Z

Moon, K. K., Lim, J., & Kim, J. S. (2024). Examining the effect of organizational justice on turnover intention and the moderating role of generational differences: Evidence from Korean public employees. *Sustainability, 16*(6), 2454. doi:10.3390/su16062454

Moritz, J. M., Pomrehn, L., Steinmetz, H., & Wehner, M. (2023). A meta-analysis and review about the (un)fairness perceptions of algorithmic decision-making. In *Academy of Management Proceedings* (Vol. 2023, No. 1, Article 15483). Academy of Management. doi:10.5465/amproc.2023.15483

Nicolls, C., Haar, J., & Wallis, A. (2022). Job demands and resources predict flourishing and turnover intentions among Aotearoa New Zealand employees during the COVID-19 pandemic. *New Zealand Journal of Employment Relations, 48*(1), 1–25.

Özkan, A. H. (2023). Organizational justice perceptions and turnover intention: a meta-analytic review. *Kybernetes, 52*(8), 2886–2899. doi:10.1108/k-01-2022-0119

Palma-Moreira, A., Dias, A. L., Pereira, B., & Au-Yong-Oliveira, M. (2024). Competence development and affective commitment as mechanisms that explain the relationship between organizational culture and turnover intentions. *Administrative Sciences, 14*(9), 223. doi:10.3390/admsci14090223

Pham, T. T. (2024). Linking family supports and Vietnamese employee performance: the mediator role of work engagement. *Journal of Organizational Behavior Research, 9*(1), 15–31. doi:10.51847/w3dmjbbfqq

Piovani, D., Sokou, R., Tsantes, A. G., Vitello, A. S., & Bonovas, S. (2023). Optimizing clinical decision making with decision curve analysis: insights for clinical investigators. *Healthcare, 11*(16), 2244. doi:10.3390/healthcare11162244

Poku, C. A., Bayuo, J., Agyare, V. A., Sarkodie, N. K., & Bam, V. (2025). Work engagement, resilience, and turnover intentions among nurses: a mediation analysis. *BMC Health Services Research, 25*(1), 71. doi:10.1186/s12913-025-12242-6

Rajopadhye, B. D., Londhe, V. A., Pingle, N. A., & Dhande, P. P. (2023). Exploring the impact of PowerPoint lectures in pharmacology: insights from phase II medical students. *Annals of Pharmacy Education, Safety, and Public Health Advocacy, 3*, 37–42.

Raza, A., Munir, K., Almutairi, M., Younas, F., & Fareed, M. M. S. (2022). Predicting employee attrition using machine learning approaches. *Applied Sciences, 12*(13), 6424. doi:10.3390/app12136424

Rehman, T. (2025). Job demands, job resources and occupational turnover intentions: testing a stress-mediated model grounded in JD-R and conservation of resources theories in higher education institutions of Pakistan. *Journal of Business and Management Research, 4*(3), 727–747. doi:10.64105/jbmr.04.03.534

Sartori, R., Ceschi, A., Zene, M., Scipioni, L., & Monti, M. (2023). The relationship between perceived organizational support (POS) and turnover intention: the mediating role of job motivation, affective and normative commitment. *Informing Science: The International Journal of an Emerging Transdiscipline, 26*, 5–21. doi:10.28945/5070

Shafie, M. R., Khosravi, H., Farhadpour, S., Das, S., & Ahmed, I. (2024). A cluster-based human resources analytics for predicting employee turnover using optimized artificial neural networks and data augmentation. *Decision Analytics Journal, 11*, 100461. doi:10.1016/j.dajour.2024.100461

Shlash Mohammad, A. A., Alkhazali, Z., Shelash Mohammad, S. I., Al Oraini, B., Vasudevan, A., Alqahtani, M. M., & Alshurideh, M. T. (2025). Machine learning models for predicting employee attrition: a data science perspective. *Data and Metadata, 4*, 669. doi:10.56294/dm2025669

Sienkiewicz, L. (2024). Algorithmic human resources management. In *HRM 5.0: Unpacking the digitalisation of human resource management* (pp. 57–85). Springer. doi:10.1007/978-3-031-58912-6_4

Singh, K., Shah, N. H., & Vickers, A. J. (2023). Assessing the net benefit of machine learning models in the presence of resource constraints. *Journal of the American Medical Informatics Association, 30*(4), 668–673. doi:10.1093/jamia/ocad006

Strohmeier, S., Becker, M., & Scheer-Weller, E. (2026). Beyond aversion – principles of appropriate algorithmic decision-making in human resource management. *Expert Systems with Applications, 296*, 128954. doi:10.1016/j.eswa.2025.128954

Sugimori, T., Yamaguchi, M., Kikuta, J., Shimizu, M., & Negishi, S. (2022). The biomechanical and cellular response to micro-perforations in orthodontic therapy. *Asian Journal of Periodontics and Orthodontics, 2*(1), 1–15. doi:10.51847/z9adsj59rj

Talebi, H., Khatibi Bardsiri, A., & Bardsiri, V. K. (2025). Machine learning approaches for predicting employee turnover: a systematic review. *Engineering Reports, 7*(8), e70298. doi:10.1002/eng2.70298

Tang, H., An, S., Zhang, L., Xiao, Y., & Li, X. (2024). The antecedents and outcomes of public service motivation: a meta-analysis using the job demands–resources model. *Behavioral Sciences, 14*(10), 861. doi:10.3390/bs14100861

Van Calster, B., McLernon, D. J., Van Smeden, M., Wynants, L., Steyerberg, E. W., Bossuyt, P., Collins, G. S., MacAskill, P., Moons, K. G. M., & Vickers, A. J. (2019). Calibration: The achilles heel of predictive analytics. *BMC Medicine, 17*(1), 230. doi:10.1186/s12916-019-1466-7

Van Heerden, J., Du Plessis, M., & Becker, J. R. (2022). Walking the tightrope of job demands and resources: leveraging work engagement to counter turnover intentions of information technology professionals. *Frontiers in Psychology, 13*, 660308. doi:10.3389/fpsyg.2022.660308

Vinh, N. Q. (2023). The influence of personal resources on job engagement of employees in tourism companies. *Journal of Organizational Behavior Research, 8*(1), 231–243. doi:10.51847/is94kjrwdn

Wong, K. P., Zhang, B., Xie, Y. J., Wong, F. K. Y., Lai, C. K. Y., Chen, S. C., & Qin, J. (2024). Impacts of job demands on turnover intention among registered nurses in Hong Kong public hospitals: exploring the mediating role of burnout and moderating effect of pay level satisfaction. *Journal of Nursing Management, 2024*(1), 3534750. doi:10.1155/2024/3534750

Wu, W., & Fukui, S. (2024). Using human resources data to predict turnover of community mental health employees: prediction and interpretation of machine learning methods. *International Journal of Mental Health Nursing, 33*(6), 2180–2192. doi:10.1111/inm.13387

Yakupoğlu, E. (2025). Impact of job satisfaction and motivation on intention to quit and organizational commitment among airline employees. *Journal of Organizational Behavior Research, 10*(3), 98–110. doi:10.51847/xftqrhjplf

Zhao, S., Ma, Z., Li, H., Wang, Z., Wang, Y., & Ma, H. (2024). The impact of organizational justice on turnover intention among primary healthcare workers: the mediating role of work motivation. *Risk Management and Healthcare Policy, 17*, 3017–3028. doi:10.2147/rmhp.s486535

Zhu, L. L., Wang, H. J., Xu, Y. F., Ma, S. T., & Luo, Y. Y. (2023). The effect of work engagement and perceived organizational support on turnover intention among nurses: a meta-analysis based on the Price–Mueller model. *Journal of Nursing Management, 2023*(1), 3356620. doi:10.1155/2023/3356620