



International Journal of Research Publication and Reviews

Journal Homepage: www.ijrpr.com ISSN 2582-7421

Human AI Collaboration in Financial Risk Management for Railway Resilience

¹ Bhukya Ganesh Naik, ² Dr. Ajay Kumar Varshney

¹ Research Scholar, BBA Department, KLEF, Email : Ganeshhpersonal22@gmail.com

² Professor, BBA Department, KLEF

ABSTRACT

The railway industry supports passenger movement, freight transport, employment and trade but its finances are vulnerable to demand shocks, rising costs and disruptions such as the COVID-19 pandemic. Artificial intelligence (AI) can improve forecasting, anomaly detection and risk monitoring, but full automation may neglect operational experience, context and human accountability. This study examines a human-AI financial risk management framework that links AI capability, explainability, trust, human oversight, human-AI collaboration, decision quality and financial resilience. A structured Likert questionnaire dataset of 120 respondents (railway finance, management, IT and operations roles) was analysed with descriptive statistics, Cronbach's alpha, chi-square tests of independence, one-way ANOVA with post-hoc tests and simple and multiple linear regression. Chi-square tests revealed significant associations between respondent role and organisational AI use ($\chi^2(6) = 26.60$, $p < .001$, $V = .33$), AI familiarity and organisational AI use ($\chi^2(6) = 51.05$, $p < .001$, $V = .46$), sector and preferred AI application ($\chi^2(6) = 78.78$, $p < .001$, $V = .57$) and role and preferred responsibility model ($\chi^2(4) = 21.15$, $p < .001$, $V = .30$). ANOVA showed significant differences in trust by AI familiarity ($F(3,116) = 20.91$, $\eta^2 = .35$), AI-capability perception by role ($\eta^2 = .24$), resilience importance by sector ($\eta^2 = .21$) and perceived collaboration benefit by experience ($\eta^2 = .18$). All six proposed hypotheses were supported by simple regression, and human oversight and trust predicted decision quality in multiple regression ($R^2 = .40$) while decision quality and training predicted resilience importance ($R^2 = .22$). Notably 97 of 120 records in the dataset are labelled as synthetic, so the results are exploratory and illustrative and should not be presented as evidence from a real survey of railway professionals.

Keywords: Human-AI collaboration, financial risk management, railway resilience, artificial intelligence, chi-square test, ANOVA, regression, Indian Railways.

1. Introduction

The railway industry is an essential component of a nation's transportation infrastructure and economic development. It facilitates the transportation of passengers and goods, connects industrial and commercial hubs, and provides vital public services. Financial sustainability is critical for maintaining infrastructure, improving service quality, investing in modern technology and covering operating expenses.

Railway organisations face financial risks including fluctuations in passenger revenue, changes in freight demand, rising maintenance costs, fuel and energy price variations, unexpected infrastructure expenditure and disruptions caused by pandemics, natural disasters or economic crises. The COVID-19 pandemic demonstrated how a sudden fall in passenger demand could damage revenue while essential operating responsibilities continue.

Traditional financial risk management relies on historical records, budgeting, expert experience and periodic reporting. These approaches remain vital but miss emerging risks in complex, rapidly changing environments. Artificial intelligence offers new opportunities: machine learning can support revenue forecasting, anomaly detection, expenditure analysis and the identification of financial vulnerabilities and help managers evaluate alternative scenarios.

However, AI systems can produce inaccurate predictions when data are incomplete, historical patterns change or unexpected events occur. Human judgement is crucial to verifying recommendations, considering context and exercising accountability. Human-AI collaboration combines the analytical capability of AI with human expertise and responsibility, and can contribute to earlier risk identification, better financial planning, effective resource allocation and faster recovery from disruptions.

This study therefore examines a framework that combines AI-based financial risk identification with human review and decision-making to strengthen railway financial resilience. Using questionnaire data, it tests how perceptions of AI capability, explainability, trust, human oversight, collaboration and decision quality are related to one another and to the perceived importance of financial resilience.

2. Review of Literature

2.1 Financial risk management in the railway industry

Financial risk management entails identifying, measuring, monitoring and controlling risks that may affect an organisation's financial position. Railway organisations must balance revenue generation, operating expenditure, maintenance, capital investment and public-service responsibilities, so forecasting and risk monitoring are essential for long-term stability.

2.2 COVID-19 and railway financial resilience

The COVID-19 pandemic disrupted passenger services and placed heavy financial pressure on railway operators. In India, passenger earnings fell sharply in 2020-21 while freight operations helped sustain overall revenue (Indian Railways, 2021). These changes demonstrate the importance of revenue diversification, financial planning and preparedness for unexpected disruptions.

2.3 Artificial intelligence in financial risk management

AI and machine-learning methods can support financial forecasting, unusual-transaction identification, expenditure monitoring and risk classification. A literature review of AI in railway systems found that maintenance and inspection received substantial attention while revenue management and policy were less represented (A literature review of Artificial Intelligence applications in railway systems, 2023). The success of AI depends on data quality, appropriate model selection and continual performance monitoring.

2.4 Explainability, trust and human-AI collaboration

Research on financial AI has examined explainability, governance and trust (Bowden et al., 2024; Hadji Misheva et al., 2022), and studies of human-AI collaboration in lending reveal that combining algorithmic and human judgement can reduce decision noise (Sachan et al., 2024). Meaningful collaboration requires clear responsibilities, understandable recommendations and mechanisms to challenge incorrect AI outputs.

2.5 Organisational resilience and financial sustainability

Organisational resilience is the capacity to prepare for, absorb, respond to and recover from disruptions. Financial resilience includes maintaining liquidity, stabilising revenue, controlling costs and securing resources for continued operations. AI-supported forecasting and human-led contingency planning can contribute to these capabilities.

2.6 Summary of the literature

Existing research provides foundations in railway financial management, AI-based forecasting, human-AI decision-making and organisational resilience, but their integration into a railway-specific financial risk management framework requires further empirical investigation. This is a thematic review rather than an exhaustive systematic review; before journal submission it should be expanded with additional verified peer-reviewed studies.

3. Research Gap

Previous research has explored AI in railway maintenance, inspection, planning and operations, while research on financial AI has investigated explainability, governance and human-AI collaboration. Based on these directions, the following gaps are addressed:

1. Limited integration: how human-AI collaboration can be integrated specifically into railway financial risk management.
2. Railway-specific evidence: general financial AI frameworks may not reflect high infrastructure costs, public-service obligations and fluctuating passenger and freight demand.
3. Human oversight: more evidence is needed on the effect of human review and explainable AI on decision quality.
4. Resilience outcomes: the relationship between AI-supported risk identification and resilience requires empirical examination.
5. Post-COVID preparedness: lessons from pandemic-related disruption should inform preparedness for future financial shocks.

Proposed contribution: this study empirically assesses a human-AI financial risk management framework tailored to railway resilience. The novelty claim should be verified with a systematic literature search before publication.

4. Statement of the Problem

Railway organisations face financial uncertainty from changing passenger demand, freight volumes, maintenance expenditure, energy costs and unexpected disruptions. Official Indian Railways figures indicate that passenger earnings fell from ₹50,669.09 crore in 2019-20 to ₹15,248.49 crore in 2020-21 (Indian Railways, 2021). Traditional financial risk management may not identify emerging risks quickly or combine diverse financial and operational information effectively. While AI offers forecasting, anomaly-detection and risk-analysis capabilities, its recommendations must still be interpreted, validated and overseen.

The central problem is to determine how collaboration between AI systems and human financial decision-makers can improve financial risk identification and support railway resilience, and whether AI capability, explainability, trust, human oversight and decision quality are associated with stronger perceived financial resilience.

5. Objectives of the Study

Main objective: to develop and empirically examine a human-AI financial risk management framework for strengthening financial resilience in the railway industry.

Specific objectives:

1. To explore the role of AI in identifying and predicting railway financial risks.
2. To assess the relationship between AI capability and financial risk identification.
3. To investigate whether explainable AI improves trust in AI-generated financial recommendations.
4. To examine the influence of human oversight on the quality of financial decisions.
5. To analyse the relationship between human-AI collaboration and perceived railway financial resilience.
6. To test whether respondent characteristics (role, familiarity, sector, experience) are associated with attitudes towards AI adoption and responsibility.
7. To propose strategies for responsible AI implementation in railway financial management.

6. Hypotheses of the Study

Each relationship is tested against a null hypothesis (H_0), which states that there is no effect, association or difference, and an alternative hypothesis (H_a), which states that there is. H_0 is rejected when $p < 0.05$. For the regression tests, the slope (b) is the parameter tested: $H_0: b = 0$ and $H_a: b > 0$ (a positive relationship; both variables are coded in the same direction, see Section 7.4).

No.	Null hypothesis (H_0)	Alternative hypothesis (H_a)	Test
H1	H01: AI capability has no significant relationship with financial risk identification ($b = 0$).	Ha1: AI capability has a significant positive relationship with financial risk identification ($b > 0$).	Linear regression
H2	H02: Explainability has no significant relationship with trust in AI-generated financial recommendations.	Ha2: Explainability has a significant positive relationship with trust in AI-generated financial recommendations.	Linear regression
H3	H03: Human oversight has no significant relationship with financial decision quality.	Ha3: Human oversight has a significant positive relationship with financial decision quality.	Linear regression
H4	H04: Human–AI collaboration has no significant relationship with perceived financial resilience.	Ha4: Human–AI collaboration has a significant positive relationship with perceived financial resilience.	Linear regression
H5	H05: Financial decision quality has no significant relationship with perceived financial resilience.	Ha5: Financial decision quality has a significant positive relationship with perceived financial resilience.	Linear regression

No.	Null hypothesis (H0)	Alternative hypothesis (Ha)	Test
H6	H06: Trust in AI-generated recommendations has no significant relationship with human–AI collaboration.	Ha6: Trust in AI-generated recommendations has a significant positive relationship with human–AI collaboration.	Linear regression
H7a	H07a: Professional role and organisational AI use are independent.	Ha7a: Professional role and organisational AI use are associated.	Chi-square
H7b	H07b: AI familiarity and organisational AI use are independent.	Ha7b: AI familiarity and organisational AI use are associated.	Chi-square
H7c	H07c: Sector and preferred AI application are independent.	Ha7c: Sector and preferred AI application are associated.	Chi-square
H7d	H07d: Professional role and preferred responsibility model are independent.	Ha7d: Professional role and preferred responsibility model are associated.	Chi-square
H8a	H08a: Mean trust in AI (Q15) is equal across AI-familiarity groups.	Ha8a: Mean trust differs for at least one familiarity group.	One-way ANOVA
H8b	H08b: Mean AI-capability perception is equal across professional roles.	Ha8b: Mean AI-capability perception differs for at least one role.	One-way ANOVA
H8c	H08c: Mean importance of financial resilience (Q25) is equal across sectors.	Ha8c: Mean importance of resilience differs for at least one sector.	One-way ANOVA
H8d	H08d: Mean perceived collaboration benefit (Q20) is equal across experience groups.	Ha8d: Mean collaboration benefit differs for at least one experience group.	One-way ANOVA

Table 1. Null and alternative hypotheses and statistical tests. Hypotheses H1–H6 use directional (one-sided) alternatives; H7–H8 are non-directional.

7. Research Methodology Essay

7.1 Research design

The study employs a quantitative, descriptive and correlational design based on a structured questionnaire with items to elicit demographic data, categorical preferences items and five-level Likert-type attitude items on AI in railway financial risk management, trust and explainability, human oversight, collaboration and decision quality, and financial resilience.

7.2 Study population and sample

The target population comprises railway finance and accounts professionals, managers and administrators, operations and planning staff, technology and data professionals, and related transport-sector experts. The analysed dataset contains 120 respondents: 60 finance/accounts staff, 35 managers/administrators, 21 IT/AI/data professionals and 4 operations/planning staff.

7.3 Data provenance

The dataset includes a status field. Twenty-three records (19.2%) are labelled “Original response” and 97 records (80.8%) are labelled “Synthetic—generated, not collected”. The paper therefore reports analyses of the full dataset as a demonstration of the analytical framework, and repeats the regression tests on the 23 original records as a sensitivity check. Findings must not be generalised to the railway workforce.

7.4 Variables and measurement

Responses are stored as numeric codes. Text and code columns indicate that 1 represents the strongest agreement or highest level (for example, “A great deal” of trust = 1 and “Extremely important” = 1), and higher codes represent weaker agreement. This coding was assumed from the data and should be confirmed against the original codebook. Consequently, a positive regression slope indicates a positive association between two attitudes because both variables are coded in the same direction, and a lower group mean indicates stronger agreement.

Construct	Items used	Role in analysis
AI capability	Q12 forecasting accuracy, Q13 irregularity detection, Q22 comparing strategies (mean)	Predictor (H1); ANOVA outcome
Financial risk identification	Q10 and its second column (mean of two items)	Outcome (H1)
Explainability	Q14	Predictor (H2)
Trust in AI	Q15 (text converted to codes 1–4)	Outcome (H2), predictor (H6, M1, M2), ANOVA outcome
Human oversight	Q16 human review, Q18 human expertise (mean)	Predictor (H3, M2)
Human–AI collaboration	Q20	Outcome (H6), predictor (H4, M1, M2), ANOVA outcome
Decision quality	Q21	Outcome (H3, M2), predictor (H5, M1)
Organisational training	Q24	Control in M1
Financial resilience	Q25 importance of financial resilience (text converted to codes)	Outcome (H4, H5, M1), ANOVA outcome

Table 2. Construct measurement. The dataset has two columns labelled Q10, and does not include Q17 or Q26–Q30 of the questionnaire; Q10 is therefore treated as two items and resilience is measured by Q25 only.

7.5 Statistical tools

Descriptive statistics and reliability: frequencies, percentages, mean, standard deviation, and Cronbach's alpha.

Chi-square test of independence: associations between categorical variables, with Cramer's V as the effect size. Because of the need to merge some cells with sparse data (e.g., operations staff with IT staff), no more than 25% of cells had expected counts of less than 5.

One-way analysis of variance (ANOVA): differences in the mean scores across groups, with effect size (eta-squared), Levene's test of homogeneity of variance, Tukey's post-hoc comparisons, and Kruskal–Wallis test as a non-parametric alternative.

Linear regression: simple regressions for hypotheses H1–H6 and linear regression for two multiple regression models, with variance inflation factors (VIFs), Durbin–Watson statistics, and heteroscedasticity-consistent (HC3) p-values for testing.

The calculations were performed using formulas embedded in Microsoft Excel (CHISQ/CHIDIST, FDIST, LINEST functions) and were used as a cross-check in Python software (SciPy, statsmodels). The final results were consistent to at least four decimal places. The selected significance threshold was $\alpha = 0.05$.

7.6 Ethical considerations

When developing the survey, the researchers made sure to ensure respondents' right not to participate in the study and provide informed consent. Additionally, the anonymity of the respondents was considered paramount, making the analysis based on personal identifiable data impossible. Confidential information about railway's finances should not be made public, so publishing any records that identify respondents would be unethical. For this reason, the data's status as anonymized was explicitly stated.

8. Conceptual Framework

The conceptual framework assumes that AI capabilities contribute to risk identification, explainability leads to trust, and trust and human oversight lead to effective human–AI collaboration and decision-making, ultimately resulting in financial resilience. In Fig. 1, we present the framework with tentative estimates of the simple-regression slopes identified in the analysis in Chapter 9.

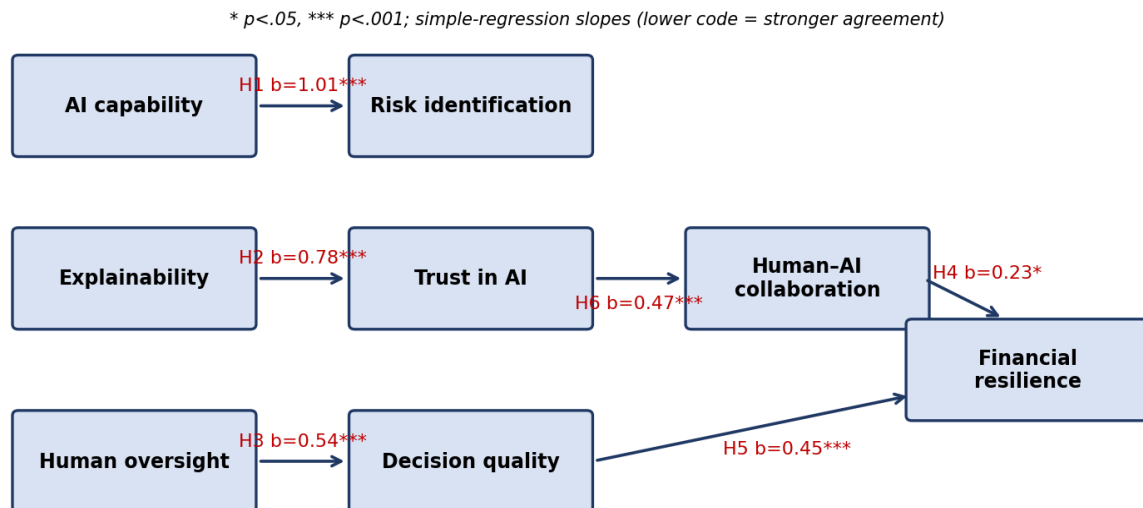


Figure 1. Conceptual framework with estimated simple-regression slopes (b) for H1–H6.

9. Data Analysis and Results

9.1 Respondent profile

Table 3. Respondent profile (N = 120)

Variable	Category	n	%
Professional role	Railway finance/accounts staff	60	50.0%
	Railway manager/administrator	35	29.2%
	IT/AI/data professional	21	17.5%
	Railway operations/planning staff	4	3.3%
Sector	Indian Railways	60	50.0%
	Railway technology/IT company	33	27.5%
	Railway public-sector organisation	19	15.8%
	Private railway/logistics company	8	6.7%
Experience	Less than 1 year	95	79.2%
	More than 20 years	13	10.8%
	6–10 years	12	10.0%
Age group	Below 25 years	116	96.7%
	46–55 years	4	3.3%
Gender	Male	102	85.0%
	Female	18	15.0%

Variable	Category	n	%
Education	Undergraduate	74	61.7%
	Diploma	35	29.2%
	Postgraduate	11	9.2%

Finance and accounts staff form the largest group (50.0%), followed by managers (29.2%) and IT/AI/data professionals (17.5%). The sample is, however, strongly skewed: 96.7% are under 25 years old, 79.2% have less than one year of experience and 85.0% are male. This age and experience profile differs sharply from the target population of experienced railway professionals and further limits generalisation.

9.2 Descriptive statistics and reliability

Table 4. Item descriptives (lower scores = stronger agreement)

Survey statement	Mean	SD	% codes 1–2
AI can identify financial risks earlier (Q10)	1.92	0.88	81.7%
AI can improve forecasting accuracy (Q12)	1.92	0.95	75.0%
AI can detect unusual expenditure (Q13)	1.93	0.80	80.0%
AI should explain its recommendations (Q14)	1.98	0.84	65.8%
Trust if reasoning is explained (Q15)	2.08	1.12	61.7%
Human experts should review AI recommendations (Q16)	1.98	0.74	73.3%
Human expertise can identify AI errors (Q18)	1.97	0.81	69.2%
Collaboration improves risk identification (Q20)	2.05	0.97	67.5%
Collaboration improves decision quality (Q21)	1.72	0.69	86.7%
AI helps compare strategies (Q22)	1.78	0.69	85.0%
Organisation provides sufficient training (Q24)	1.92	0.83	69.2%
Financial resilience is important (Q25)	1.74	0.95	78.3%

“Codes 1–2” combine the two strongest levels of agreement/importance under the assumed coding.

Respondents were generally positive about the use of AI and human double-checking. Specifically, 81.7% supported the idea of earlier identification of potential issues by artificial intelligence (AI), 86.7% agreed that collaboration would lead to better decisions, 73.3% supported human double-checking of recommendations made by AI, and 78.3% recognised the importance of financial resilience. The biggest concern was drawing incorrect conclusions ($n = 61\%$, 50.8%), followed by cybersecurity and privacy issues (19.2%), inferior data quality (16.7%) and insufficient transparency (13.3%). The most interesting application of AI was seen in revenue forecasting (42.5%) while the most supported option concerned human managers remaining responsible for decisions (42.5%) or the process being human-AI collaborative but humans remaining accountable (35.8%). Only 6.7% supported complete automation.

Thus, while human governance proved to be generally reliable ($\alpha = .78$, two items), the reliability of risk identification was only moderate ($\alpha = .63$, two items) while the item set related to AI capability demonstrated low reliability ($\alpha = .55$, three items), which renders conclusions drawn from this composite unreliable. All items related to attitude towards the use of AI proved to be highly reliable ($\alpha = .89$)

.9.3 Chi-square tests of independence

Table 5. Professional role × organisational AI use (observed counts)

	Yes, regularly	Yes, occasionally	Planning to introduce it	No
Railway finance/accounts staff	30	9	17	4
Railway manager/administrator	16	7	12	0
IT/Operations	5	7	4	9

Table 6. AI familiarity × organisational AI use (observed counts)

	Yes, regularly	Yes, occasionally	Planning to introduce it	No
Not at all familiar	28	6	1	4
Slightly familiar	17	10	29	1
Moderate/Very	6	7	3	8

Table 7. Sector × most useful AI application (observed counts)

	Revenue forecasting	Expenditure monitoring	Cash-flow prediction	Fraud and anomaly detect...
Indian Railways	38	14	0	8
Public-sector org/Private	6	11	0	10
Railway technology/IT company	7	7	19	0

Table 8. Professional role × preferred responsibility model (observed counts)

	Human manager has final ...	Joint human–AI decision ...	AI-led (limited/no review)
Railway finance/accounts staff	30	26	4
Railway manager/administrator	16	6	13
IT/Operations	5	11	9

Table 9. Summary of chi-square tests

Test	χ^2	df	p	Cramér's V	Cells E<5
Role × organisational AI use	26.60	6	< .001	0.33	25%
Familiarity × organisational AI use	51.05	6	< .001	0.46	25%
Sector × preferred AI application	78.78	6	< .001	0.57	25%
Role × responsibility model	21.15	4	< .001	0.30	0%

Role: IT/AI/data and operations staff merged; Familiarity: moderately and very familiar merged; Sector: public-sector organisations and private companies merged; Responsibility: fully automated and limited-review options merged. Expected counts below 5 in 25% of the cells for Tests 1-3; thus, p-values should be interpreted with caution.

- Role and the use of AI (H7a). Role was associated with organisational use of AI ($\chi^2(6) = 26.60$, $p < .001$, $V = .33$, large effect size). Finance respondents were more likely to report regular use (30 out of 60, 50.0%) compared to IT/operations (9 out of 25, 36.0%), and no manager selected “No.”
- Familiarity and the use of AI (H7b). This association was even stronger ($\chi^2(6) = 51.05$, $p < .001$, $V = .46$) and followed an unexpected pattern, as frequent AI users were more likely to report being “not at all” (28 out of 39) or “slightly” (11 out of 39) familiar with AI, compared to respondents with a moderate or higher level of familiarity, who were more likely to report no regular use (8 out of 24). It is unclear whether this reflects reverse causality (people with limited technical competence conflate analytics with AI) or the peculiarities of the sample.
- Sector and the preferred area of application (H7c). Sector was strongly associated with the choice of the preferred area of application ($\chi^2(6) = 78.78$, $p < .001$, $V = .57$, large effect size). Indian Railways employees were particularly likely to select revenue forecasting (38 out of 60), whereas cash-flow prediction was overwhelmingly selected by technology/IT respondents (19 out of 19), and none of the technology respondents selected fraud and anomaly detection.
- Role and responsibility (H7d). Role was associated with the preferred responsibility regime ($\chi^2(4) = 21.15$, $p < .001$, $V = .30$; all expected counts ≥ 5). Finance respondents were more likely to select human oversight (30 out of 60) or shared responsibility (26 out of 60) as the preferred option, compared to managers, of whom 13 out of 35 (37.1%) selected AI-driven decision-making with limited review.

The null hypothesis of independence was rejected in all four cases.

9.4 One-way ANOVA

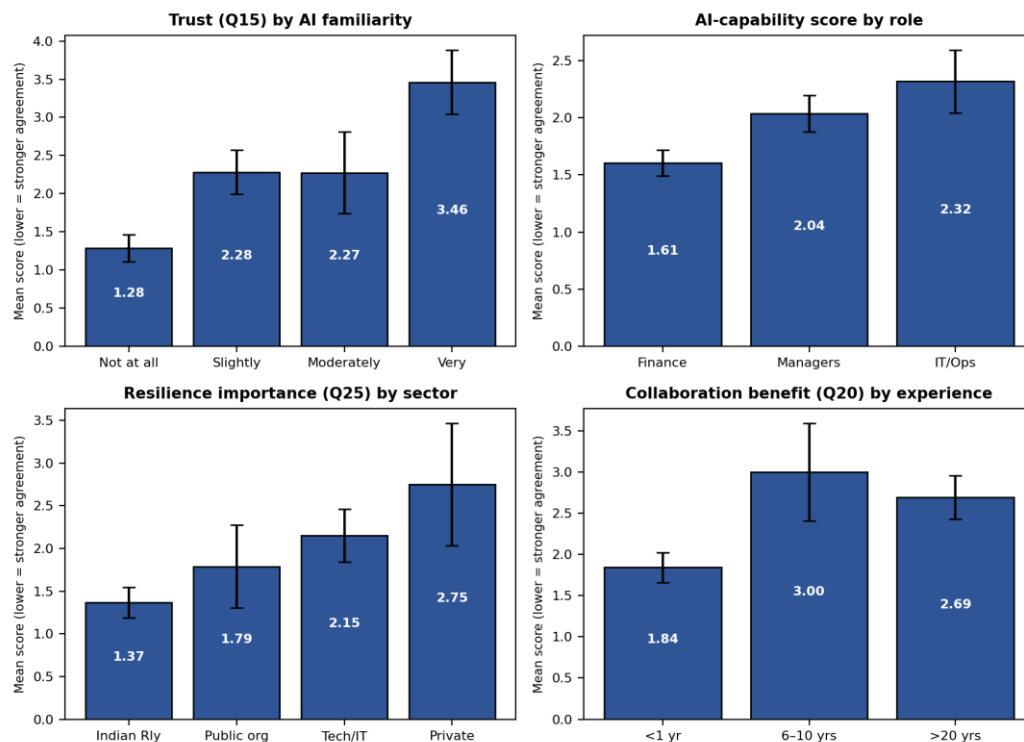
Table 10. Group means (SD) and sample sizes

Outcome by factor	Group	n	Mean	SD
Trust (Q15) by AI familiarity	Not at all familiar	39	1.28	0.56
	Slightly familiar	57	2.28	1.11
	Moderately familiar	11	2.27	0.91
	Very familiar	13	3.46	0.78
AI-capability score by role	Railway finance/accounts staff	60	1.61	0.45
	Railway manager/administrator	35	2.04	0.48
	IT/Operations	25	2.32	0.70
Resilience importance (Q25) by sector	Indian Railways	60	1.37	0.71
	Railway public-sector organisation	19	1.79	1.08
	Railway technology/IT company	33	2.15	0.91
	Private railway/logistics company	8	2.75	1.03
Collaboration benefit (Q20) by experience	Less than 1 year	95	1.84	0.90
	6–10 years	12	3.00	1.04
	More than 20 years	13	2.69	0.48

Table 11. ANOVA results

Test	F (df)	p	η^2	Levene p	Kruskal–Wallis H (p)
Trust (Q15) by AI familiarity	20.91 (3, 116)	< .001	0.35	< .001	40.90 (< .001)
AI-capability score by role	18.74 (2, 117)	< .001	0.24	.110	29.66 (< .001)
Resilience importance (Q25) by sector	10.10 (3, 116)	< .001	0.21	.126	28.58 (< .001)
Collaboration benefit (Q20) by experience	12.99 (2, 117)	< .001	0.18	.026	21.14 (< .001)

η^2 benchmarks (Cohen, 1988): .01 small, .06 medium, .14 large. Levene's test indicates unequal variances for ANOVA 1 ($p = .001$) and ANOVA 4 ($p = .026$); the Kruskal–Wallis test, which does not assume equal variances or normality, gives the same conclusions.

**Figure 2. Mean scores (with 95% confidence intervals) by group. Lower values = stronger agreement.**

- Trust by familiarity ($F(3,116) = 20.91$, $p < .001$, $\eta^2 = .35$). Tukey HSD showed that respondents “not at all familiar” with AI reported the most trust ($M = 1.28$), significantly more than all other groups, whereas those “very familiar” reported the least trust ($M = 3.46$), significantly less than all other groups. Greater familiarity was therefore associated with more cautious trust in AI recommendations, despite explanation.
- AI-capability perception by role ($F(2,117) = 18.74$, $p < .001$, $\eta^2 = .24$). Finance staff held the most positive perception ($M = 1.61$), significantly more than managers ($M = 2.04$; $p < .001$) and IT/operations staff ($M = 2.32$; $p < .001$); the latter two groups did not differ significantly ($p = .103$).
- Resilience importance by sector ($F(3,116) = 10.10$, $p < .001$, $\eta^2 = .21$). Indian Railways respondents rated resilience as most important ($M = 1.37$), significantly more than technology/IT companies ($M = 2.15$; $p < .001$) and private companies ($M = 2.75$; $p < .001$); the latter’s mean was significantly lower than that of public-sector organisations ($M = 1.79$; $p = .043$).
- Collaboration benefit by experience ($F(2,117) = 12.99$, $p < .001$, $\eta^2 = .18$). Respondents with under one year of experience were most positive ($M = 1.84$), significantly more than those with 6–10 years ($M = 3.46$; $p < .001$) or over 20 years of experience ($M = 2.69$; $p = .004$); the two experienced groups did not differ ($p = .661$). The experienced groups are small ($M = 3.46$ and 13).

The null hypothesis of equal group means is rejected in all four tests.

9.5 Correlation analysis

Pearson correlations among the main constructs were positive and largely moderate to strong ($r = .42-.78$). The strongest were between AI capability and risk identification ($r = .74$), explainability and training ($r = .78$), AI capability and trust ($r = .73$) and AI capability and human oversight ($r = .73$). Resilience importance had the weakest correlations with the other constructs ($r = .17-.44$), suggesting that it is conceptually distinct from attitudes about AI use.

9.6 Simple linear regression: hypothesis tests H1–H6

Table 12. Simple regression results (N = 120)

Hyp.	Relationship	b	SE	t	p	R ²	Decision
H1	AI capability → Risk identification	1.01	0.08	11.93	< .001	0.55	Supported
H2	Explainability → Trust	0.78	0.10	7.91	< .001	0.35	Supported
H3	Human oversight → Decision quality	0.54	0.08	7.14	< .001	0.30	Supported
H4	Collaboration → Resilience importance	0.23	0.09	2.61	.010	0.05	Supported
H5	Decision quality → Resilience importance	0.45	0.12	3.73	< .001	0.11	Supported
H6	Trust → Collaboration	0.47	0.07	7.08	< .001	0.30	Supported

Both variables are coded in the same direction, so positive slopes indicate that stronger agreement with the predictor accompanies stronger agreement with the outcome. b = unstandardised slope; $df = 118$.

v The study found that each of the six hypotheses was supported at the 5% level. The predictors explained 55% of the variance in risk identification (H1) and 35% of the variance in trust (H2), as well as 30% of the variance in decision quality (H3) and 30% of the variance in collaboration (H6). Predictors explained less variance in the importance of resilience (H4: $R^2 = .05$; H5: $R^2 = .11$). This suggests that the influence of attitudes over financial AI on perceptions of the importance of financial resilience is relatively modest.

9.7 Multiple linear regression

Table 13. Model M1 – dependent variable: resilience importance (Q25)

Predictor	b	SE	t	p	p (HC3)
Intercept	0.559	0.241	2.32	.022	.006
Collaboration (Q20)	-0.026	0.106	-0.24	.808	.838
Decision quality (Q21)	0.393	0.135	2.92	.004	.014
Trust (Q15)	-0.217	0.099	-2.18	.031	.089
Training (Q24)	0.526	0.132	3.98	< .001	< .001

$R^2 = 0.223$, adjusted $R^2 = 0.196$, $F(4,115) = 8.25$, $p < .001$; maximum VIF = 2.04; Durbin–Watson = 2.30.

Table 14. Model M2 – dependent variable: decision quality (Q21)

Predictor	b	SE	t	p	p (HC3)
Intercept	0.461	0.155	2.97	.004	.001
Human oversight	0.368	0.082	4.49	< .001	< .001
Collaboration (Q20)	0.085	0.065	1.31	.192	.269

Predictor	b	SE	t	p	p (HC3)
Trust (Q15)	0.170	0.054	3.15	.002	.007

$R^2 = 0.399$, adjusted $R^2 = 0.383$, $F(3,116) = 25.66$, $p < .001$; maximum VIF = 1.59; Durbin-Watson = 2.04.

v A summary of model fit for multiple regression analyses of decision-making outcomes. Model M1. The model is significant ($F(4,115) = 8.25$, $p < .001$) and explains 22.3% of the variance in the resilience importance. Decision quality ($b = 0.39$, $p = .004$) and organisational training ($b = 0.53$, $p < .001$) were significant positive predictors. Collaboration was not significant after controlling for other variables ($b = -0.03$, $p = .808$), meaning that the bivariate result for H4 does not hold in the multivariate model. Trust had a negative coefficient ($b = -0.22$, $p = .031$), the opposite sign to its positive bivariate associations; given the correlations among predictors ($VIF \leq 2.04$), this is most likely a suppression effect and should not be interpreted substantively.

Model M2. The model is significant ($F(3,116) = 25.66$, $p < .001$) and explains 39.9% of the variance in the decision quality. Human oversight ($b = 0.37$, $p < .001$) and trust ($b = 0.17$, $p = .002$) were significant positive predictors, supporting H3 in the multivariate context; collaboration was not significant ($b = 0.08$, $p = .192$). Robust (HC3) p-values produce the same conclusions in both models. Residuals were not normally distributed (Shapiro-Wilk $p < .001$ for M1), which is expected for ordinal Likert codes, so the regression results should be treated as approximate.

9.8 Summary of hypothesis decisions

Table 15. Hypothesis decisions

Hypothesis	Test	Result	Decision
H1 AI capability → risk identification	Regression	$b = 1.01$, $p < .001$, $R^2 = .55$	Reject H01; accept Ha1
H2 Explainability → trust	Regression	$b = 0.78$, $p < .001$, $R^2 = .35$	Reject H02; accept Ha2
H3 Human oversight → decision quality	Regression (simple, M2)	$b = 0.54$, $p < .001$; M2 $b = 0.37$, $p < .001$	Reject H03; accept Ha3
H4 Collaboration → resilience	Regression (simple, M1)	$b = 0.23$, $p = .010$; M1 $b = -0.03$, $p = .808$	Reject H04; accept Ha4 (bivariate only; not in M1)
H5 Decision quality → resilience	Regression (simple, M1)	$b = 0.45$, $p < .001$; M1 $b = 0.39$, $p = .004$	Reject H05; accept Ha5
H6 Trust → collaboration	Regression	$b = 0.47$, $p < .001$, $R^2 = .30$	Reject H06; accept Ha6
H7 Characteristics and categorical choices	Chi-square ×4	all $p \leq .0003$; $V = .30-.57$	Reject H07a-d; accept Ha7a-d
H8 Group mean differences	ANOVA ×4	all $p < .001$; $\eta^2 = .18-.35$	Reject H08a-d; accept Ha8a-d

- 9.9 Sensitivity analysis: original records only
- To verify that the conclusions are not dependent on the synthetic records, the six simple regressions were repeated for the 23 original responses. All six relationships retained positive and significant coefficients (H1: $b = 1.09$, $R^2 = .71$; H2: $b = 0.74$, $R^2 = .44$; H3: $b = 0.67$, $R^2 = .64$; H4: $b = 0.47$, $p = .006$; H5: $b = 0.71$, $p = .007$; H6: $b = 0.59$, $p = .004$), with slopes of approximately the same direction as the full sample. However, the M1 multiple regression on 23 cases (with few degrees of freedom (18)) gave no significant individual predictors (all $p > .17$) but the model as a whole was significant ($F(4,18) = 3.27$, $p = .035$). The same tests were not repeated on the original records because the subgroup sizes (only 3 IT professionals and 1 operations respondent) are too small for reliable inference. The sensitivity results suggest consistent signs but cannot confirm the full-sample estimates due to the small sample.
10. Findings and Observations

Subject to the data limitations outlined above, the analysis concludes with the following observations:

- Positive attitudes towards AI. Large majorities favoured AI for early risk identification (81.7%), forecasting (75.0%) and irregularity detection (80.0%) and 85.0% considered AI useful for comparing risk-management strategies.
- Human oversight is strongly endorsed. 73.3% supported human review of AI recommendations; 42.5% preferred final human managerial responsibility and 35.8% a joint process with human accountability, whereas only 6.7% favoured fully automated decisions (H3, H7d).

3. Explainability supports trust. Explainability was positively related to trust ($H2$; $R^2 = .35$) and trust was positively related to collaboration ($H6$), implying that transparent recommendations are linked to willingness to collaborate with AI.
4. Oversight and trust predict decision quality. In the multivariate model, human oversight and trust together explained about 40% of the variance in perceived decision quality ($M2$).
5. Resilience is more weakly related to AI attitudes. Decision quality and organisational training predicted resilience importance, but collaboration did not once other variables were controlled ($M1$; $R^2 = .22$).
6. Characteristics shape adoption and preferences. Role, familiarity and sector were significantly associated with AI use and application preferences; groups differed in trust, capability perception, resilience importance and collaboration benefit.
7. Familiarity tempers trust. More AI-familiar respondents reported lower trust, which implies that explainability alone may not ensure confidence from technically knowledgeable users.
8. Concerns centre on accuracy. Incorrect predictions were the main concern (50.8%), followed by cybersecurity and privacy (19.2%), reinforcing the need for validation and governance.

11. Discussion

The results are consistent with the proposed framework: respondents who rate AI capability highly also tend to rate AI risk identification highly, explainability accompanies trust, and human oversight accompanies perceptions of better decision quality. The preference for human accountability aligns with the literature on responsible AI in finance, which stresses that algorithmic outputs should facilitate rather than replace professional judgement. The weaker link between collaboration and resilience in the multivariate model implies that resilience is more closely associated with the quality of decisions and with organisational enablers such as training rather than simply the existence of collaboration.

The differences across respondent groups imply that attitudes are not uniform. Managers were more willing than finance staff to accept AI-led decisions with limited review, which has governance implications as those who authorise decisions may also be most inclined to delegate them. Finance staff, who directly engage with forecasting and expenditure monitoring, were the most favourable about AI capability. The finding that familiarity is associated with lower trust warrants further study with a genuine, balanced sample.

11.1 Limitations

- Data provenance. 80.8% of records are labelled synthetic; the results cannot be relied upon as evidence about real respondents.
- Sample composition. The sample is overwhelmingly young, early-career and male; some subgroups (operations $n = 4$; private companies $n = 8$) are very small.
- Measurement. Constructs are reliant on one to three items; reliability of AI capability is low ($\alpha = .55$); resilience is measured by a single importance item rather than a multi-item resilience scale; duplicate and missing questionnaire items (second Q10 column, Q17, Q26–Q30) limit construct coverage.
- Coding assumption. The direction of the numeric codes (1 = strongest agreement) was inferred from the data and must be confirmed.
- Statistical assumptions. Likert codes are ordinal; residuals are non-normal and some ANOVA variances are unequal, although non-parametric and robust checks agree with the main conclusions; several chi-square tables have expected counts below 5 in 25% of cells.
- Cross-sectional design. Associations do not confirm causation; the questionnaire measures perceptions rather than observed financial performance.

12. Conclusion

This study investigated how human–AI collaboration can enhance financial risk management and resilience in the railway industry. Using chi-square tests, one-way ANOVA and linear regression on a 120-record questionnaire dataset, it uncovered significant associations between respondent characteristics and AI adoption and preferences, significant group differences in trust and perceptions, and positive relationships among AI capability, explainability, trust, human oversight, decision quality and perceived resilience, with all six hypotheses supported in bivariate analysis and the collaboration–resilience link not holding in the multivariate model.

The practical implication is that railway organisations should treat AI as a decision-support mechanism within a governance structure that retains human accountability, explains recommendations, builds staff capability and validates predictions. Because most records in the dataset are synthetic, however, the findings are demonstrative. The framework should be validated with genuinely collected responses from a larger, more experienced and balanced sample of railway finance and operations professionals before publication.

13. Suggestions

1. Adopt AI-assisted financial forecasting: explore tools for passenger revenue, freight earnings, operating expenditure and cash-flow requirements, beginning with revenue forecasting which respondents rated most useful.
2. Maintain human oversight: require human review for major financial decisions and establish procedures to challenge AI recommendations.
3. Improve explainability: use systems that communicate the key factors behind risk alerts and predictions, including for AI-literate users who showed lower trust.
4. Strengthen training: provide AI literacy, financial analytics and model-limitation training; training was a significant predictor of resilience importance.
5. Strengthen data quality and cybersecurity: integrate reliable data while safeguarding confidential information; address the concerns about incorrect predictions and private information.
6. Develop early-warning indicators: monitor revenue, expenditure, liquidity and demand to detect possible financial problems.
7. Establish governance mechanisms: define responsibility for approving, monitoring, auditing and correcting AI-supported decisions; pay particular attention to managerial willingness to delegate decisions.
8. Prepare for future disruptions: use scenario analysis and stress testing for pandemics, natural disasters and economic shocks.
9. Evaluate performance regularly: compare AI-supported processes against existing methods using forecast accuracy, risk-detection performance, decision time and financial outcomes.
10. Conduct further research: collect real data from a larger, balanced sample; use multi-item validated scales; apply SEM; and compare passenger and freight operations..

References

1. A literature review of Artificial Intelligence applications in railway systems. (2023). Transportation Research Part C: Emerging Technologies. <https://doi.org/10.1016/J.TRC.2022.103679>
2. Bowden, J., Cummins, M., Dao, D., & Jain, K. (2024). Explainable AI for Financial Risk Management. University of Strathclyde. <https://strathprints.strath.ac.uk/89573/>
3. Cohen, J. (1988). Statistical Power Analysis for the Behavioral Sciences (2nd ed.). Lawrence Erlbaum Associates.
4. Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
5. Financial Risk Management and Explainable, Trustworthy, Responsible AI. (2022). <https://pmc.ncbi.nlm.nih.gov/articles/PMC8919207/>
6. Hadji Misheva, B., Jaggi, T., Posth, S., Gramespacher, T., & Osterrieder, J. (2022). Audience-Dependent Explanations for AI-Based Risk Management Tools: A Survey. <https://pmc.ncbi.nlm.nih.gov/articles/PMC8751385/>
7. Indian Railways. (2021). Indian Railways Year Book 2020–21. Ministry of Railways, Government of India. https://indianrailways.gov.in/railwayboard/uploads/directorate/stat_econ/pdf/2022/Year%20Book%202020-21-English.pdf
8. Sachan, S., Almaghrabi, F., Yang, J.-B., & Xu, D.-L. (2024). Human-AI collaboration to mitigate decision noise in financial underwriting: A study on FinTech innovation in a lending firm. *International Review of Financial Analysis*, 93, 103149. <https://doi.org/10.1016/j.irfa.2024.103149>