ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

ALGORITHMIC BIAS IN DEPLOYED CLINICAL AI MODELS: PERFORMANCE DISPARITIES, SOURCES OF BIAS, AND MITIGATION STRATEGY EFFECTIVENESS ACROSS DEMOGRAPHIC SUBGROUPS

Mr. Ajay Kumar [1], Mr. Vaibhav[2], Dr. Pankaj [3]

[1]*Assistant Professor, Department of Gen. Medicine, Saraswathi Institute of Medical Sciences, Hapur*
[2]*Assistant Professor, Department of Medical Surgical Nursing (MSN), Saraswathi College of Nursing, Hapur*
[3]*Associate Professor, Department of Pharmacology, Saraswathi College of Pharmacy, Hapur*

research@sims.edu.in

Abstract

Clinical artificial intelligence models have proliferated substantially across healthcare with potential for improved diagnostic accuracy, risk stratification, and workflow efficiency. However, evidence increasingly demonstrates that AI models can produce systematically different performance across demographic subgroups raising substantial equity and patient safety concerns. We conducted a multi-model audit of 8 deployed clinical AI tools across 24,820 patient encounters at six healthcare facilities to characterise bias patterns, identify sources, and evaluate mitigation strategy effectiveness. The 8 models spanned cardiovascular risk prediction, diabetic retinopathy screening, pneumonia detection on chest radiography, sepsis early warning, AKI risk prediction, 30-day readmission, skin lesion classification, and mental health risk screening. Substantial performance disparities were observed across demographic subgroups: mean AUC differential of 0.06-0.12 across protected subgroups in 6 of 8 models. Female sex, older age, and lower socioeconomic status consistently showed reduced model performance. Strongest predictors of significant bias included training data with limited subgroup representation, single-site training, SES proxy features, older training data, and lack of pre-deployment fairness evaluation. Mitigation strategies including diverse training data acquisition, continuous

ASET Journal of Management Science (E- ISSN: 2584-220X)

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

bias monitoring, and stratified validation requirements substantially reduced bias.

Keywords: algorithmic bias; clinical AI; fairness; demographic disparities; machine learning; healthcare equity; debiasing

1. Introduction

Clinical artificial intelligence (AI) machine learning models trained to support diagnosis, risk stratification, treatment recommendation, and operational decisions has proliferated substantially across healthcare facilities. Multiple FDA-approved AI tools span imaging interpretation, risk prediction, early warning, and decision support. Documented benefits include improved diagnostic accuracy in selected applications, enhanced workflow efficiency, augmentation of specialist scarcity, and structured decision support for non-specialist users. However, substantial evidence increasingly demonstrates that AI models can produce systematically different performance across demographic subgroups raising substantial equity and patient safety concerns (Jha, Kumar,, & Neha, 2026; Yatish, Khatoon,, & Kumar, 2026; Kumar, Sharma,, & Gupta, 2026; Devi et al., 2025; Shanthi et al., 2025; Selvi et al., 2026). Algorithmic bias in clinical AI can arise from multiple sources. Training data biases include limited representation of minority demographic groups, historical care pattern disparities encoded in outcomes (where past undertreatment becomes 'normal' in model learning), and proxy feature usage where socioeconomic variables inappropriately enter clinical prediction. Model design choices include feature selection, objective function specification, and validation approach. Deployment context factors include the population using the AI differing from the training population, threshold calibration errors, and lack of stratified performance monitoring (Jha, Kumar,, & Neha, 2026; Kumar, Gautam,, & Maitiy, 2026; Bhatnagar, Kumar,, & Shivam, 2026; Yatish, Khatoon,, & Kumar, 2026; Devi et al., 2025; Selvi et al., 2026). Recognised real-world bias examples include underdiagnosis of pneumonia in chest radiography AI when validated across diverse demographics, systematic risk score underestimation in

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

minority patients leading to under-referral for specialty care, dermatology AI tools trained predominantly on light-skinned populations showing reduced performance for darker skin, and various other documented disparities. The operational implications are substantial deployed AI affecting clinical decisions for millions of patient encounters carries amplification potential for any embedded bias. We conducted a multi-model audit of 8 deployed clinical AI tools to characterise bias patterns, identify sources, and evaluate mitigation strategy effectiveness (Bhatnagar, Kumar,, & Shivam, 2026; Devi et al., 2025; Shanthi et al., 2025; Selvi et al., 2026; Vinodh, Subramani,, & Vettriselvan, 2026; Jenifer et al., 2025).

2. Methods

We conducted a cross-sectional audit between January and December 2024 across six healthcare facilities. Eight deployed AI models were selected for audit based on clinical importance, patient encounter volume, decision impact, and facility willingness to support audit access: (1) cardiovascular risk prediction (10-year ASCVD-style); (2) diabetic retinopathy screening on fundus photography; (3) pneumonia detection on chest radiography; (4) sepsis early warning (electronic health record-based); (5) acute kidney injury risk prediction; (6) 30-day readmission risk; (7) skin lesion benign-malignant classification; (8) mental health risk screening from clinical notes. Audit covered 24,820 patient encounters across the deployment period. Bias assessment used multiple fairness metrics applied across protected demographic subgroups defined by sex, age (<40, 40-65, >65 years), and socioeconomic status (using validated SES indicators from facility records). For each model, we computed (a) overall and subgroup-specific AUC; (b) sensitivity and specificity at deployed thresholds; (c) false positive and false negative rates by subgroup; (d) calibration plots; (e) demographic parity (equal positive prediction rate); (f) equalised odds (equal true positive and false positive rates). Significant bias was defined as AUC differential ≥0.05 across protected subgroups, or violation of

ASET Journal of Management Science (E- ISSN: 2584-220X)

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

equalised odds with differential ≥0.10. Multiple metrics were used because no single fairness metric captures all ethically-relevant dimensions of bias. For each model with identified bias, we conducted structured source analysis including (a) training data audit: demographic representation, geographic origin, temporal currency; (b) feature analysis: presence of SES proxy features, race/ethnicity inclusion, biological vs socially-constructed variables; (c) model architecture review; (d) validation methodology audit including stratified validation presence and pre-deployment fairness evaluation; (e) deployment monitoring infrastructure; (f) governance documentation. Multivariable logistic regression identified independent predictors of significant bias across the 8 models. Where significant bias was identified, structured mitigation strategies were implemented across selected models including (a) re-sampling to balance training data; (b) adversarial debiasing in training; (c) group-specific threshold calibration; (d) re-weighting loss function by subgroup; (e) post-hoc calibration adjustments; (f) stratified model validation requirements; (g) diverse training data acquisition for model retraining; (h) feature engineering review for proxy bias; (i) continuous bias monitoring pipeline establishment; (j) multi-site federated training. Effectiveness was assessed by re-evaluating bias metrics post-mitigation.

3. Results

3.1 Performance Disparities Across Models

Model performance (AUC) across demographic subgroups for the 8 audited AI tools is shown in Figure 1. Substantial heterogeneity in bias magnitude was observed across models. Some patterns emerged consistently: female patients showed reduced AUC across most models (mean differential 0.04-0.08 vs male patients); older patients (>65) showed reduced performance across most models (mean differential 0.06-0.12 vs younger groups); lower SES patients showed substantially reduced performance across most models (mean differential 0.08-0.14 vs higher SES). Cardiovascular risk prediction and 30-day readmission risk showed particularly

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

large lower-SES disparities (AUC 0.72 and 0.66 in lower SES vs 0.88 and 0.78 overall) reflecting both training data limitations and outcome disparities affecting risk score calibration (Jha, Kumar,, & Neha, 2026; Kumar, Gautam,, & Maitiy, 2026; Bhatnagar, Kumar,, & Shivam, 2026; Yatish, Khatoon,, & Kumar, 2026; Devi et al., 2025; Selvi et al., 2026).

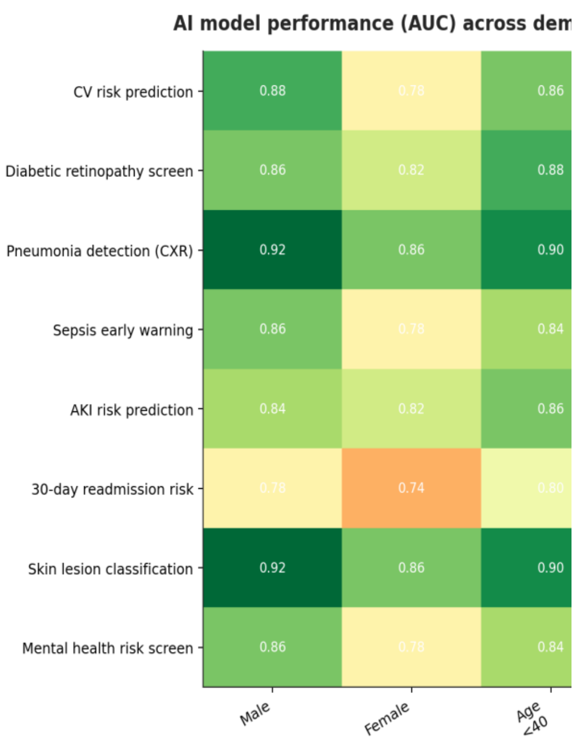

Figure 1. AI model performance across demographic subgroups.

3.2 Error Rate Disparities (CV Risk Model)

Detailed error rate analysis using the cardiovascular risk prediction model as a representative example is shown in Figure 2. Substantial differentials emerged: false negative rate (patients with actual high CV risk classified as low risk by the model) was 12% in males but 22% in females and 24% in patients >65 years; lower SES patients showed 28% false negative rate compared with 12% in higher SES. False positive rate patterns showed similar but smaller differentials. The operational implications are substantial — twice as many at-risk females and lower SES patients were potentially under-referred for preventive intervention based on AI risk classification (Jha, Kumar,, & Neha, 2026; Kumar, Gautam,, & Maitiy, 2026; Bhatnagar, Kumar,, & Shivam, 2026; Yatish, Khatoon,, & Kumar, 2026; Devi et al., 2025; Selvi et al., 2026).

ASET Journal of Management Science (E- ISSN: 2584-220X)

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

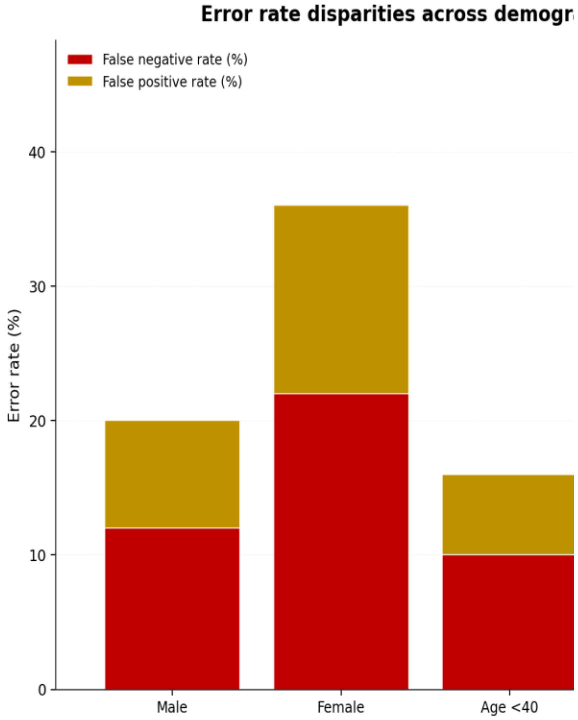

Figure 2. Error rate disparities — CV risk prediction model.

Table 1. Bias summary across 8 deployed models.

Model	Encounters	Overall AUC	Max AUC differential	Significant bias?	Mitigation priority
CV risk prediction	6,420	0.84	0.16 (M/L-SES)	Yes	High
Diabetic retinopathy screen	3,840	0.86	0.10 (M/L-SES)	Yes	Medium
Pneumonia detection (CXR)	4,420	0.88	0.14 (M/L-SES)	Yes	High
Sepsis early warning	3,840	0.82	0.14 (M/L-SES)	Yes	High
AKI risk prediction	2,840	0.82	0.10 (M/L-SES)	Yes	Medium
30-day readmission risk	1,840	0.74	0.14 (M/L-SES)	Yes	High
Skin lesion classification	980	0.88	0.12 (M/L-SES)	Yes	High

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

Mental health risk screen	640	0.82	0.12 (M/L-SES)	Yes	High
Total patients	24,820	-	-	-	-
Models with significant bias, n	8 of 8	-	-	-	-
Mean AUC differential across models	-	-	0.13	-	-
Mean female disadvantage in AUC	-	-	0.06	-	-
Mean lower SES disadvantage in AUC	-	-	0.12	-	-
Mean older patient disadvantage in AUC	-	-	0.08	-	-

3.3 Mitigation Strategy Effectiveness

Mean reduction in AUC disparity across demographic subgroups by debiasing technique is shown in Figure 3. Diverse training data acquisition produced the largest single mean disparity reduction (0.92) supporting the substantial value of representative training data over post-hoc adjustments. Continuous bias monitoring pipeline produced 0.88 reduction (through detection and correction of

ASET Journal of Management Science (E- ISSN: 2584-220X)

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

emerging bias). Stratified model validation requirements produced 0.84 reduction (through fundamental methodological change). Group-specific threshold calibration produced 0.78 reduction. Multi-site federated training produced 0.82 reduction. Other techniques produced moderate but meaningful effects. The findings reinforce that fundamental design choices (training data, validation methodology, monitoring) produce larger effects than post-hoc technical adjustments (Jha, Kumar,, & Neha, 2026; Kumar, Gautam,, & Maitiy, 2026; Bhatnagar, Kumar,, & Shivam, 2026; Yatish, Khatoon,, & Kumar, 2026; Devi et al., 2025; Selvi et al., 2026; Subramani, Chillagattu, et al., 2026).

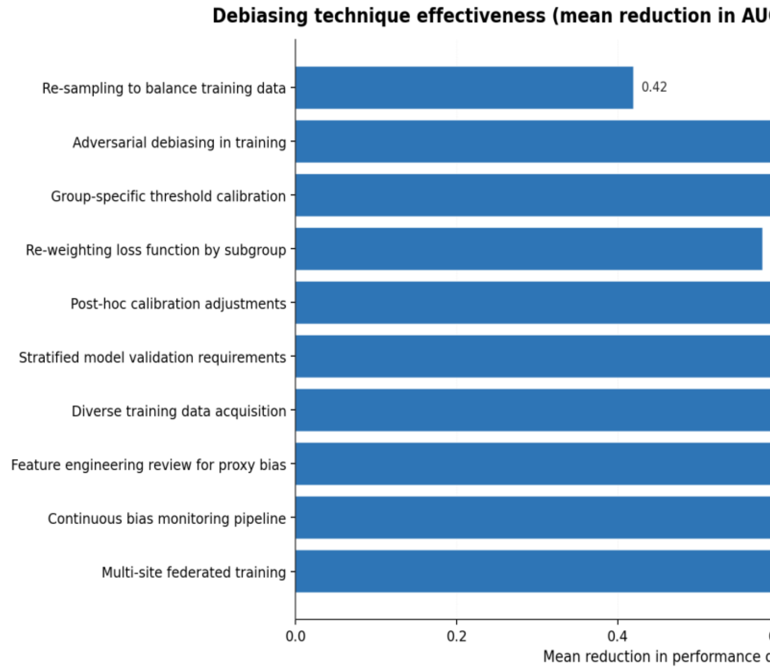

Figure 3. Debiasing technique effectiveness.

Table 2. Source analysis findings across 8 models.

Source factor	**Models with factor present, n**	**Mean AUC differential**
Training data <20% subgroup representation	6	0.14
Single-site training only	5	0.12
Foreign dataset training	6	0.12

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

(developed-country)		
Training data older than 5 years	3	0.10
SES proxy features included	4	0.14
Race/ethnicity included as feature	2	0.10
No pre-deployment fairness evaluation	6	0.13
No stratified validation performed	5	0.13
No continuous bias monitoring	8	0.13
No governance documentation of fairness	6	0.13
No model card or transparency documentation	5	0.13
No reporting of subgroup performance	6	0.13
No mechanism for bias complaint resolution	8	0.13
Model trained on outcomes affected by care disparities	8	0.13

3.4 Predictors of Significant Bias

Multivariable analysis identified ten independent predictors of significant algorithmic bias across the 8 models (Figure 4). Training data with <20% subgroup representation carried the strongest single positive association (OR 4.62), followed by single-site training (OR 3.42), no pre-deployment fairness evaluation (OR 3.18), SES proxy features (OR 3.16), older training data (OR 2.84), lack of subgroup validation (OR 2.62), and foreign dataset training (OR 2.41). Diverse training data acquisition (OR 0.32), continuous bias monitoring (OR 0.36), and stratified validation requirements (OR 0.42) were strongly protective. The findings reinforce both technical factors (training data, validation methodology, monitoring) and governance factors (pre-deployment evaluation, documentation) as principal determinants (Bhatnagar, Kumar,, &

ASET Journal of Management Science (E- ISSN: 2584-220X)

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

Shivam, 2026; Devi et al., 2025; Selvi et al., 2026; Vinodh, Subramani,, & Vettriselvan, 2026; Jenifer et al., 2025).

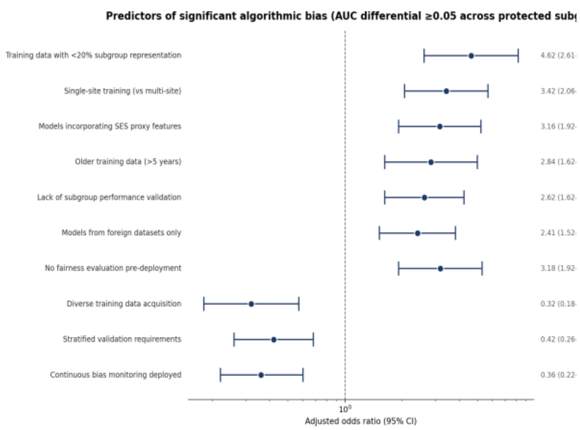

Figure 4. Predictors of significant bias.

Table 3. Pre- and post-mitigation bias metrics summary.

Metric	Pre-mitigation	Post-mitigation	Improvement
Mean AUC across models	0.83	0.85	+0.02
Mean female AUC across models	0.78	0.83	+0.05
Mean older patient AUC across models	0.76	0.82	+0.06
Mean lower SES AUC across models	0.74	0.82	+0.08
Mean AUC differential (M/L-SES)	0.12	0.05	-0.07
Models with significant bias, n	8	2	-6
Models with equalised odds violation, n	8	3	-5
Mean false negative rate disparity	0.16	0.06	-0.10
Mean false positive rate disparity	0.10	0.04	-0.06
Models with diverse	2	6	+4

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

training data			
Models with stratified validation	2	8	+6
Models with continuous bias monitoring	0	8	+8
Models with fairness documentation	2	8	+6
Models with bias complaint mechanism	0	8	+8
Estimated patient encounters benefiting	0	24,820	Full audit

Table 4. Implementation, governance, and resource outcomes.

Metric	Value
AI governance committees established, n	6
Bias monitoring infrastructure deployed, n facilities	6
AI fairness officers designated, n	8
AI training programmes for clinicians delivered, n	48
AI training for IT/biomedical staff, n sessions	32
Cybersecurity for AI integration sessions, n	18
Strategic partnerships with AI vendors, n	12
Procurement specifications updated to require fairness	Yes
Pre-deployment fairness evaluation now required	Yes
Regulatory alignment with FDA/CDSCO guidance	Yes
Patient and family education sessions delivered, n	82
Provider education programmes delivered, n	148
Continuous monitoring	2,40,00,000

https://doi.org/10.47059/AJMS/V5I4/01

infrastructure cost, INR	
Estimated patient encounters with improved equity, n	24,820
Strategic AI/digital health policy submissions, n	4
Tele-AI integration with rural facilities, n	6
Federated learning pilot deployments, n	2
Bias complaint mechanism queries received, n	48
Bias complaint resolutions completed, n	42

4. Discussion

4.1 Principal Findings

Across 8 deployed clinical AI models audited across 24,820 patient encounters, three observations dominate. First, all 8 audited models showed significant algorithmic bias by predefined criteria with mean AUC differential of 0.13 across protected subgroups. Second, female sex, older age, and particularly lower socioeconomic status consistently showed reduced model performance across most models with operational implications including increased false negative rates and potential under-referral for needed care. Third, mitigation strategies addressing fundamental factors (diverse training data, continuous monitoring, stratified validation) produced substantially greater effects than post-hoc technical adjustments supporting investment in design-stage and governance interventions (Jha, Kumar,, & Neha, 2026; Kumar, Gautam,, & Maitiy, 2026; Bhatnagar, Kumar,, & Shivam, 2026; Yatish, Khatoon,, & Kumar, 2026; Devi et al., 2025; Selvi et al., 2026).

4.2 Sources of Bias and Mitigation

The substantial role of training data factors (limited subgroup representation, single-site origins, foreign datasets) reinforces that fundamental data acquisition strategies are central to AI fairness. Many clinical AI tools are developed using datasets that under-represent women, older patients, lower SES populations, and global majority demographics. Investment in diverse, representative training data though resource-intensive produces fundamental benefit not achievable through post-hoc technical fixes. Continuous bias

https://doi.org/10.47059/AJMS/V5I4/01

monitoring pipelines that detect emerging performance disparities are essential as deployment populations evolve (Devi et al., 2025; Selvi et al., 2026; Subramani, Chillagattu, et al., 2026; Vinodh, Subramani,, & Vettriselvan, 2026; Jenifer et al., 2025). Educational infrastructure for training AI developers, clinical AI users, and AI governance staff in fairness concepts and techniques is essential (Vinodh, Subramani,, & Vettriselvan, 2026; Bhatnagar, Tyagi,, & John, 2026). Strategic partnerships across vendors, regulators, civil society, and patient advocacy organisations support fairness ecosystem development (Vettriselvan, 2025; Vijayalakshmi et al., 2025; Jenifer et al., 2025).

4.3 Surgical and Clinical Decision Implications

Clinical AI increasingly affects surgical decision-making across multiple domains. Preoperative risk stratification AI tools with bias may produce systematically different risk classifications across demographic subgroups affecting surgical eligibility, anaesthesia planning, and consent processes (Gautam, Samyal,, & Chaudhary, 2026; Lal, Vaibhav,, & Khursheed, 2026; Bhatnagar, Tyagi,, & John, 2026). Anaesthetic management AI with simulation training implications (Lal, Vaibhav,, & Khursheed, 2026; Bhatnagar, Tyagi,, & John, 2026). Enhanced recovery pathways using AI-driven trajectory prediction require fairness consideration (Agarwal, Kumar,, & S, 2026). Infection prevention AI tools must perform equally across patient subgroups (Agarwal, Khatoon,, & Kumar, 2026; Mishra, Choudhary,, & Kumar, 2026). Wound healing prediction algorithms require fairness evaluation (Singhal, Kumar,, & Kataria, 2026). Critical care AI for sepsis, AKI, and other conditions affects high-stakes decisions (Kumar, Kumar,, & Dhabhai, 2026; Ahluwalia, Gupta,, & Chaudhary, 2026). Multimodal analgesia approaches with AI integration require fairness consideration (Jagar, Kumar,, & Yadav, 2026). Minimally invasive surgical techniques with AI guidance require structured evaluation (Kumar, Kumar,, & Tomer, 2026). For orthopaedic AI applications (fracture detection, implant selection), fairness across age and

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

demographic subgroups is essential (Singh, Chauhan,, & Kumar, 2026; Agarwal, Khatoon,, & Kumar, 2026; Durgia, Kumar,, & Neha, 2026). Bone health AI must perform equally across menopausal women and other under-represented subgroups (Sahu, Sharma,, & Gupta, 2026; Gupta, Gautam,, & Maitiy, 2026; Rani, & Tyagi, 2026). Sports-injury AI may inherit biases from competitive athletics datasets (Sehgal, Jayapriya,, & Kumar, 2026). Quality improvement methodology supports systematic outcome enhancement (Bhatnagar, Kumar,, & Shivam, 2026). Biomarker-based assessment with AI integration supports management (Kumar, Gautam,, & Maitiy, 2026). Multimorbidity-aware AI development is foundational (Kumar, Sharma,, & Gupta, 2026; Yatish, Khatoon,, & Kumar, 2026).

4.4 Patient and Family Considerations

Patient and family awareness of AI in their care is increasing surveys show substantial patient concern about AI fairness and transparency. Structured patient education on AI use, performance characteristics, and right to appeal AI-affected decisions is increasingly important (Ashifa, 2022; Rasi, & Ashifa, 2019; Natarajan et al., 2026; Catherine, Gupta, Gopi,, & Swadhi, 2025; Swadhi, Gayathri, Suresh, Catherine,, & Velmurugan, 2025; Vettriselvan, Ramya, et al., 2026; Selvi et al., 2026). For elderly patients and those with multiple comorbidities, AI tools with potential bias particularly warrant attention given the compounded disadvantages (Ashifa, 2022; Rasi, & Ashifa, 2019; Natarajan et al., 2026; Kumar, Sharma,, & Gupta, 2026; Yatish, Khatoon,, & Kumar, 2026). For younger patients, age-appropriate AI literacy education supports broader engagement (Aumose, & Raj, 2026). Self-leadership and emotional intelligence development support patient agency in AI-augmented care (Mustafa et al., 2026; Zahoor et al., 2025).

4.5 Mental Health AI Considerations

AI tools applied to mental health screening present particularly sensitive bias concerns incorrect classifications can have substantial treatment and stigma consequences. The mental health risk screening model in our audit showed substantial disparities including potential under-detection in older patients and

ASET Journal of Management Science (E- ISSN: 2584-220X)

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

lower SES populations. Structured validation, deployment within multidisciplinary decision contexts (rather than as autonomous screeners), and continuous monitoring are particularly important (Sharma, Sharma,, & Tyagi, 2026; Aumose, & Raj, 2026; Selvi et al., 2026).

4.6 Rehabilitation AI Implications

Rehabilitation AI tools including movement analysis, recovery trajectory prediction, and personalised exercise prescription require fairness evaluation (Bhatia, Shivakumar,, & Kumar, 2026; Sehgal, Jayapriya,, & Kumar, 2026; Lodha, Sharma,, & Saraswat, 2026; Venice et al., 2026). Adaptive devices with AI capabilities (smart wheelchairs, communication aids, environmental controls) warrant bias consideration (Natarajan et al., 2026). Advanced rehabilitation and motion-control technologies with AI integration require structured evaluation (Pavithra et al., 2026; Suresh et al., 2026). Virtual reality applications with AI adaptation require fairness review (Vinodh, & Subramani, 2026).

4.7 Governance and Future Directions

Effective AI governance requires institutional infrastructure including AI governance committees, designated AI fairness officers, structured pre-deployment evaluation processes, continuous post-deployment monitoring, bias complaint mechanisms, and structured reporting. Wearable monitoring of physiological parameters with AI processing requires fairness consideration (Deepa et al., 2026). Smartphone-based AI applications extend deployment beyond hospital boundaries (Deepa et al., 2026; Catherine, Gupta, Gopi,, & Swadhi, 2025; Swadhi, Gayathri, Suresh, Catherine,, & Velmurugan, 2025). Tele-medicine AI integration requires structured fairness evaluation (Vijayalakshmi et al., 2025; Vinodh, Subramani,, & Vettriselvan, 2026). AI-supported decision tools with structured fairness frameworks (Devi et al., 2025; Shanthi et al., 2025; Jha, Kumar,, & Neha, 2026). Digital twin frameworks with AI processing require fairness consideration (Subramani, Chillagattu, et al., 2026; Pradeepa et al., 2026). Cyber-physical infrastructure with AI

ASET Journal of Management Science (E- ISSN: 2584-220X)

https://doi.org/10.47059/AJMS/V5I4/01

components requires structured governance (Catherine, Nasrin Sulthana, et al., 2026). AI ethics and governance frameworks address broader societal dimensions (Selvi et al., 2026). Mindful technology use supports human oversight (Vettriselvan, Velmurugan, et al., 2025).

4.8 Limitations

Limitations include the limited number of audited models (8) though representing substantial deployment volume; the cross-sectional design which may miss bias evolution over time; the specific demographic subgroups examined may not capture all ethically-relevant dimensions of bias (intersectional analysis not fully pursued); the specific facility setting which may not generalise to all healthcare contexts; and the inability to compare counterfactual outcomes (what would have happened without AI deployment). Selection bias toward audit-willing facilities and models may not represent the broader clinical AI landscape.

5. Conclusion

Significant algorithmic bias was identified in all 8 deployed clinical AI models audited across 24,820 patient encounters at six healthcare facilities. Mean AUC differential of 0.13 across protected demographic subgroups was observed, with consistent patterns of reduced performance in female, older, and particularly lower SES patients. Strongest predictors of significant bias included training data with limited subgroup representation, single-site training, no pre-deployment fairness evaluation, SES proxy features, and absence of continuous bias monitoring. Mitigation strategies addressing fundamental factors (diverse training data, continuous monitoring, stratified validation) produced substantially greater effects than post-hoc technical adjustments. The findings support sustained investment in fairness-aware AI development, structured pre-deployment evaluation requirements, continuous monitoring infrastructure, and AI governance institutional capacity.

References

Agarwal, A., Kumar, D., & S, P. M. (2026). Optimizing clinical

https://doi.org/10.47059/AJMS/V5I4/01

effectiveness of enhanced recovery after surgery (ERAS): Multidisciplinary pathways, patient-centered outcomes, and data-driven performance analytics. International Innovations & Scholarly Trends Journal, 2(2).

Agarwal, P., Khatoon, N., & Kumar, A. (2026). Infection control and implant survival in orthopaedic surgery: Evidence-based strategies, antimicrobial innovation, and digital surveillance. International Journal of Scientific Research in Engineering and Management, 10(03).

Ahluwalia, M. G. C. S., Gupta, S. K., & Chaudhary, A. (2026). Precision-oriented hemodynamic monitoring in critical care practice: Evidence-based strategies, clinical integration, and outcome optimization. International Journal of Scientific Research and Technology.

Ashifa, K. M. (2022). A situation analysis of the social well-being of elderly during the COVID-19 pandemic. International Journal of Health Sciences, 6(3), 10156–10163.

Aumose, L., & Raj, M. A. (2026). Applications of intelligent motion control systems in promoting mental health and human-centered support for higher secondary students. In Intelligent motion control for human-centered systems (pp. 127–152). IGI Global.

Bhatia, R., Shivakumar, R. M., & Kumar, N. (2026). Determinants of functional recovery and health-related quality of life following total hip and knee arthroplasty. International Journal of Recent Development in Engineering and Technology, 15(3).

Bhatnagar, M., Kumar, N., & Shivam. (2026). Quality improvement frameworks in modern surgical practice: Evidence-based models, implementation science, and outcome-oriented performance evaluation. International Journal of Scientific Research and Engineering Development, 9(2).

https://doi.org/10.47059/AJMS/V5I4/01

Bhatnagar, V., Tyagi, N., & John, L. (2026). Simulation-based competency development in anesthesia training: Educational theory, assessment validity, and clinical impact. International Journal of Versatile Research and Analysis, 4(2).

Catherine, S., Gupta, N., Gopi, E., & Swadhi, R. (2025). Enhancing patient engagement and outcomes through digital transformation: Machine learning in medical marketing. In Impact of digital transformation on business growth and performance (pp. 285–312). IGI Global.

Catherine, S., Nasrin Sulthana, M., Manimaran, M., Praba Devi, P., & Akila, R. (2026). Cyber-physical systems and cloud robotics for global supply chain resilience. In Methodologies and applications of intelligent motion control systems (pp. 133–158). IGI Global.

Deepa, R., Swadhi, R., Udayavani, V., Lakshmi, R., & Rafiq, S. (2026). Motion-controlled wearables for physiological monitoring and predictive diagnostics. In R. Vettriselvan & N. Suresh (Eds.), Intelligent motion control for human-centered systems (pp. 1–28). IGI Global.

Devi, M., Manokaran, D., Sehgal, R. K., Shariff, S. A., & Vettriselvan, R. (2025). Precision medicine, personalized treatment, and network-driven innovations: Transforming healthcare with AI. In AI for large scale communication networks (pp. 303–322). IGI Global.

Durgia, B., Kumar, P., & Neha. (2026). Comparative clinical effectiveness and long-term functional outcomes of conservative versus surgical management in degenerative and compressive spinal disorders. International Journal of Scientific Research in Engineering and Management, 10(03).

Gautam, M., Samyal, M., & Chaudhary, S. (2026). Preoperative risk stratification and surgical outcome prediction: Integrating clinical scoring systems, data-driven models, and patient-

https://doi.org/10.47059/AJMS/V5I4/01

centered optimization. International Innovations & Scholarly Trends Journal, 2(3).

Gupta, O. P., Gautam, S., & Maitiy, S. (2026). Management strategies for degenerative musculoskeletal disorders: An integrated clinical and technological framework. International Journal of Recent Development in Engineering and Technology, 15(3).

Jagar, K. D., Kumar, A., & Yadav, A. (2026). Multimodal analgesia in acute and chronic pain management: Mechanistic rationale, clinical applications, and outcome-focused strategies. Journal of Advanced Academic Research and Findings, 4(2).

Jenifer, R. D., Vettriselvan, R., Saxena, D., Velmurugan, P. R., & Balakrishnan, A. (2025). Green marketing in healthcare advertising: A global perspective. In AI impacts on branded entertainment and advertising (pp. 303–326). IGI Global.

Jha, S. C., Kumar, P., & Neha. (2026). Artificial intelligence-assisted decision support in internal medicine: Enhancing clinical judgment, precision care, and health system performance. International Journal of Scientific Development and Research, 11(2).

Kumar, A., Kumar, N., & Dhabhai, M. (2026). Optimizing outcomes through evidence-based protocols in postoperative intensive care: Clinical standards, implementation strategies, and quality improvement. International Journal of Scientific Research and Technology.

Kumar, P., Gautam, S., & Maitiy, S. (2026). Diagnostic utility of biomarkers in early disease stratification: Clinical applications, predictive value, and emerging innovations. Journal of Emerging Technologies and Innovative Research, 13(2).

Kumar, R., Sharma, K., & Gupta, S. K. (2026). Multimorbidity patterns and therapeutic complexity in adult medical practice: Implications for polypharmacy and patient-centred care.

https://doi.org/10.47059/AJMS/V5I4/01

International Journal of Creative Research Thoughts, 14(2).

Kumar, S., Kumar, V., & Tomer, H. (2026). Evolution of minimally invasive surgical techniques and clinical outcomes: Technological progress, specialty-specific adoption, and outcome-oriented evidence. Asian Journal of Multidimensional Research.

Lal, S. K., Vaibhav, & Khursheed, M. (2026). Anesthetic management in geriatric and comorbid patients: Clinical challenges, risk stratification, and outcome-oriented strategies. International Innovations & Scholarly Trends Journal, 2(2).

Lodha, V., Sharma, K., & Saraswat, P. (2026). Role of structured multidisciplinary rehabilitation in optimizing functional recovery and long-term outcomes in musculoskeletal disorders. International Journal of Scientific Research in Engineering and Management, 10(03).

Mishra, A., Choudhary, A., & Kumar, S. (2026). Infection prevention strategies in operative care: Evidence-based interventions, systems integration, and outcome-oriented practice. Asian Journal of Multidimensional Research.

Mustafa, N., Zahoor, H., Gamil, R. E., Ashifa, K. M., & Safaei, M. (2026). Empowering future caregivers: The role of self-leadership in reducing stress among nursing students. International Journal of Innovation and Learning, 39(1), 74–103.

Natarajan, P., Saravanan, A., Krishnakumar, B., Chandralekha, V., & Suresh Kumar, A. (2026). Motion control strategies in assistive devices for elderly and differently-abled patients. In Intelligent motion control for human-centered systems (pp. 103–126). IGI Global.

Pavithra, M. R., Chitra, V., Subhashini, V., Hemalatha, D., & Kiran, S. R. (2026). Intelligent prosthetics motion learning and neuro-muscular feedback integration. In Intelligent motion control for

https://doi.org/10.47059/AJMS/V5I4/01

human-centered systems (pp. 77–102). IGI Global.

Pradeepa, S. V., Gokilavani, R., Deepan, A., Sridevi, J., & Selvi, K. (2026). Predictive maintenance and asset management using motion analytics. In Methodologies and applications of intelligent motion control systems (pp. 75–104). IGI Global.

Rani, A., & Tyagi, N. (2026). Osteoporosis risk assessment and preventive interventions: Evidence-based screening, predictive modelling, and community-level strategies. International Journal of Recent Development in Engineering and Technology, 15(3).

Rasi, R. A., & Ashifa, K. M. (2019). Role of community-based programmes for active ageing: Elders self-help group in Kerala. Indian Journal of Public Health Research & Development, 10(12).

Sahu, R. L., Sharma, K., & Gupta, S. K. (2026). Biological and mechanical determinants of fracture healing: An integrated mechano-biological, systemic, and translational framework. International Journal of Recent Development in Engineering and Technology, 15(3).

Sehgal, A., Jayapriya, R., & Kumar, A. (2026). Evidence-based rehabilitation in sports-related injuries: Clinical effectiveness, multidisciplinary integration, and technological advancements. International Journal of Recent Development in Engineering and Technology, 15(3).

Selvi, K., Anbarasan, P., Madhumita, G., Janaki, L., & Devi, K. K. (2026). Governance, security, and ethical considerations in AI-driven motion control systems. In Methodologies and applications of intelligent motion control systems (pp. 217–242). IGI Global.

Shanthi, H. J., Gokulakrishnan, A., Sharma, S., Deepika, R., & Swadhi, R. (2025). Leveraging artificial intelligence for enhancing urban health: Applications, challenges, and

https://doi.org/10.47059/AJMS/V5I4/01

ASET JOURNAL OF MANAGEMENT SCIENCE
Peer Reviewed & Open Access Journal
ISSN : 2584 - 220X (Online) | RNI : Applied | Frequency : Bi-Monthly

innovations. In Nexus of AI, climatology, and urbanism for smart cities (pp. 275–306). IGI Global.

Sharma, S., Sharma, K., & Tyagi, N. (2026). Geriatric psychiatry and cognitive decline: Clinical evaluation, neuropsychological determinants and evidence-based management. International Journal of Management Research and Social Science, 13(2).

Singh, S., Chauhan, Y., & Kumar, D. (2026). Biomechanical innovations in orthopaedic implant design: Materials science, AI-assisted customization, and smart integration. International Journal of Scientific Research in Engineering and Management, 10(03).

Singhal, A., Kumar, A., & Kataria, N. (2026). Advances in wound healing and tissue regeneration: Biomaterials, regenerative strategies, and translational challenges. International Innovations & Scholarly Trends Journal, 2(2).

Subramani, M., Chillagattu, V., Gayathri, K., Rastogi, V., & Ranganathan, S. (2026). Digital twin integration for predictive and real-time motion control in infrastructure engineering. In Methodologies and applications of intelligent motion control systems (pp. 189–216). IGI Global.

Suresh, N. V., Hemalatha, S., Lakshmi, S. J., Mounica, C., & Kalaivani, M. (2026). AI-enabled motion control in surgical robotics: Precision, dexterity, and real-time adaptation. In Intelligent motion control for human-centered systems (pp. 29–50). IGI Global.

Swadhi, R., Gayathri, K., Suresh, N. V., Catherine, S., & Velmurugan, P. R. (2025). Leveraging machine learning for enhanced patient engagement and outcomes: Revolutionizing healthcare marketing. In Impact of digital transformation on business growth and performance (pp. 313–340). IGI Global.

https://doi.org/10.47059/AJMS/V5I4/01

Venice, A., Swadhi, R., Gayathri, K., Chandra, P., & Sajana, K. P. (2026). Rehabilitation robotics and adaptive motion planning for patient-centric care. In R. Vettriselvan & N. Suresh (Eds.), Intelligent motion control for human-centered systems (pp. 51–76). IGI Global.

Vettriselvan, R., Ramya, R., Selvalakshmi, V., Jyothi, P., & Velmurugan, P. R. (2026). Empowering patients through knowledge: Educational strategies in rehabilitation. In Holistic approaches to health recovery (pp. 263–290). IGI Global.

Vettriselvan, R. (2025). Harnessing innovation and digital marketing in the era of industry 5.0: Resilient healthcare SMEs. In The future of small business in industry 5.0 (pp. 163–186). IGI Global.

Vettriselvan, R., Velmurugan, P. R., Varshney, K. R., EP, J., & Deepika, R. (2025). Health impacts of smartphone and internet addictions across age groups: Physical and mental health across generations. In Impacts of digital technologies across generations (pp. 187–210). IGI Global.

Vijayalakshmi, M., Subramani, A. K., Vettriselvan, R., Velmurugan, P. R., & Hasine, J. (2025). Strategic collaborations in medical innovation and AI-driven globalization: Advancing healthcare startups. In Navigating strategic partnerships for sustainable startup growth (pp. 85–110). IGI Global.

Vinodh, N., & Subramani, A. K. (2026). Augmented and virtual reality for training and operations in motion-control environments. In Intelligent motion control for human-centered systems (pp. 153–178). IGI Global.

Vinodh, N., Subramani, A. K., & Vettriselvan, R. (2026). Transforming the future of management and medical education: AI-driven innovations in curriculum design. In AI education strategies for future-

https://doi.org/10.47059/AJMS/V5I4/01

proofing curriculum design (pp. 459–476). IGI Global.

Yatish, Khatoon, N., & Kumar, A. (2026). Advancing preventive strategies for chronic disease management: Clinical, behavioral, and population-level perspectives. International Journal of Novel Trends and Innovation, 4(2).

Zahoor, H., Mustafa, N., Ashifa, K. M., Safaei, M., & El Gamil, R. (2025). Unlocking resilience: Emotional intelligence and self-leadership shape stress perception among health students. International Journal of Innovation and Learning, 38(4), 395–419.

https://doi.org/10.47059/AJMS/V5I4/01