REVIEW **Open access**

Machine Learning for Healthcare Workforce Analytics: A Review of Models for Staffing Prediction, Burnout Detection, Skill Matching, Productivity Monitoring, and Retention Risk Forecasting

Diego Morales[1*], Andres Gutierrez[1], Lucia Navarro[2], Pablo Rios[1]

Abstract

Healthcare workforce management faces persistent challenges, including staffing shortages, clinician burnout, skill misalignment, productivity pressure, and turnover. Machine learning has been proposed as a way to support earlier prediction, better allocation, and more responsive workforce decision-making. This systematic review synthesizes peer-reviewed evidence from 2017 to 2025 on machine learning applications in healthcare workforce analytics. The review focuses on staffing prediction, burnout detection, skill matching, productivity monitoring, and retention risk forecasting. A PRISMA 2020-aligned review approach was used to structure database searching, screening, eligibility assessment, and narrative synthesis. The review considered studies from PubMed, Scopus, IEEE Xplore, and Web of Science, with emphasis on workforce-relevant outcomes and healthcare operational settings. Staffing prediction and burnout detection were the most developed domains in the reviewed literature. Skill matching and productivity monitoring were less mature, while retention risk forecasting showed growing interest but limited implementation evidence. Machine learning offers useful tools for healthcare workforce analytics, but the evidence base remains uneven. Stronger prospective validation, fairness assessment, and implementation reporting are needed before these models can be considered mature workforce decision-support systems.

Keywords Machine learning, Healthcare workforce analytics, Staffing prediction, Clinician burnout, Skill matching, Productivity monitoring

*Correspondence:
Diego Morales
diego.morales@gmail.com
[1] Department of Intelligent Healthcare Informatics, Faculty of Medicine, University of Lima, Lima, Peru
[2] Department of Clinical Data Analytics, Faculty of Engineering, Pontifical Catholic University of Peru, Lima, Peru

Introduction

Healthcare organizations increasingly face workforce strain from staffing shortages, rising workload, occupational dissatisfaction, and employee turnover. Nurse staffing research has shown that workforce adequacy cannot be understood only through headcounts, because patient acuity and unit-level demand strongly shape staffing requirements [1]. Physician scheduling research similarly shows that workforce planning must address availability, workload distribution, and service continuity rather than relying on static staffing assumptions [2]. These pressures have made predictive workforce analytics more relevant to healthcare management.

Traditional healthcare workforce analytics often rely on descriptive staffing ratios, retrospective productivity reports, and manual planning processes. Machine learning can extend these approaches by identifying patterns in appointment

demand, bed demand, clinical workload, and productivity variation before shortages become operationally visible [3]. For example, outpatient appointment modeling has demonstrated how predictive and prescriptive analytics may support scheduling decisions that influence workforce needs [4]. Such applications suggest that machine learning can help shift workforce management from retrospective reporting toward anticipatory decision support.

The literature remains fragmented because studies often focus on one workforce problem rather than the full analytics cycle. Burnout detection studies have used electronic health record activity logs to infer work strain, while retention studies have modeled turnover or departure risk using workforce and administrative data [5, 6]. Skill matching and assignment studies have developed decision-support approaches for nurse-patient allocation, but this work is less common than forecasting and burnout prediction [7]. A holistic review is therefore needed to synthesize how machine learning supports staffing, burnout, skill matching, productivity, and retention decisions as connected workforce challenges.

This systematic review was designed to synthesize machine learning applications for healthcare workforce analytics from 2017 to 2025 using principles. The review includes empirical studies, systematic reviews, scoping reviews, and implementation-relevant papers that address one or more workforce analytics domains. Existing reviews on burnout prediction, turnover modeling, nursing workforce management, and workforce projection provide useful foundations but do not fully integrate the five domains considered here. The objective is to identify the maturity, data sources, validation practices, implementation gaps, and ethical considerations in this emerging field.

Materials and Methods

Search strategy

The search strategy targeted peer-reviewed literature published from 2017 through 2025 on machine learning, artificial intelligence, optimization, and predictive analytics applied to healthcare workforce management. Searches were organized around staffing prediction, burnout detection, skill matching, productivity monitoring, retention risk forecasting, and broader workforce analytics. Staffing-related searches were informed by literature on nurse staffing and physician scheduling, both of which use operational data to support workforce decisions. Burnout and retention searches were informed by studies using EHR audit logs, clinical activity records, survey data, and employment-related outcomes.

Inclusion and exclusion criteria

Studies were eligible when they applied machine learning, artificial intelligence, optimization, forecasting, or predictive analytics to a healthcare workforce outcome. Eligible outcomes included staffing need, shift scheduling, workload assessment, burnout risk, clinician assignment, productivity, turnover, retention, practice location, or workforce duration. Studies of nurse turnover prediction were included when they used workforce data to model turnover or turnover intention. Studies were excluded when they focused exclusively on patient diagnosis, clinical prognosis, hospital finance, or operational throughput without a workforce decision point.

Screening and selection

Screening followed principles, with title and abstract review followed by full-text eligibility assessment. A realistic flow for **Figure 1** includes 1,942 records identified, 312 duplicates removed, 1,630 titles and abstracts screened, 1,380 records excluded, 250 full texts assessed, and 74 studies included in the broader evidence map. Full-text exclusions most often occurred because articles lacked a workforce outcome, did not apply predictive or optimization methods, or focused on patient-level clinical decision-making rather than workforce analytics. Studies of health workforce projection were retained when they informed workforce planning methods and reporting practices [8].

Figure 1 presents the PRISMA 2020 flow diagram for study identification, screening, eligibility assessment, and final inclusion.

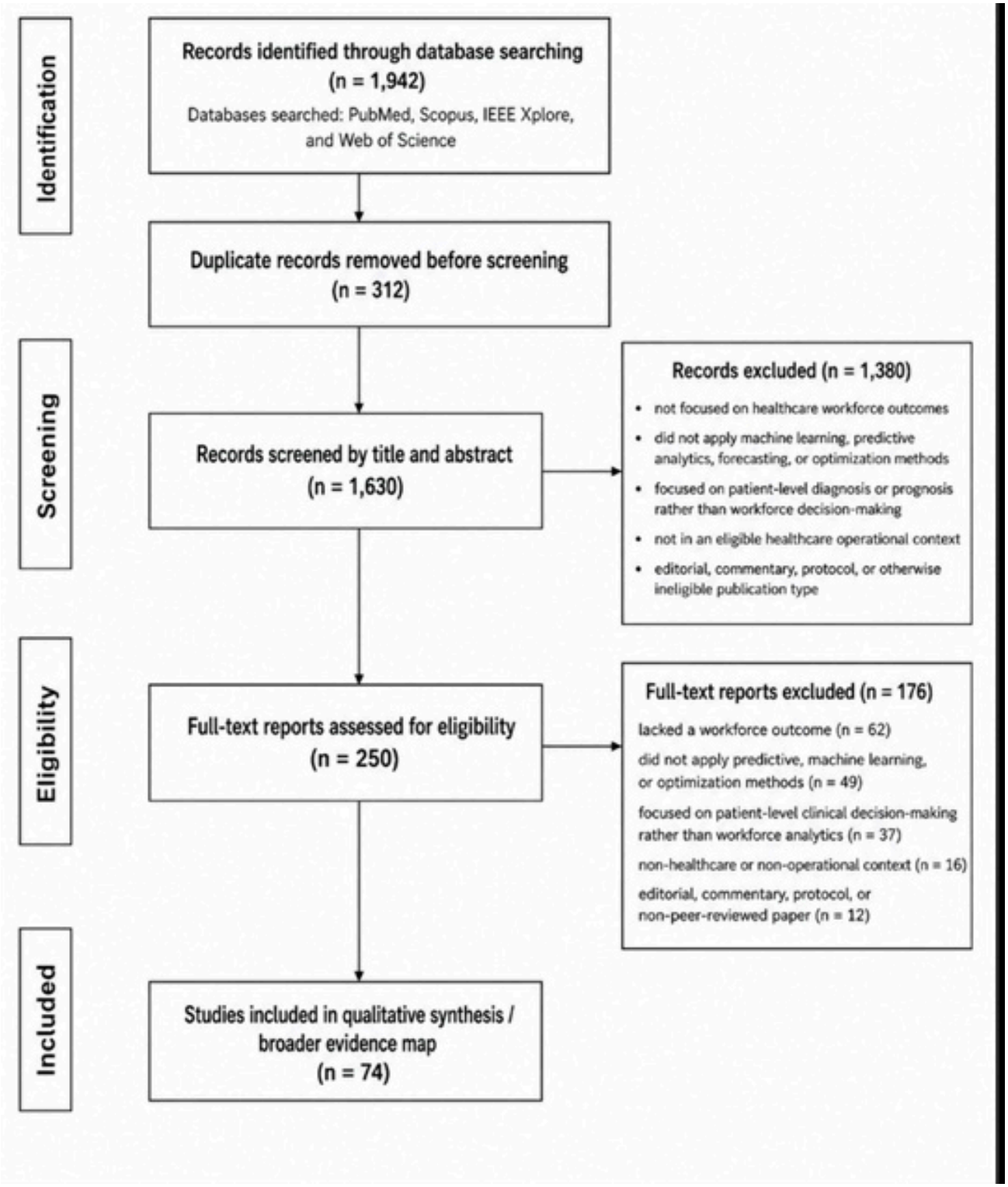


Figure 1. PRISMA 2020 Flow Diagram for Study Selection in a Systematic Review of Machine Learning for Healthcare Workforce Analytics

Data extraction

Data extraction captured the workforce domain, healthcare setting, workforce group, data source, modeling approach, outcome, validation method, and implementation status. For burnout studies, extracted variables included EHR activity features, audit log signals, survey outcomes, and links to validated burnout measures where reported. For staffing and productivity studies, extraction emphasized patient demand, appointment flow, census, scheduling constraints, and time-varying clinician productivity. For retention studies, extraction focused on turnover outcomes, length of stay, practice location, turnover intention, and workforce stability indicators.

Risk of bias assessment

Risk of bias was assessed qualitatively using principles adapted from prediction-model appraisal frameworks. The assessment considered participant selection, predictor measurement, outcome definition, missing data, temporal validation, external validation, and transportability across institutions. Burnout models were judged as higher risk when EHR activity patterns could reflect local documentation culture rather than generalizable clinician strain. Retention models were also assessed for class imbalance and context specificity because turnover events may be unevenly distributed across roles, settings, and time periods.

Synthesis methods

A narrative synthesis was used because the included evidence differed substantially in setting, workforce group, data source, model type, and outcome definition. Studies were first grouped into the five target domains and then examined for cross-cutting patterns in data sources, validation, implementation, and ethical reporting. Scoping reviews of artificial intelligence in nursing care organization and nursing leadership were used to contextualize implementation maturity and leadership implications. Evidence on AI's broader influence on the nursing workforce was also used to interpret how workforce analytics may reshape management responsibilities.

Results and Discussion

Study selection

The study selection process showed that machine learning for healthcare workforce analytics is expanding but unevenly distributed across domains. Staffing prediction, burnout detection, and retention forecasting were represented more consistently than skill matching and productivity monitoring. Nurse-focused workforce analytics appeared frequently, especially in staffing, assignment, and turnover studies [1]. Physician-focused studies appeared most often in scheduling, productivity, EHR workload, burnout, and departure prediction [9].

Study characteristics

The reviewed literature included studies of nurses, physicians, healthcare workers in underserved communities, and broader clinical workforces. Nurse studies addressed staffing metrics, nurse-patient assignment, turnover prediction, and AI-supported workforce management [10]. Physician studies examined EHR workload, scheduling, productivity variation, burnout prediction, and departure risk [11]. Broader healthcare workforce studies included resilience, sickness absence, burnout, practice location, and health workforce projection [12].

Staffing prediction models

Staffing prediction models addressed workforce demand through nurse staffing metrics, operational forecasting, physician scheduling, and hospital capacity planning. The adjusted nurse staffing metric developed for neonatal intensive care showed how staffing adequacy can be modeled in relation to unit-level care requirements [1]. A review of operational research techniques in nurse staffing showed that staffing analytics has a methodological base that predates many newer

machine learning applications [13]. Physician scheduling literature similarly emphasized that staffing prediction must account for constraints such as coverage, continuity, workload, and availability [2].

Staffing prediction: features and methods

Staffing prediction studies commonly used features such as patient census, historical demand, bed occupancy, acuity, seasonal variation, appointment flow, and shift constraints. Machine learning approaches to inpatient bed demand showed how hospital capacity forecasts can indirectly support staffing decisions by anticipating service pressure [3]. Bed capacity planning studies also illustrated the importance of combining machine learning with operational context so that predicted demand can inform workforce allocation [14]. In staffing applications, feature engineering was therefore closely tied to the operational setting rather than to generic workforce variables alone.

Burnout detection models

Burnout detection models increasingly used EHR audit logs, clinical activity logs, and survey-linked data to identify work patterns associated with clinician strain. Lou and colleagues reported physician burnout prediction using clinical activity logs, showing how digital work traces can be converted into workforce risk indicators [5]. The HiPAL framework further demonstrated that deep learning can be applied to EHR activity logs for burnout prediction among physicians [15]. These studies suggest that burnout analytics is moving from exclusive reliance on surveys toward combined behavioral and self-report data sources.

Burnout detection: validation and outcomes

Burnout prediction studies were usually validated against survey-based burnout measures or related workforce outcomes rather than prospective intervention effects. Research predicting primary care physician burnout from EHR-use measures showed that clinical system activity can be linked to burnout-related outcomes, although interpretation remains dependent on context [16]. A systematic review and meta-analysis of machine learning for healthcare-worker burnout found growing interest but substantial heterogeneity in populations, predictors, and validation approaches [17]. Current evidence therefore supports model development but does not yet establish that burnout predictions reduce burnout when acted upon.

Skill matching models

Skill matching was less common than staffing prediction and burnout detection, but several studies addressed related assignment problems. Integrated patient-to-room and nurse-to-patient assignment work demonstrated how assignment decisions can be modeled across interacting constraints in hospital wards [10]. Nurse-patient assignment research also showed that continuity of care can be formalized as a decision objective alongside workload considerations [7]. A later nursing-focused paper argued that machine learning may influence care assignment dynamics, but direct evidence on competency-based clinician-to-task matching remains limited [18].

Productivity monitoring models

Productivity monitoring models were concentrated in physician scheduling, emergency department operations, outpatient appointment systems, and EHR workload assessment. Productivity-driven physician scheduling in emergency departments showed how workforce output can be incorporated into scheduling decisions rather than treated as a fixed characteristic [11]. Studies of time-varying productivity further suggested that physician performance may fluctuate across shifts and operational conditions [19]. EHR-based workload assessment complemented this literature by showing that digital activity data can describe the burden associated with clinical work [9].

Retention risk forecasting

Retention risk forecasting was an emerging domain, with studies modeling nurse turnover, physician departure, practice location, and workforce duration. A machine learning model for nurse turnover in Korea demonstrated how employment-related data can be used to estimate turnover risk in hospital nursing workforces [20]. Another study addressed imbalanced nurse turnover data using oversampling and machine learning algorithms, highlighting a recurring technical challenge in retention modeling [21]. Physician departure prediction using EHR-use patterns suggested that digital work behavior may also contain signals relevant to clinician retention [6].

Data sources and feature engineering

Data sources included EHR logs, clinical activity logs, scheduling data, appointment systems, staffing records, surveys, HR-related variables, and administrative workforce data. Studies of healthcare workers in underserved South African communities used workforce duration data to model length of stay, illustrating how retention analytics can be adapted to workforce distribution problems [22]. Practice-location modeling similarly used machine learning to identify factors associated with where healthcare workers practice [23]. These studies show that workforce feature engineering depends on both the employment outcome and the policy or management question being addressed.

ML algorithms employed

The reviewed literature used supervised machine learning, deep learning, forecasting, optimization, heuristic assignment, and prescriptive analytics. Burnout studies included deep learning and log-based approaches, whereas turnover studies more often used structured-data classifiers suited to administrative or survey variables [15, 24]. Staffing and assignment studies relied more heavily on operations research, optimization, and scheduling methods that support constrained decision-making [7]. Because methods varied across domains, direct comparison of algorithms was less informative than assessing whether each method aligned with the workforce decision being supported.

Evaluation and validation

Evaluation and validation practices varied widely across the reviewed studies. Predictive studies commonly reported internal validation, while external validation and temporal validation were less consistently described. A multinational study of nurse turnover intention showed that machine learning can be applied across broader workforce samples, although such work still requires careful attention to setting differences and implementation context [24]. Across the evidence base, validation was generally stronger for model development than for demonstrating effects on staffing decisions, burnout reduction, productivity improvement, or retention outcomes.

Implementation and deployment

Few studies reported full operational deployment of workforce machine learning tools in routine management workflows. Staffing and assignment models were closest to implementation because they addressed concrete decisions such as coverage, scheduling, allocation, and care continuity [10, 13]. Burnout and retention models were often framed as early-warning tools, but most did not report prospective use by managers or occupational well-being teams [5, 25]. Reviews of AI in nursing leadership emphasized that deployment requires governance, staff engagement, workflow integration, and leadership readiness [26].

Figure 2 presents an evidence-to-implementation map showing how the reviewed studies connect workforce domains, data sources, analytic methods, validation patterns, implementation gaps, and future priorities.

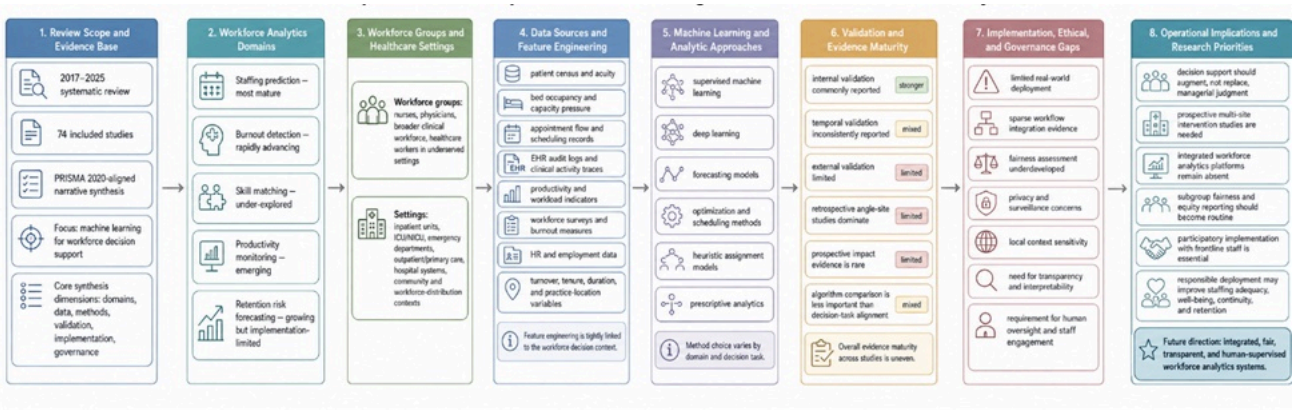


Figure 2. Evidence-to-Implementation Map of Machine Learning for Healthcare Workforce Analytics

Table 1 compares the five workforce analytics domains in terms of managerial purpose, data foundation, analytic approach, operational output, evidence maturity, and unresolved challenges.

Table 1. Cross-Domain Architecture of Machine Learning Applications in Healthcare Workforce Analytics

Workforce analytics domain	Core managerial question	Primary data foundation	Predominant analytic approaches	Main output delivered to decision-makers	Current evidence maturity	Principal unresolved issue
Staffing prediction	How many staff are needed, of what type, at what time, and in which location?	Patient census, acuity, bed occupancy, appointment flow, historical demand, seasonal trends, shift constraints	Forecasting, supervised ML, optimization, scheduling models, prescriptive analytics	Demand forecasts, shift staffing recommendations, coverage scenarios, allocation plans	**Highest maturity** across the five domains; closest to operational use	Limited prospective evidence that model-guided staffing improves workforce and patient outcomes across settings
Burnout detection	Which clinicians may be at risk of strain, and when should supportive intervention be triggered?	EHR audit logs, after-hours work indicators, inbox burden, documentation patterns, survey-linked burnout measures, workload signals	Supervised ML, deep learning, log-mining, pattern recognition	Burnout risk flags, workload strain profiles, prioritization for supportive review	**Moderate-to-high maturity** in model development, but implementation remains limited	Construct validity, privacy concerns, and risk of punitive or stigmatizing use
Skill matching	Which clinician should be assigned to which	Credentials, competencies, patient acuity, continuity needs,	Optimization, heuristic assignment, constrained decision	Assignment recommendations, continuity-preserving allocation options,	**Lowest maturity**; evidence remains sparse	Inadequate modeling of real competency profiles, fairness constraints, and

	patient, task, or shift under operational and care constraints?	workload balance, assignment history, staffing rules, stated preferences	models, early ML-assisted matching	workload-balanced staffing plans	and fragmented	explainable assignment logic
Productivity monitoring	How does workforce output vary across shifts and conditions, and how should schedules or workload distribution adapt?	Scheduling records, encounter volume, throughput indicators, documentation time, EHR workload traces, operational conditions	Predictive modeling, time-varying productivity analytics, scheduling optimization, prescriptive analytics	Productivity forecasts, schedule adjustments, workload balancing insights, operational alerts	**Emerging maturity**; conceptually important but not yet well standardized	Risk of overemphasizing volume while underrepresenting quality, coordination work, and clinician well-being
Retention risk forecasting	Which workforce groups or individuals may be at risk of leaving, and what supportive action is justified?	HR and administrative records, tenure, turnover history, survey data, location variables, EHR-use patterns, workforce duration data	Structured-data classifiers, imbalance-aware ML, survival-oriented prediction, retention-risk modeling	Risk stratification, retention planning priorities, workforce stability alerts, targeted support planning	**Growing maturity**, but heavily context-specific and weakly implemented	Poor transportability, fairness sensitivity, and danger of discriminatory or punitive interpretation

Staffing prediction is the most mature domain

Staffing prediction appeared to be the most mature domain because it builds on established nurse staffing metrics, physician scheduling research, and operational forecasting. Nurse staffing studies demonstrated that workforce requirements can be modeled in relation to care needs rather than fixed ratios alone [1]. Physician scheduling studies showed that workforce models must balance coverage, workload, and service constraints in ways that are familiar to healthcare operations [2]. However, even in this mature domain, prospective evidence on workforce and patient outcomes remains limited.

Burnout detection is advancing with EHR audit logs

Burnout detection is advancing through the use of EHR audit logs and clinical activity data. These data sources can capture workload patterns that may not be visible in periodic surveys, such as after-hours work, documentation burden, and system-use intensity [5]. Studies mining EHR audit logs also show that unsupervised learning can characterize clinical tasks and work patterns relevant to burnout interpretation [27]. Nevertheless, digital workload measures should be treated as supportive signals rather than definitive evidence of individual burnout.

Skill matching remains under-explored

Skill matching remains under-explored despite its relevance to patient safety, workforce satisfaction, and care continuity. Current evidence is strongest for nurse-patient assignment and related optimization problems rather than full competency-based matching across clinicians and patient needs [10]. Continuity-oriented assignment models show that workforce allocation can incorporate care relationship factors as well as workload balance [7]. Future studies need to integrate credentials, competencies, preferences, acuity, and fairness constraints into assignment models that can be understood by clinical leaders and frontline staff.

Productivity monitoring models: from descriptive to predictive

Productivity monitoring is beginning to move from retrospective reporting toward predictive and prescriptive modeling. Emergency department studies show that clinician productivity can be incorporated into scheduling models, especially when productivity varies across shifts or operational conditions [11]. Outpatient appointment analytics similarly indicate that demand prediction and scheduling rules can influence workforce utilization and throughput [4]. The main concern is that productivity models may overemphasize volume-based output unless they are balanced with quality, complexity, coordination work, and clinician well-being.

Retention risk prediction: data and privacy sensitivity

Retention risk prediction has grown through studies of turnover, turnover intention, physician departure, workforce duration, and practice location. Nurse turnover models demonstrate the feasibility of applying machine learning to structured workforce data, but the results are highly dependent on local employment conditions and outcome definitions [20]. Physician departure prediction using EHR-use patterns raises additional privacy and interpretability concerns because digital work traces may reflect both disengagement and workload burden [6]. Retention models should therefore be deployed only as supportive planning tools with strong safeguards against punitive or discriminatory use.

Implementation gaps across all domains

Implementation gaps were evident across staffing, burnout, skill matching, productivity, and retention domains. Many studies developed models but did not report how predictions were embedded into workforce workflows, manager decision-making, or staff-facing interventions. Scoping reviews of AI in nursing care organization suggest that technical feasibility is advancing faster than organizational implementation evidence [28]. Health workforce projection guidance also highlights the need for transparent reporting when predictive models are used to inform planning decisions [8].

Ethical and fairness considerations

Ethical and fairness considerations were underdeveloped across much of the evidence base. Workforce models can reproduce historical inequities if they are trained on biased staffing patterns, uneven workload documentation, unequal access to development opportunities, or differential turnover conditions. Reviews of AI and nursing leadership emphasize that governance and accountability are central when algorithms shape workforce management [29]. Broader discussion of AI's influence on the nursing workforce also suggests that implementation should protect professional autonomy, well-being, and equity [30].

Table 2 outlines a governance and implementation readiness framework for translating workforce machine learning models into responsible healthcare decision-support systems.

Table 2. Governance and Implementation Readiness Framework for Machine Learning–Enabled Workforce Decision Support in Healthcare

Governance or implementation dimension	Why this dimension matters in workforce analytics	Minimum reporting expectation for journal publication	Operational safeguard required before deployment	Likely consequence if neglected

Data provenance and context	Workforce models are highly sensitive to local staffing culture, documentation practice, HR policies, and operational workflows	Clearly describe data sources, setting, workforce group, period of capture, and known contextual constraints	Perform local data-quality review and confirm that training data reflect the intended deployment environment	A technically accurate model may fail or mislead when moved across units or institutions
Outcome definition	Workforce outcomes such as burnout, productivity, and retention are conceptually complex and can be measured inconsistently	Define outcomes precisely, including label source, timing, and whether outcomes are direct, proxy, or self-reported	Review outcome definitions with operational leaders and workforce stakeholders before use	The model may optimize an administratively convenient label rather than the real workforce problem
Validation strategy	Internal validation alone does not establish whether a model is reliable over time or across settings	Report internal, temporal, and external validation whenever feasible, plus subgroup performance	Conduct temporal re-testing and local pilot evaluation before using the model in workflow	Apparent performance may collapse under real-world deployment conditions
Fairness and equity	Workforce models may encode historical understaffing, biased evaluation structures, or unequal opportunity patterns	Report subgroup performance, fairness checks, and any bias-mitigation strategy	Require fairness audit review before deployment and repeat it during model updates	Existing disparities may be reproduced or amplified in staffing, monitoring, or retention decisions
Privacy and surveillance risk	Burnout, productivity, and retention models may use sensitive digital work traces that staff could experience as monitoring	Explain whether EHR logs, activity traces, or employment data were used and describe privacy protections	Establish role-based access, purpose limitation, and explicit governance around non-punitive use	Staff trust may deteriorate and models may be perceived as instruments of surveillance
Explainability and interpretability	Workforce decisions affect professionals directly and therefore require understandable rationale	Report interpretable features, explanation methods, or rationale pathways suitable for end users	Present explanations in manager-facing and staff-facing formats that support contestability and review	Users may reject, misuse, or overtrust model outputs
Human oversight and action pathway	A prediction is only useful if it leads to a defined, appropriate, and accountable human response	Specify who receives outputs, what decision is supported, and what action pathway follows	Define escalation rules, override authority, and documentation of human review	Predictions may accumulate without triggering action or may drive inconsistent managerial responses

Workflow integration	Even strong models have limited value if they do not fit scheduling, leadership, occupational well-being, or HR workflows	Describe the intended operational workflow and decision point for model use	Pilot the tool within an actual workflow with end-user feedback before scale-up	Implementation failure occurs despite acceptable technical performance
Monitoring and recalibration	Workforce conditions change over time because of seasonality, staffing shortages, policy shifts, and organizational change	Report plans for drift monitoring, update frequency, and post-deployment evaluation	Establish ongoing performance monitoring and periodic recalibration	Model degradation may go unnoticed and produce outdated recommendations
Staff engagement and acceptability	Workforce analytics affects clinicians and managers whose buy-in is essential for safe use	Report whether staff, managers, or leadership were involved in model design or evaluation	Use participatory design, communication plans, and feedback channels during implementation	Resistance, fear, and low adoption may undermine otherwise useful decision-support tools

Limitations

Review limitations

This review is limited by its English-language focus, heterogeneity across domains, and reliance on peer-reviewed studies available within the 2017–2025 window. Because staffing, burnout, skill matching, productivity, and retention use different outcomes and methods, meta-analysis was not appropriate for this synthesis. The review instead used narrative synthesis, drawing on systematic and scoping reviews where they clarified the state of the evidence [15]. Some workforce analytics implementations may also be unpublished within health systems or reported in internal operational documents rather than academic literature.

Evidence base limitations

The evidence base was limited by retrospective designs, single-site development, inconsistent validation, and sparse implementation reporting. Burnout and retention studies often relied on existing EHR, survey, or administrative data, which may reflect local culture, documentation practices, and employment conditions [16]. Staffing and assignment studies were more directly connected to operational decisions but still often lacked prospective evaluation of workforce outcomes [19]. Across domains, fairness assessment, staff acceptability, and governance reporting were not yet routine features of healthcare workforce machine learning research.

Comparison with prior reviews

Prior reviews have usually focused on narrower segments of healthcare workforce analytics rather than treating the field as an integrated workforce decision-support domain. Reviews of nurse staffing and physician scheduling emphasized operational research, workforce planning, and scheduling constraints, but they did not comprehensively cover burnout detection, productivity monitoring, or retention risk forecasting [2, 13]. Burnout-focused reviews synthesized machine learning models for healthcare-worker burnout, but they generally treated burnout prediction as a distinct occupational health problem rather than part of a broader workforce analytics architecture [17]. Similarly, turnover-focused reviews concentrated on nurse turnover and turnover intention without fully integrating staffing, skill assignment, productivity, and burnout analytics [25].

This review differs from prior syntheses by organizing the evidence around five connected workforce analytics domains. Staffing prediction studies showed how workforce demand can be anticipated using patient census, capacity, appointment, and scheduling data, while burnout studies demonstrated how EHR activity logs and survey-linked models can identify work strain [3, 5]. Skill matching and assignment studies added a prescriptive workforce allocation perspective, although they were less developed than prediction-oriented domains [7, 10]. Retention studies then extended workforce analytics toward longer-term organizational stability by modeling turnover, practice location, length of stay, and departure risk [6, 20, 22, 23].

The comparison with prior reviews suggests that the next phase of the field should move from isolated models toward integrated workforce analytics platforms. Reviews of AI in nursing leadership and care organization emphasize that implementation requires governance, workflow integration, staff participation, and leadership readiness [28, 29]. Health workforce projection guidance also shows that predictive models must be reported transparently if they are to inform workforce planning decisions across settings [8]. The distinctive contribution of this review is therefore to connect technical model development with operational implementation, ethical oversight, and cross-domain workforce management.

Recommendations

For researchers

Researchers should prioritize prospective, multi-site studies that evaluate whether machine learning-informed decisions improve staffing adequacy, burnout prevention, productivity management, and retention outcomes. Existing burnout and turnover studies demonstrate that predictive models can be developed from EHR logs, surveys, and administrative data, but evidence of intervention impact remains limited [16, 21]. Future studies should report model transportability, subgroup performance, temporal validation, and the organizational actions triggered by predictions. Researchers should also include fairness metrics and participatory design methods so that workforce analytics tools are evaluated not only for technical validity but also for their effects on workers.

For journal editors

Journal editors should require workforce machine learning studies to report deployment context, ethical safeguards, data provenance, and decision pathways. Studies that use digital work traces or employment data should explain how privacy, consent, surveillance risk, and organizational power dynamics were addressed [6, 27]. Papers should also specify whether models were retrospective, prospectively validated, piloted in workflow, or used to guide management actions. Requiring this information would help distinguish technical proof-of-concept studies from workforce decision-support systems that are ready for real-world evaluation.

For healthcare leaders

Healthcare leaders should treat machine learning workforce analytics as decision support rather than autonomous decision-making. Staffing prediction and scheduling models can help anticipate demand, but managers should combine predictions with frontline expertise, patient acuity, staff fatigue, and local operational knowledge [1, 11]. Burnout and retention models should be used to trigger supportive organizational responses, not punitive monitoring of individual clinicians [5, 20]. Health systems should invest in interoperable workforce data infrastructure, human oversight, model monitoring, and transparent communication with staff before deploying analytics tools.

For policymakers

Policymakers should develop guidance for ethical and accountable use of predictive workforce analytics in healthcare. This guidance should address high-risk domains such as burnout prediction, retention-risk modeling, and productivity monitoring, where predictions may influence employment conditions, workload allocation, or managerial attention [24, 31]. Regulatory and professional bodies should encourage fairness audits, privacy protections, documentation standards, and

independent evaluation of implemented tools. Policy support is also needed for multi-site research infrastructure that allows workforce models to be tested across diverse hospitals, regions, and workforce groups.

Research gaps

Prospective intervention trials

A central gap is the absence of prospective intervention trials showing that acting on machine learning predictions improves workforce outcomes. Staffing studies have developed forecasting and scheduling approaches, but few have tested whether deployment reduces understaffing, overtime, missed care, or staff dissatisfaction [13, 19]. Burnout studies have identified signals from EHR logs and surveys, but they generally have not evaluated whether early-warning systems reduce burnout when linked to organizational interventions [15, 16]. Retention models similarly need prospective testing to determine whether risk prediction can support career development, workload adjustment, or retention programs without producing unintended harms.

Integrated workforce analytics platforms

No reviewed evidence demonstrated a single integrated platform that combines staffing prediction, burnout detection, skill matching, productivity monitoring, and retention forecasting into one coordinated workforce decision-support system. Current studies usually model one workforce outcome at a time, such as bed demand, nurse assignment, physician productivity, burnout risk, or turnover probability [10, 14, 20]. This fragmentation makes it difficult for healthcare leaders to understand trade-offs between staffing efficiency, clinician well-being, care continuity, and long-term retention. Future systems should connect operational, HR, EHR, survey, and scheduling data while preserving transparency and appropriate human oversight.

Fairness and equity

Fairness and equity remain major unresolved gaps in healthcare workforce machine learning. Workforce models may encode historical understaffing, unequal workload distribution, biased documentation patterns, or differential access to professional development if these patterns are present in training data [23, 29]. Retention and productivity models are especially sensitive because they may influence how managers perceive employee commitment, value, or risk [6, 25]. Future studies should report subgroup validation, bias mitigation, governance structures, and staff-facing explanations so that workforce analytics does not reinforce existing disparities.

Implications

For research practice

For research practice, the field should shift from retrospective model development toward implementation science and real-world evaluation. The reviewed evidence shows substantial progress in using EHR logs, scheduling records, surveys, and administrative data to model workforce problems, but fewer studies demonstrate how predictions change decisions or outcomes [5, 22]. Researchers should design studies that follow the full pathway from data extraction to prediction, human interpretation, intervention, and workforce impact. This would make healthcare workforce analytics more relevant to management practice and less dependent on isolated technical benchmarks.

For healthcare management

For healthcare management, machine learning should be used to augment judgment, not replace professional and managerial accountability. Staffing and productivity models can support capacity planning and scheduling, but they must be interpreted alongside patient complexity, team dynamics, and clinician well-being [2, 4]. Skill matching systems may improve alignment between patient needs and clinician competencies, but they should preserve transparency, flexibility, and staff voice in assignment decisions [7, 18]. Leaders should also monitor whether predictive tools alter workload distribution or create unintended pressure on already strained workforce groups.

For society

For society, responsible machine learning in healthcare workforce analytics could contribute to safer staffing, healthier clinicians, better continuity of care, and more stable health systems. The potential benefits are especially important in settings facing shortages, maldistribution, burnout, and retention challenges [8, 22]. However, societal value depends on whether models are deployed ethically, evaluated prospectively, and aligned with workforce well-being rather than narrow efficiency goals. The broader literature on AI and the nursing workforce underscores that technological change should support professional practice, not weaken autonomy or trust [30].

Conclusion

Machine learning offers powerful tools to predict staffing needs, detect burnout, match skills, monitor productivity, and forecast retention risk in the healthcare workforce. Across the reviewed domains, the strongest evidence was found in staffing prediction and burnout detection, while skill matching and productivity monitoring remained less mature.

However, the field is dominated by retrospective, single-site, and domain-specific studies with limited evidence of real-world implementation or impact. The ability to build predictive models has advanced faster than the ability to evaluate how those models affect workforce decisions, clinician well-being, or organizational outcomes.

The most pressing gaps are the absence of prospective intervention trials, integrated workforce analytics systems, and equity analyses. Future research should examine not only whether models predict workforce outcomes, but also whether acting on those predictions improves staffing adequacy, reduces burnout, supports fair assignment, and strengthens retention.

A collaborative effort among researchers, healthcare leaders, journal editors, policymakers, and frontline clinicians is needed to translate workforce machine learning into responsible practice. With careful governance, transparent implementation, and sustained evaluation, these tools may help build a healthier, fairer, and more sustainable healthcare workforce.

Acknowledgements

None

Conflict of interest

None

Financial support

None

Ethics statement

None

Received: 13 Nov 2025 Revised: 10 Jan 2026 Accepted: 09 Feb 2026
Published online: 25 February 2026

Rights and permissions

Open Access The author(s) retain copyright. This article is licensed under the Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License. It may be shared and adapted for non-commercial purposes with appropriate attribution, an indication of changes, and distribution of adaptations under the same license. Third-party material may be subject to separate terms identified in its credit line. View the license at https://creativecommons.org/licenses/by-nc-sa/4.0/.

References

Tawfik DS, Profit J, Lake ET, Liu JB, Sanders LM, Phibbs CS. Development and use of an adjusted nurse staffing metric in the neonatal intensive care unit. Health Serv Res. 2020;55(2):190-200.

Erhard M, Schoenfelder J, Fügener A, Brunner JO. State of the art in physician scheduling. Eur J Oper Res. 2018;265(1):1-18.

Tello M, Reich ES, Puckey J, Maff R, Garcia-Arce A, Feijoo F, et al. Machine learning based forecast for the prediction of inpatient bed demand. BMC Med Inform Decis Mak. 2022;22(1):55.

Srinivas S, Ravindran AR. Optimizing outpatient appointment system using machine learning algorithms and scheduling rules: a prescriptive analytics framework. Expert Syst Appl. 2018;102:245-61.

Lou SS, Liu H, Warner BC, Harford D, Lu C, Kannampallil T. Predicting physician burnout using clinical activity logs: model performance and lessons learned. J Biomed Inform. 2022;127:104015.

Lopez K, Li H, Paek H, Williams B, Nath B, Loza AJ, et al. Predicting physician departure with machine learning on EHR use patterns: a longitudinal cohort from a large multispecialty ambulatory practice. PLoS One. 2023;18(2):e0280251.

Jiang H, Gomes P, Vander Meer D. Promoting continuity of care in nurse-patient assignment: a multiple objective heuristic algorithm. Decis Support Syst. 2023;167:113926.

Lee JT, Crettenden I, Tran M, Miller D, Cormack M, Li J, et al. Methods for health workforce projection model: systematic review and recommended good practice reporting guideline. Hum Resour Health. 2024;22(1):25.

Arndt BG, Beasley JW, Watkinson MD, Temte JL, Tuan WJ, Sinsky CA, et al. Tethered to the EHR: primary care physician workload assessment using EHR event log data and time-motion observations. Ann Fam Med. 2017;15(5):419-26.

Brandt T, Klein TL, Reuter-Oppermann M, Schäfer F, Thielen C, de Vrugt MV, et al. Integrated patient-to-room and nurse-to-patient assignment in hospital wards. OR Spectr. 2025;47:1-44.

Camiat F, Restrepo MI, Chauny JM, Lahrichi N, Rousseau LM. Productivity-driven physician scheduling in emergency departments. Health Syst. 2021;10(2):104-17.

Lieslehto J, Rantanen N, Oksanen LM, Oksanen SA, Kivimäki A, Paju S, et al. A machine learning approach to predict resilience and sickness absence in the healthcare workforce during the COVID-19 pandemic. Sci Rep. 2022;12(1):8055.

Saville CE, Griffiths P, Ball JE, Monks T. How many nurses do we need? A review and discussion of operational research techniques applied to nurse staffing. Int J Nurs Stud. 2019;97:7-13.

Mahmoudian Y, Nemati A, Safaei AS. A forecasting approach for hospital bed capacity planning using machine learning and deep learning with application to public hospitals. Healthc Anal. 2023;4:100245.

Liu H, Lou SS, Warner BC, Harford DR, Kannampallil T, Lu C. HiPAL: a deep framework for physician burnout prediction using activity logs in electronic health records. In: Proc 28th ACM SIGKDD Conf Knowl Discov Data Min; 2022. p. 3377-87.

Tawfik D, Bayati M, Liu J, Nguyen L, Sinha A, Shanafelt T, et al. Predicting primary care physician burnout from electronic health record use measures. Mayo Clin Proc. 2024;99(9):1411-21.

Shi H, Liu J, Yang C, Shang J, Zeng Y. Machine learning for predicting burnout among healthcare workers: a systematic review and meta-analysis. Contemp Nurse. 2025;61:1-21.

Othman MI, Nashwan AJ, Abujaber AA. Optimising nurse-patient assignments: the impact of machine learning model on care dynamics. Nurs Open. 2025;12(4):e70195.

Zaerpour F, Bijvank M, Ouyang H, Sun Z. Scheduling of physicians with time-varying productivity levels in emergency departments. Prod Oper Manag. 2022;31(2):645-67.

Kim SK, Kim EJ, Kim HK, Song SS, Park BN, Jo KW. Development of a nurse turnover prediction model in Korea using machine learning. Healthcare (Basel). 2023;11(11):1583.

Xu Y, Park Y, Park JD, Sun B. Predicting nurse turnover for highly imbalanced data using the synthetic minority over-sampling technique and machine learning algorithms. Healthcare (Basel). 2023;11(24):3173.

Moyo S, Doan TN, Yun JA, Tshuma N. Application of machine learning models in predicting length of stay among healthcare workers in underserved communities in South Africa. Hum Resour Health. 2018;16(1):68.

Bounsanga J, Lipsky MS, Hon ES, Licari FW, Ruttinger C, Salt A, et al. Using machine learning to identify factors associated with practice location of the healthcare workforce. Rural Remote Health. 2022;22(1):1-13.

Baris VK, Fu Y, Gilbreath B, Rainbow J, Fiorini LA, Love P. Harnessing machine learning to predict nurse turnover intention and uncover key predictors: a multinational investigation. J Adv Nurs. 2025.
https://doi.org/10.1111/jan.16678.

Hudays A. Machine learning models for predicting nurse turnover and turnover intention: a systematic review. Saudi J Nurs Health Care. 2025;8(6). [In press].

Burford JS, Booth RG, McIntyre A. Exploring the intersection of nursing leadership and artificial intelligence: scoping review. JMIR Nurs. 2025;8:e80085.

Chen B, Alrifai W, Gao C, Jones B, Novak L, Lorenzi N, et al. Mining tasks and task characteristics from electronic health record audit logs with unsupervised machine learning. J Am Med Inform Assoc. 2021;28(6):1168-77.

Ventura-Silva J, Martins MM, Trindade LD, Faria AD, Pereira S, Zuge SS, et al. Artificial intelligence in the organization of nursing care: a scoping review. Nurs Rep. 2024;14(4):2733-45.

Kiwanuka F, Stevanin S, Ahtisham Y, Owusu B, Nurmeksela A, Kvist T. Nurse leadership and artificial intelligence integration in nursing workforce management: a scoping review. J Adv Nurs. 2025.
https://doi.org/10.1111/jan.16789.

Yakusheva O, Bouvier MJ, Hagopian CO. How artificial intelligence is altering the nursing workforce. Nurs Outlook. 2025;73(1):102300.

Liu C, Chuang YC, Qin L, Ren L, Chien CW, Tung TH. Machine learning-based model for analysing and accurately predicting factors related to burnout in healthcare workers. BMJ Public Health. 2025;3(2).