International Journal for Multidisciplinary Research (IJFMR)

E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

The Organisational Effects of Machine-Learning Burnout Prediction Models on Employees' Psychological Well-Being

Nasser bin Said bin Mohammed Al Ismaili

Assistant Professor, College of Commerce, University of Buraimi, Sultanate of Oman

Abstract

This paper examines how the organisational deployment of machine-learning-based burnout prediction models may affect employees' psychological well-being. It addresses a growing gap in digital human resource management and AI ethics: existing work often privileges predictive accuracy and managerial utility, while the lived psychological consequences of being monitored, classified and acted upon through probabilistic systems remain comparatively under-theorised. The study adopts a theoretical-analytical design based on a critical synthesis of peer-reviewed literature on burnout, psychological well-being, HR analytics, algorithmic management and responsible AI. The Job Demands-Resources model and Conservation of Resources theory provide the primary explanatory lenses, while organisational care versus algorithmic surveillance is used as an interpretive contrast rather than an additional variable in the core model. The analysis indicates that the effect is neither inherently beneficial nor inherently harmful. When prediction is bounded by a preventive purpose, uses proportionate data, is interpretable, subject to meaningful human review and linked to supportive interventions such as workload redesign, recovery time and access to assistance, it can operate as an organisational resource and protect employees from cumulative strain. When it is opaque, individualising, punitive or surveillance-oriented, it can become an additional job demand, threaten privacy and professional identity, intensify anxiety and weaken trust. The paper therefore shifts evaluation from the narrow question of whether a model is accurate to the broader question of what its organisational use does to employees. It proposes a human-centred governance and empirical-testing framework that separates probabilistic risk estimation from clinical diagnosis, verifies actual organisational use through multiple data sources and treats well-being, rights and redress as core performance criteria.

Keywords: burnout prediction; machine learning; predictive HR analytics; psychological well-being; digital human resource management; algorithmic ethics; organisational care; algorithmic surveillance

1. Introduction

Over the past two decades, digital human resource management has moved from the administration of records towards the construction of data-rich decision infrastructures. Attendance, payroll and personnel files remain important, but they now sit alongside dashboards, workforce analytics and probabilistic models that promise to detect patterns, anticipate risks and inform decisions before problems become visible through conventional reporting. This shift has been presented as a route to more evidence-based management and a reduction in undisciplined intuition. At the same time, it has enlarged the organisation's

capacity to observe, infer and classify employees, thereby raising questions about the limits of measurement, the ethics of data collection and the position of the human person within increasingly automated systems of management (Marler and Boudreau, 2017; Angrave et al., 2016; Minbaeva, 2018).
Machine learning represents a more advanced stage of this development. In human resource settings, it can move analysis from description and diagnosis towards prediction and recommendation. Applications now extend beyond recruitment and performance appraisal to turnover, retention, absence, safety, engagement and occupational strain. Burnout prediction is particularly attractive because it appears to offer an early-warning capacity: rather than waiting for prolonged absence, performance deterioration or resignation, an organisation may seek to identify accumulating risk and intervene earlier. Yet technical feasibility does not settle organisational legitimacy. Every increase in the capacity to predict an employee's behaviour or condition is accompanied by an increase in the power to define, monitor and act upon that person through a data-derived representation (Tambe et al., 2019; Meijerink et al., 2021; Hülter et al., 2024).
The sensitivity of this development becomes most evident in relation to burnout. Burnout is not adequately described as ordinary fatigue or a temporary dislike of work. It is a persistent occupational experience commonly associated with exhaustion, cynicism or psychological distancing from work and a reduced sense of professional efficacy. It emerges through a complex interaction between chronic job demands, insufficient resources, impaired recovery and a poor fit between the worker and the organisational environment. Its consequences are simultaneously human and institutional: burnout affects meaning, dignity and the capacity to continue, while also influencing commitment, absence, turnover, performance and workplace relationships (Maslach et al., 2001; Maslach and Leiter, 2016; Demerouti et al., 2001).
Employees' psychological well-being is therefore an appropriate outcome against which to evaluate technologies presented as instruments of prevention. Psychological well-being is not a public-relations label and is not reducible to job satisfaction. It concerns the quality of psychological functioning at work, including emotional balance, meaning, competence, adaptability, positive relationships, perceived recognition and the capacity to remain engaged without chronic depletion. Nor is it equivalent to the mere absence of mental disorder. A person may not meet a clinical threshold and still experience diminished meaning, impaired resilience or a sustained loss of psychological vitality. Any system that claims to protect employees from burnout must therefore be examined in terms of its effect on this wider and more positive conception of well-being (Ryff, 1989, 2014; Dagenais-Desmarais and Savoie, 2012; Guest, 2017).
A central tension follows. A burnout prediction model can be framed as an instrument of anticipatory care that helps an organisation recognise cumulative strain, correct workload imbalances and direct support towards those who need it. The same model can also be experienced as a new form of algorithmic surveillance that turns fatigue, interaction patterns and behavioural traces into material for continuous monitoring and classification. The employee may therefore ask a question that is not answered by technical performance metrics: does the organisation see me as a person whose working conditions should be improved, or as a risk profile to be managed? The answer depends less on the mathematical form of the model than on purpose, data scope, transparency, level of inference, human oversight and the consequences attached to its output (Meijerink and Bondarouk, 2023; Giermindl et al., 2022; Leicht-Deobald et al., 2019).
The relevant literature remains fragmented. Studies of burnout prediction have largely concentrated on model construction, feature importance and comparative predictive performance. Research on psychological well-being has developed rich conceptualisations of positive functioning but has rarely

examined the experience of being subject to burnout prediction. People-analytics scholarship has debated managerial value and analytical credibility, while the ethics literature has highlighted opacity, bias, privacy and personal integrity across recruitment and algorithmic management more broadly. Direct analysis of how the actual organisational use of burnout prediction affects employees' psychological well-being is still limited, and the gap is particularly visible in Arabic-language research (Grządzielewska, 2021; Johnson and Shamroukh, 2024; Van Zyl-Cillié et al., 2024).

This paper addresses that gap by developing a conditional, human-centred explanation of the relationship. It combines the Job Demands-Resources model and Conservation of Resources theory, using the contrast between organisational care and algorithmic surveillance as an interpretive lens. Methodologically, it is a theoretical-analytical foundational study: it does not claim original field results, but it synthesises the available literature, derives analytical propositions and sets out a rigorous design for later empirical testing. Its contribution is to relocate the debate from the narrow question, 'Can the model predict burnout?' to the more consequential question, 'How does the organisational use of that prediction reshape employees' psychological experience, their resources and their relationship with the organisation?'

2. Organisational and technological context

Predictive human resource analytics should be understood as a sociotechnical practice rather than a detached statistical tool. Descriptive analytics summarise what has happened; diagnostic analytics explore possible explanations; predictive analytics estimate what may happen; and prescriptive systems recommend or trigger action. In practice, these categories overlap. A burnout risk score is produced through technical choices about labels, variables, thresholds and algorithms, but it becomes organisationally meaningful only when someone decides who may see it, how it is interpreted and what response follows. The model is therefore embedded in a chain of authority and intervention rather than operating in an institutional vacuum (Marler and Boudreau, 2017; Meijerink et al., 2021).

Burnout prediction may draw on several types of data. Human resource information systems can provide working hours, overtime, leave, absence, tenure and role changes. Operational platforms may indicate workload volume, scheduling instability, queues, shift intensity or repeated deadline compression. Pulse surveys can capture exhaustion, perceived support, psychological safety and work-life interference, while climate and culture surveys may reveal trust, voice and perceived responsiveness. Some systems also seek to infer strain from digital activity or communication metadata. The technical availability of such data does not make every use necessary or legitimate. The principle of data minimisation is especially important where signals approach the psychological domain, and content-level monitoring should not be treated as an ordinary extension of workforce analysis (Bankins, 2021; Bujold et al., 2024).

Table 1. Organisational data sources, potential signals and safeguards

Data source	Potential burnout-related signal	Analytical value	Principal risk	Minimum safeguard
Working-time and scheduling records	Long hours, overtime, unstable shifts, insufficient recovery	Identifies chronic demand and temporal pressure	Context-free interpretation; surveillance	Use work-design context; aggregate where possible; limit retention

Data source	Potential burnout-related signal	Analytical value	Principal risk	Minimum safeguard
Leave and absence data	Changes in absence or delayed leave use	May indicate accumulating strain at team or cohort level	Health inference; stigma; reverse causality	Do not treat as diagnosis; protect access; examine alternative explanations
Workload and operational data	Backlogs, queue intensity, deadline compression, task volatility	Directs attention to job demands	Productivity monitoring can displace well-being purpose	Separate well-being analysis from performance control
Pulse and climate surveys	Exhaustion, support, autonomy, fairness and recovery perceptions	Captures lived experience and organisational resources	Low participation; response fear; common-method bias	Voluntary confidential collection; minimum reporting groups
HR and career records	Role change, tenure, mobility, managerial span	Provides occupational and organisational context	Proxy discrimination; historical bias	Necessity test; subgroup fairness review; exclude protected attributes from adverse use
Communication or collaboration metadata	Meeting load, after-hours activity, fragmented work	May reveal coordination demands	Intrusive inference; chilling effects	Prefer metadata aggregates; exclude message content; explicit purpose notice
Employee-assistance or health-related data	Demand for support or occupational-health contact	Could inform aggregate capacity planning	Extreme sensitivity; deterrence from help-seeking	Strict separation, confidentiality and no individual model use

Source: Author synthesis from the literature reviewed in Sections 2 and 3.

The level at which predictions are produced also matters. Team-level or unit-level analysis can direct attention towards the organisational sources of risk, such as persistent understaffing, unstable scheduling or a managerially produced overload. Individual-level classification creates a different ethical and psychological situation because it can affect reputation, opportunity and self-understanding. An aggregated signal that a department is experiencing sustained pressure may invite work redesign; an individual label of 'high burnout risk' may invite support, but it may also produce stigma or inappropriate employment consequences. The closer the model moves towards the individual, the stronger the requirements for necessity, explanation, human review, confidentiality and redress.

A further distinction concerns actual use. An experimental model that is not connected to decisions is different from an operational system that generates regular alerts and changes work allocation, referral, supervision or performance management. Studies based only on employees' general perceptions of artificial intelligence may therefore mis-specify the independent variable. The relevant construct in this paper is the actual organisational use of a burnout prediction model: the model is active, produces outputs, is integrated into a recognised process and is capable of generating a real response. Perception remains important because employees live the consequences of the system, but organisational verification is needed to establish what the technology actually does.

The organisational context also shapes whether a model is read as care or surveillance. High-trust environments with credible confidentiality, participatory governance and a history of supportive intervention may make employees more willing to interpret prediction as protection. In low-trust environments, even a well-intentioned model may be understood as a mechanism for identifying weakness, intensifying supervision or transferring responsibility for structural overload to individuals. The same technical architecture can therefore produce different psychological outcomes across organisations. This contextual variability explains why predictive accuracy cannot be treated as a sufficient indicator of organisational value.

3. Conceptual foundations and literature review

The study brings together four bodies of scholarship that are frequently discussed separately: burnout and occupational health, psychological well-being at work, predictive HR analytics and algorithmic ethics. Integrating them requires conceptual discipline. Burnout is the phenomenon of concern; a machine-learning model is an organisational instrument that estimates risk; psychological well-being is the human outcome; and governance conditions explain how the instrument acquires a supportive or harmful meaning. The following sections establish these distinctions and identify the unresolved analytical problem.

3.1 Burnout as an occupational phenomenon

The modern literature conceptualises burnout as an occupational syndrome shaped by the relationship between the individual and the work environment. The most influential account identifies exhaustion, cynicism or distancing and reduced professional efficacy as central dimensions. This formulation does not deny individual variation, but it resists explaining burnout as a personal weakness detached from job design, workload, control, recognition and social support. The Job Demands-Resources tradition similarly treats burnout as a likely outcome when demanding aspects of work chronically consume physical, emotional or cognitive energy without sufficient compensating resources (Maslach et al., 2001; Demerouti et al., 2001; Bakker et al., 2014).

This understanding matters for prediction. If burnout is represented only through individual attributes, the model may reproduce an individualising account of what is partly an organisational problem. By contrast, models that identify workload concentration, weak managerial support or recurring scheduling instability can help reveal modifiable conditions. The distinction is not merely technical. It determines whether the output is used to ask, 'What is wrong with this employee?' or 'What is happening in this work system?' A credible approach must therefore maintain the systemic nature of burnout even when individual data are used.

Table 2. Conceptual distinctions relevant to the study

Concept	Defining feature	What it is not	Relevance to this research
Occupational burnout	A work-related syndrome associated with exhaustion, mental distance or cynicism and reduced professional efficacy	A momentary mood, personal weakness or automatic psychiatric diagnosis	The occupational phenomenon whose risk the model seeks to estimate
Burnout prediction model	A probabilistic system trained to estimate elevated burnout-related risk	A clinical diagnosis or certain statement about an individual's inner state	The independent construct when operationally deployed by an organisation
Predictive HR analytics	Use of workforce data to estimate future or latent organisational outcomes	Merely descriptive reporting or a value-neutral technical activity	The broader organisational field in which burnout prediction is located
Psychological well-being at work	Positive functioning involving meaning, competence, adaptation, relationships and psychological balance	Job satisfaction alone or the mere absence of disorder	The dependent construct and human outcome of interest
Organisational care	Institutional practices that protect employees and provide resources through credible action	Supportive rhetoric without resource-bearing intervention	Interpretive lens for the preventive pathway
Algorithmic surveillance	Systematic data-driven observation and inference that expands managerial visibility and control	Only overt monitoring or malicious intent	Interpretive lens for the strain-producing pathway
Probabilistic prediction	An estimate whose validity depends on data, context and threshold	A causal explanation or deterministic judgement	Defines the epistemic limits of model outputs

Source: Author synthesis based on Maslach et al. (2001), Ryff (1989), Dagenais-Desmarais and Savoie (2012), Marler and Boudreau (2017) and the algorithmic-management literature.

Burnout risk estimation must also remain distinct from clinical or medical diagnosis. A predictive model produces a probability, classification or warning based on the variables available to it. It does not conduct a clinical interview, apply a complete diagnostic framework or assume the responsibilities of a qualified health professional. Treating a probability score as a medical fact would be an epistemic error and an ethical overreach. The model may inform prevention or prompt a supportive conversation, but it should not be allowed to confer a clinical label or determine treatment (Maslach and Leiter, 2016; Grządzielewska, 2021).

3.2 Machine learning and predictive HR analytics

Machine-learning models learn statistical patterns from historical or current data and use those patterns to estimate an outcome for new observations. In burnout research, the target may be a high score on an exhaustion measure, membership in a risk category or a composite outcome associated with emotional exhaustion and disengagement. Studies have explored decision trees, Bayesian approaches and other supervised learning techniques, often comparing predictive performance and identifying organisational or cultural variables that contribute to risk (Johnson and Shamroukh, 2024; Van Zyl-Cillié et al., 2024).

The attraction of machine learning lies in its ability to combine multiple signals and detect interactions that may be difficult to observe through ordinary supervision. Yet its outputs remain conditional on the quality of the target, the representativeness of the data and the stability of the organisational context. A model trained in one sector, occupational group or culture may not transfer reliably to another. Historical data may encode unequal treatment, while proxy variables may reproduce differences that were never intended as legitimate grounds for action. Model drift can also occur as workload, technology, staffing or employee behaviour changes. Technical validation is therefore a continuing responsibility rather than a one-off approval (Tambe et al., 2019; Minbaeva, 2018; Bujold et al., 2024).

Interpretability varies across techniques. Some models can provide relatively accessible explanations, such as the variables most associated with a prediction, whereas others are difficult for non-specialists to understand. Even where a technical explanation is possible, organisational explanation remains necessary: employees need to know the purpose of the system, the types of data used, who receives the output and what consequences may follow. Opacity is not confined to code. It can arise from secrecy about purpose, access, thresholds and decision rights (Burrell, 2016).

Predictive HR analytics is consequently not equivalent to objective knowledge. It is a structured form of probabilistic inference produced through human choices about what to measure and how to act. Analytical credibility requires more than high accuracy. It requires construct validity, calibration, error analysis, fairness assessment, transparent limitations and a clear account of organisational contribution. A model that predicts a survey score while worsening employee trust may be statistically successful and organisationally unsuccessful at the same time.

3.3 Psychological well-being at work

Psychological well-being has been developed through both general and work-specific traditions. Ryff's eudaimonic account shifted attention from momentary pleasure towards positive functioning, including purpose, autonomy, growth, self-acceptance and positive relationships. Organisational research has adapted this logic to the workplace by examining interpersonal fit, thriving, competence, recognition and the desire to remain involved. These perspectives make well-being a substantive outcome rather than a residual category defined only by the absence of distress (Ryff, 1989, 2014; Dagenais-Desmarais and Savoie, 2012).

The distinction from job satisfaction is important. Job satisfaction usually concerns an evaluative attitude towards aspects of the job, such as pay, supervision, promotion or task content. It overlaps with well-being but does not exhaust it. An employee can be reasonably satisfied with contractual conditions while experiencing depleted meaning, anxiety or weak relational safety. Conversely, a person may be dissatisfied with a particular condition yet retain a strong sense of competence and purpose. Reducing well-being to satisfaction would therefore conceal the deeper psychological consequences of predictive monitoring (Guest, 2017).

Well-being is also relational and institutional. It is influenced by job demands, resources, leadership, justice, voice, workload, recognition and the degree to which the organisation treats employees as ends rather than merely as instruments of performance. HR practices can produce trade-offs among performance, health and relationships, and beneficial outcomes cannot be assumed simply because a practice is labelled innovative or evidence-based (Grant et al., 2007; van de Voorde et al., 2012). This relational character makes psychological well-being especially suitable for evaluating burnout prediction: the technology may change not only workload or support, but the employee's understanding of the employment relationship itself.

3.4 Organisational care, algorithmic surveillance and ethical governance

The contrast between organisational care and algorithmic surveillance is used in this paper as an interpretive lens. Organisational care refers to the credible use of information to reduce preventable harm, improve working conditions and provide resources in a manner that respects dignity, privacy and employee voice. Algorithmic surveillance refers to the use of data-driven systems to observe, classify and influence workers under conditions of limited transparency or control. These are not mutually exclusive properties of a technology; they are alternative meanings and practices that can coexist or compete within the same organisation (Meijerink and Bondarouk, 2023).

Research on people analytics and algorithmic management has documented a duality. Algorithms can support consistency, identify patterns and improve resource allocation, but they can also intensify measurement, centralise power and reduce the space in which employees explain their circumstances. The 'dark sides' of people analytics include privacy intrusion, decontextualised classification, function creep and the transformation of data collected for one purpose into material for another. Decisions may appear objective even when they reflect contested choices about variables, thresholds and acceptable errors (Giermindl et al., 2022; Leicht-Deobald et al., 2019).

Fairness, trust and emotion are central to how employees respond. Individuals are more likely to accept an algorithmic decision when they understand its basis, perceive a fair procedure and retain some form of voice or challenge. Conversely, opacity and the inability to contest a classification can produce anxiety, anger and distrust. Evidence from algorithmic management, digital Taylorism and automated hiring is not identical to burnout prediction, but it provides a relevant analytical warning: systems that continuously evaluate people can affect psychological well-being even before any formal adverse decision occurs (Lee, 2018; Konuk et al., 2023; Kinowska and Sienkiewicz, 2023; Zhang and Yencha, 2022; Ochmann et al., 2024).

Responsible AI frameworks respond through principles such as purpose limitation, proportionality, privacy, fairness, explainability, human oversight and accountability. However, principles become meaningful only when converted into governance arrangements, decision rights and enforceable boundaries. A statement that a model is 'for well-being' is insufficient if data are extensive, outputs are hidden or high-risk classifications influence promotion and performance processes. Ethical governance must therefore address the complete chain from data collection to intervention and redress (Mittelstadt et al., 2016; Bankins, 2021; Bar-Gil et al., 2024; Bujold et al., 2024).

3.5 Research gap

The research gap is theoretical, organisational and methodological. Theoretical work has not adequately connected burnout prediction to a positive and multidimensional account of employee psychological well-being. Organisational research often discusses predictive systems as if their meaning resided in the algorithm, while the decisive mechanisms may lie in who controls the model, how it is framed and what follows from the output. Methodologically, studies frequently examine prediction performance, general attitudes towards AI or hypothetical scenarios without verifying actual organisational use and its consequences.

Table 3. Literature synthesis and research gap

Theme	Established insight	Unresolved limitation	Implication for this study
Burnout and work design	Chronic demands and insufficient resources contribute to exhaustion and disengagement	Prediction studies can individualise risk and understate structural causes	Interpret model outputs through job demands, resources and organisational context
Psychological well-being	Well-being is broader than satisfaction and includes positive psychological functioning	Technology studies often use narrow or inconsistent outcome measures	Treat work-related psychological well-being as an integrated positive construct
HR analytics and AI	Predictive tools can support evidence-informed decisions and earlier intervention	Research often privileges adoption, accuracy or managerial value	Focus on actual organisational use and employee psychological consequences
Burnout prediction	Machine learning can identify patterns associated with burnout-related outcomes	Evidence is sectorally concentrated and rarely follows deployment into organisational action	Separate technical feasibility from organisational and human effectiveness
Algorithmic management	Opacity, monitoring and automated classification can affect autonomy, trust and well-being	Much evidence comes from recruitment, platforms or broader management systems	Transfer mechanisms cautiously and test them in burnout-specific settings
Responsible AI governance	Transparency, fairness, privacy, explainability and human oversight are central	Principles are frequently detached from intervention and measurable employee outcomes	Evaluate the complete prediction-to-action pathway
Research gap	Relevant literatures offer complementary insights	They are seldom integrated into a direct model linking operational burnout prediction to psychological well-being	Develop a conditional framework and a multi-source empirical design

Source: Author synthesis from the literature reviewed in Section 3.

This paper addresses the gap by treating organisational deployment as the substantive independent variable. The model is not a neutral object added to the workplace; it can alter job demands, access to resources, perceived control, privacy, trust and professional identity. The paper also preserves a strict distinction between probability and diagnosis and proposes multiple-source verification for later field research. The result is a conditional framework capable of explaining why similar systems may improve well-being in one setting and undermine it in another.

4. Research questions and analytical propositions

The central research question is: How does the organisational use of machine-learning-based burnout prediction models affect employees' psychological well-being?

Five sub-questions organise the analysis. They examine the conceptual and technical characteristics of burnout prediction, the organisational processes through which models are built and used, the potential benefits and risks, the conditions that explain variation in employee experience and the methodological design required for later field testing.

1. What is meant by a machine-learning-based burnout prediction model, and which technical and organisational characteristics distinguish probabilistic prediction from psychological or medical diagnosis?
2. How are these models built and deployed within organisations, what types of data and indicators do they use, and what preventive or intervention-related uses can legitimately follow from their outputs?
3. What benefits may the models offer to psychological well-being when they are used within a supportive preventive framework, and what psychological, ethical and organisational risks arise when use is surveillance-oriented, punitive or stigmatising?
4. How do transparency, privacy, predictive fairness, explainability, employee participation and the nature of the organisational intervention help explain variation between supportive and harmful effects?
5. What methodological design can test the relationship between actual organisational use and psychological well-being without overextending the foundational model?

The paper advances one central analytical proposition. The organisational use of burnout prediction models has a dual and conditional effect on psychological well-being. It tends to support well-being when the system is preventive, proportionate, transparent, privacy-preserving and linked to humane organisational interventions. It tends to weaken well-being when the system is opaque, surveillance-oriented, stigmatising or punitive, or when it intrudes on privacy without a credible supportive purpose.

Two directional propositions follow without expanding the core model into a complex network of mediators and moderators at this foundational stage. First, supportive and preventive organisational use is expected to be positively associated with psychological well-being. Secondly, surveillance-oriented, punitive or stigmatising use is expected to be negatively associated with psychological well-being. Transparency, privacy, fairness, explainability, employee participation and the nature of the intervention are treated here as interpretive conditions that describe the form of use. Later research may test them as mediators or moderators, but the initial empirical model should retain analytical economy.

This formulation avoids treating the independent variable as an ethically neutral block. A predictive model never operates outside an institutional setting. Its purpose, data, access rules, intervention protocol and accountability structure define what 'use' means. The relevant question is therefore not only whether the algorithm predicts an outcome, but how the employee and the work system are redefined after prediction enters the organisation's approach to psychological risk.

5. Theoretical framework

The theoretical framework must explain how a technical system can become either a resource or a stressor. Models of technology acceptance are not sufficient because the study is not primarily concerned with whether employees intend to use the system. Employees may have little choice about its deployment. The issue is the psychological effect of being subject to an organisational practice. For that reason, the paper combines the Job Demands-Resources model with Conservation of Resources theory and uses organisational care versus algorithmic surveillance as an interpretive contrast (Bakker and Demerouti, 2007, 2017; Hobfoll, 1989).

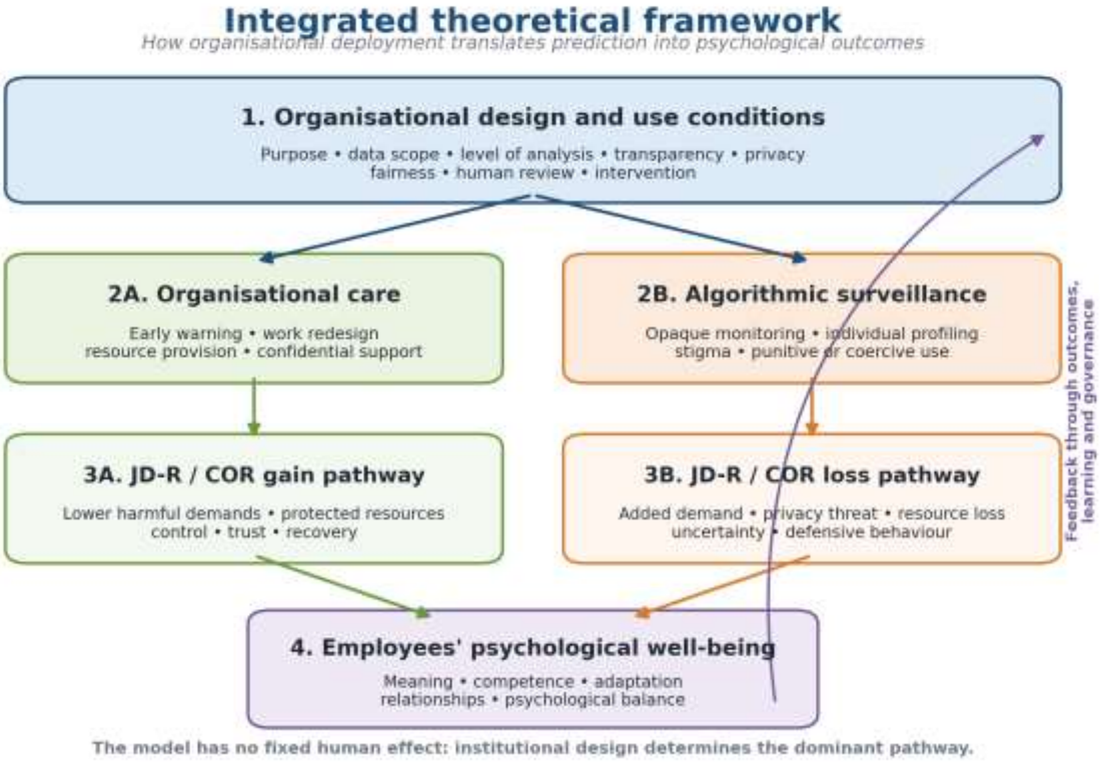
The Job Demands-Resources model proposes that jobs can be understood through demands that require sustained effort and resources that help employees achieve goals, reduce the cost of demands or support growth. Burnout becomes more likely when demands remain high and resources are insufficient. A burnout prediction system can alter this balance. When it leads to workload redistribution, improved scheduling, recovery time, managerial support or access to assistance, it functions as an organisational mechanism for creating or restoring resources. When it intensifies monitoring, produces anxiety about classification or adds unexplained expectations, it becomes an additional demand (Demerouti et al., 2001; Bakker et al., 2014).

Conservation of Resources theory adds an important subjective dimension. Individuals seek to acquire, retain and protect valued resources, and stress occurs when resources are threatened, lost or invested without adequate return. Resources include time, energy, control, privacy, status, trust, professional reputation, social support and psychological safety. A supportive prediction process may prevent further resource loss and initiate a gain cycle: the employee receives timely support, regains control and experiences the organisation as responsive. A surveillance-oriented process can threaten several resources at once, including privacy, autonomy, reputation and the ability to define one's own professional identity (Hobfoll, 1989).

The care-surveillance contrast explains the organisational meaning attached to the system. Care is credible when the declared purpose is matched by practice: limited data, clear communication, supportive consequences and meaningful safeguards. Surveillance is experienced when the organisation expands observation, conceals decision logic, denies challenge or connects risk classification to discipline and opportunity. The lens does not constitute a third independent variable. It interprets the way in which the predictive model enters the work system and thereby modifies demands and resources.

The integrated sequence is shown in Figure 1. Organisational design and use conditions shape whether the system is experienced primarily as care or surveillance. That experience alters the balance of job demands and resources and activates resource gain or loss cycles. These mechanisms influence psychological well-being, including emotional balance, meaning, competence, adaptability, trust and perceived respect. Outcomes then feed back into organisational legitimacy, data quality and the continuing validity of the model. Low trust, for example, may reduce honest survey participation and thereby weaken the data on which the system depends. The framework is therefore dynamic rather than a one-way technical intervention.

Figure 1. Integrated theoretical framework for burnout prediction and psychological well-being



Source: Author synthesis.

The framework also clarifies why predictive accuracy is necessary but insufficient. A poorly calibrated model cannot support responsible intervention. Yet a well-calibrated model may still harm well-being if it is deployed in a low-trust, punitive environment. Technical quality and organisational quality are distinct but interdependent. A full evaluation must ask whether the model identifies relevant risk, whether it changes work in a beneficial way and whether employees retain dignity, voice and control throughout the process.

6. Methodology

The present paper is a theoretical-analytical foundational study. It does not report original interviews, surveys, experiments or organisational datasets, and it does not present statistical findings as though a field study had already been conducted. Its current purpose is to synthesise relevant scholarship, clarify the relationship between the two core variables, derive testable propositions and formulate an ethically viable design for subsequent empirical research. This distinction is methodologically important because conclusions drawn from theory and prior literature must not be represented as direct causal estimates produced by the paper itself (Guest, 2017; Marler and Boudreau, 2017).

The review is structured rather than systematic in the PRISMA sense. It brings together seminal and recent peer-reviewed work in four overlapping areas: burnout and the Job Demands-Resources tradition; psychological well-being in organisational settings; HR analytics and machine learning; and algorithmic ethics, fairness, surveillance and responsible AI. The synthesis is critical rather than aggregative. Sources are organised around the central research question, and evidence from adjacent domains, such as automated hiring, is used cautiously and identified as analytically relevant rather than direct evidence about burnout prediction.

A later field study should adopt a quantitative explanatory design supported by multiple data sources. The statistical test should remain focused on the relationship between actual organisational use of burnout prediction and employee psychological well-being. An employee survey can measure lived experience and well-being, but organisational verification is required to establish whether a functioning model exists, what data it uses, how often it produces outputs and whether those outputs are integrated into real interventions. This prevents the independent variable from collapsing into general attitudes towards AI or an employee's assumption that a model may exist (Minbaeva, 2018; Hülter et al., 2024).

The proposed population consists of employees in medium and large organisations that use advanced HR information systems or workforce analytics and operate in sectors where data intensity and occupational strain make burnout prediction plausible. Relevant settings include healthcare, financial services, telecommunications, shared-service centres, higher education and selected public-sector units with high cognitive or operational demands. A multi-organisation design is preferable because organisational culture and policy shape the meaning of use. An indicative sample of 450-600 employees across six to eight organisations would provide a reasonable basis for confirmatory factor analysis and a parsimonious structural model, while a smaller group of HR, analytics or well-being officials would complete an organisational verification instrument.

Table 4. Proposed empirical design and evidence requirements

Design element	Proposed approach	Purpose	Evidence requirement
Study design	Multi-organisation, explanatory quantitative study with organisational verification	Test the direct relationship while documenting actual deployment	Operational models, not hypothetical scenarios or adoption intentions
Study population	Employees in medium and large organisations using advanced workforce analytics	Ensure exposure to a relevant organisational system	Documented model, output cycle and intervention pathway
Sample	Approximately 450-600 employees across 6-8 organisations, subject to power analysis	Support confirmatory measurement and stable direct-effect estimation	Balance sectors, roles and supervisory status where feasible
Employee instrument	Work-related psychological well-being scale plus deployment-experience items	Measure the dependent construct and lived consequences	Confidential administration separated from performance management
Organisational audit	Structured verification of target, data, model, level, users, actions and governance	Prevent perceived AI use from substituting for actual use	Named responsible function and auditable operational evidence
Model documentation	Performance, calibration, subgroup error, threshold and drift records	Characterise technical validity and unequal risk	Current model card or equivalent assurance record

Design element	Proposed approach	Purpose	Evidence requirement
Analysis	CFA followed by regression or structural equation modelling of the direct path	Test construct validity and the principal relationship	Reliability, convergent/discriminant validity and sensitivity checks
Ethical protection	Voluntary participation, data minimisation, no individual results to employer	Protect participants in a sensitive employment context	Independent ethics review and explicit separation from organisational monitoring

Source: Author methodological proposal. Sample size is indicative and must be confirmed through a formal power analysis.

Two complementary instruments are proposed. The employee questionnaire would measure psychological well-being at work and employees' experience of the model and its interventions, while avoiding requests for unnecessary sensitive information. The organisational audit card would verify the active model, data categories, periodic outputs, integration into decision processes, level of inference and available safeguards. The dependent variable should be measured as an integrated construct using a validated work-related psychological well-being scale, with subdimensions used for construct validity rather than fragmenting the main model (Dagenais-Desmarais and Savoie, 2012).

Validity and reliability procedures should include expert review, translation and back-translation where an Arabic instrument is used, a pilot study, Cronbach's alpha and McDonald's omega, confirmatory factor analysis, composite reliability and average variance extracted. Descriptive analysis and correlations would precede the direct-effect test. Structural equation modelling is appropriate if the constructs are treated as latent variables; linear regression remains acceptable for a simpler observed-score model. Ethical safeguards include voluntary participation, separation of research from performance management, no transmission of individual responses to the employer, minimal data collection and explicit communication that participation does not constitute consent to the organisation's predictive system.

7. Burnout prediction models in organisational settings

Burnout prediction models occupy a particularly sensitive position within predictive human resource analytics because they infer a risk connected to a person's psychological and occupational experience. Their organisational value cannot be assessed solely through algorithm choice or a headline accuracy score. The target definition, data provenance, level of analysis, threshold, output format, human review and intervention protocol together constitute the operational system. The same statistical model can therefore have substantially different consequences when deployed as a team-level signal for work redesign, an individual-level prompt for confidential support or an undisclosed score used in performance management. This section examines the defining characteristics of these models, the processes through which they are developed and deployed and the benefits and risks that follow.

7.1 Concept and defining characteristics

For the purposes of this study, a machine-learning-based burnout prediction model is an algorithmic system trained on organisational, human resource, survey or work-pattern data to estimate the probability that an employee, team or organisational unit is experiencing, or may develop, elevated burnout-related risk. Its output may take the form of a probability, risk band, ranking or early-warning indicator. Such an

output is an inference under uncertainty rather than a definitive statement about an individual's inner state. Research on burnout modelling demonstrates that machine-learning techniques can identify associations among organisational culture, emotional exhaustion and related risk factors, but it does not convert an organisational model into a clinical instrument (Grządzielewska, 2021; Johnson and Shamroukh, 2024).
This distinction is epistemically and ethically decisive. Burnout is an occupational phenomenon with recognised dimensions and measurement traditions, whereas clinical diagnosis requires professional judgement, validated diagnostic criteria and an appropriate healthcare context. A risk score may indicate that observed patterns resemble those associated with elevated burnout in a training dataset; it cannot establish that a particular employee has a psychiatric disorder, identify the full cause of distress or determine appropriate treatment. Treating a probability as a diagnosis would overstate what the model knows and could expose employees to stigma or inappropriate intervention (Maslach and Leiter, 2016).
These models share several technical characteristics. They are probabilistic rather than declarative; commonly multi-source rather than dependent on a single measure; capable of being updated as data and conditions change; and variable in interpretability. Some approaches, such as decision trees or regularised regression, may permit relatively accessible explanations. Others may provide stronger predictive performance in a particular dataset while making it harder to explain how a score was produced. No algorithm is context-free. A model trained in one occupation, country, organisation or time period may perform differently when job design, reporting behaviour, workforce composition or managerial practices change.
Their organisational characteristics are equally important. A model is embedded in a chain of authority: someone specifies the target, authorises data access, selects the threshold, receives the output and decides what follows. These decisions distribute informational power inside the organisation. They determine whether the system directs attention towards excessive demands and weak resources or towards allegedly high-risk individuals; whether employees can understand and contest an output; and whether a risk signal generates support, monitoring or sanction. Algorithmic HRM is therefore not merely an analytical technique but a form of organised decision-making that can reshape autonomy, integrity and the employment relationship (Meijerink et al., 2021; Leicht-Deobald et al., 2019).
Error also has a substantive meaning. A false positive may label an employee or team as high risk when the underlying condition is absent, potentially producing anxiety, altered managerial treatment or unnecessary disclosure. A false negative may leave genuine strain unrecognised and create unwarranted confidence in the system. Aggregate accuracy can conceal unequal error rates across occupational, demographic or contractual groups. Evaluation must therefore consider calibration, sensitivity, specificity, subgroup performance and the consequences attached to each type of error rather than treating technical performance as a single neutral number (Barocas and Selbst, 2016; Mittelstadt et al., 2016).

Table 5. Machine-learning model characteristics and organisational implications

Characteristic	Technical meaning	Organisational implication	Required control
Probabilistic output	Produces a score, class or risk band under uncertainty	Can be mistaken for certainty or diagnosis	Uncertainty statement; trained human interpretation

Characteristic	Technical meaning	Organisational implication	Required control
Multi-source data	Combines HR, survey, scheduling or operational variables	Can improve pattern detection but expand intrusion	Necessity and proportionality assessment; data minimisation
Context dependence	Performance depends on population, period and work setting	External models may not transfer to local occupations or cultures	Local and temporal validation; documented population limits
Threshold selection	Converts continuous risk into an alert or category	Determines false-positive and false-negative consequences	Consequence-sensitive threshold review and monitoring
Unequal error	Performance may vary across groups or roles	Can distribute stigma or missed support unequally	Subgroup calibration and error analysis; remedial action
Variable interpretability	Some models are easier to explain than others	Opacity can reduce contestability and trust	Select interpretable method where stakes require; provide reasoned explanation
Model drift	Data relationships change after deployment	Past validity may create false confidence	Scheduled monitoring, revalidation and suspension rules
Feedback effects	Decisions based on scores alter later data and behaviour	Can create self-fulfilling classifications or gaming	Separate evaluation data; monitor behavioural and organisational consequences

Source: Author synthesis based on Burrell (2016), Mittelstadt et al. (2016), Bankins (2021) and the burnout-prediction literature.

7.2 Model development, data and deployment

Model development begins with a defensible definition of the prediction target. A target may be a validated burnout-scale threshold, a subsequent survey outcome or a carefully specified composite indicator. It should not be replaced by convenient organisational proxies such as absence, reduced output or turnover intention without demonstrating their conceptual relationship to burnout. Poor target construction can cause the model to learn managerial assumptions rather than the phenomenon it purports to estimate. It may also individualise a structural problem by interpreting the consequences of understaffing, poor leadership or unstable scheduling as employee attributes.

The next stage is data selection and preparation. Potential sources include working hours, overtime, shift instability, leave patterns, workload indicators, pulse surveys, supervisory support, organisational climate and job-control measures. Some sources may be informative at team level while being disproportionately intrusive at individual level. The principle of data minimisation requires the organisation to collect and retain only what is necessary for a clearly stated preventive purpose. Digital traces should not be included merely because they are available, and highly sensitive communications or behavioural exhaust should not be treated as costless raw material. Where the objective can be met through aggregated or de-identified indicators, that design should be preferred.

Training and validation require separation between development and test data, attention to missingness and class imbalance, transparent feature documentation and external or temporal validation where possible. Feature importance can indicate which variables contributed most strongly to a model's prediction, but it does not establish causal responsibility. Findings that organisational culture, fatigue or perceptions of managerial responsiveness are predictive should prompt organisational inquiry; they should not be read as proof that changing one variable will necessarily prevent burnout. Recent sectoral studies illustrate the potential of such models while also underscoring that their performance and salient predictors are context-dependent (Johnson and Shamroukh, 2024; Van Zyl-Cillié et al., 2024).

Deployment is the decisive transition from analytical possibility to organisational consequence. The institution must specify the level at which results are reported, who receives them, how frequently they are generated and what actions are permitted. A team-level dashboard used to review staffing may pose a different risk profile from an individual alert sent to a line manager. A supportive protocol should separate risk estimation from disciplinary and performance processes, require trained human review and define a proportionate response. Without such a protocol, a technically valid model may remain an ungoverned experiment or become an informal source of consequential judgements.

Post-deployment monitoring is equally necessary. Work conditions, data-generating processes and employee responses can change, creating model drift and feedback effects. Employees may alter behaviour because they know or suspect that it is being monitored; managers may make decisions based on the score, thereby changing the outcomes on which future models are trained. Institutions should monitor calibration, subgroup error, complaints, intervention quality and unintended effects, and should suspend use when evidence no longer supports the model's purpose. Responsible AI in HRM therefore requires an auditable lifecycle rather than a one-off technical approval (Bankins, 2021; Bar-Gil et al., 2024).

7.3 Organisational benefits and ethical risks

The most persuasive potential benefit is earlier recognition. Burnout often develops through cumulative exposure to high demands, inadequate recovery and insufficient resources. Conventional indicators may become visible only after prolonged absence, deteriorating performance or resignation. A carefully designed model may help identify patterns at a stage when work can still be redesigned and resources restored. Its value lies not in predicting an inevitable breakdown, but in enabling a timely and proportionate organisational response.

A second benefit is prevention through work redesign. If elevated risk is concentrated in a unit, shift pattern, occupation or managerial context, the model may direct attention towards workload, scheduling, role conflict, staffing, autonomy and supervisory support. This reframes burnout from an individual weakness to a signal about the organisation of work. Under the Job Demands-Resources model, such action can reduce harmful demands and strengthen resources; under Conservation of Resources theory, it can interrupt loss spirals before employees exhaust time, energy, confidence and social support.

A third benefit is more deliberate allocation of support. Organisations frequently offer generic well-being programmes that are disconnected from the distribution of risk. Predictive evidence may help target manager training, occupational-health capacity, recovery time or technical assistance where need is greatest. Yet targeting is defensible only where it increases access and does not create a hidden category of supposedly fragile employees. Resource allocation should be guided by need while preserving dignity and equal opportunity.

The first major risk is privacy intrusion. Burnout-related modelling may draw on data close to employees' health, emotion, relationships and everyday behaviour. An expansive data appetite can transform normal work activity into continuous psychological inference. The fact that data can be collected does not make its collection necessary or legitimate. Purpose limitation, proportionality, retention limits, access controls and separation from unrelated HR decisions are therefore foundational rather than optional safeguards (Mittelstadt et al., 2016; Bujold et al., 2024).

The second risk is the experience of surveillance. Even a model announced as supportive can increase strain when employees believe that their pace, interactions, absence or expressions of fatigue are continuously assessed. Digital Taylorism and algorithmic management research suggests that intensive monitoring can reduce autonomy, trust and psychological safety, thereby creating a new demand in the very environment the system is intended to improve (Konuk et al., 2023; Kinowska and Sienkiewicz, 2023; Zayid et al., 2024).

The third risk is stigma and unfair classification. A high-risk label may influence a manager's expectations, restrict access to demanding assignments or promotion, or alter how employees understand their own professional identity. These consequences can occur even when the prediction is uncertain or wrong. Historical data and proxy variables may also reproduce existing inequalities, while apparently objective scores conceal value judgements embedded in target definitions and thresholds. Fairness must therefore be evaluated in both model performance and organisational consequence.

A fourth risk is function creep. Data collected for prevention may later be used for productivity management, absence control, workforce reduction or insurance-related decisions. Supportive language cannot compensate for punitive practice. Clear prohibited-use rules, independent oversight and enforceable deletion or separation requirements are required to prevent the system from migrating towards purposes that employees did not reasonably anticipate. The central organisational question is thus not simply whether prediction is possible, but whether the institution can govern the power created by prediction.

8. Effects on employees' psychological well-being

The effect of burnout prediction on psychological well-being can be understood through two analytically opposed pathways. In the supportive and preventive pathway, prediction helps the organisation identify harmful conditions, mobilise resources and intervene in ways that protect dignity and agency. In the surveillance and resource-loss pathway, prediction intensifies monitoring, uncertainty, stigma and the threat of consequential classification. These are not properties of two different technologies. They are alternative organisational translations of the same technical capability. The balance between them depends on institutional purpose, governance and the quality of post-prediction action.

Table 6. Dual pathways from burnout prediction to psychological well-being

System element	Supportive and preventive pathway	Surveillance and resource-loss pathway	Expected well-being implication
Purpose	Improve work and prevent cumulative harm	Increase visibility, control or workforce sorting	Care supports trust; control increases threat

System element	Supportive and preventive pathway	Surveillance and resource-loss pathway	Expected well-being implication
Level of analysis	Team/system level unless individual support is necessary	Individual ranking or hidden profiling	System focus reduces stigma; individual profiling may heighten anxiety
Data practice	Minimal, relevant, transparent and protected	Expansive, opaque and reused across purposes	Control over information protects resources; intrusion threatens them
Interpretation	Probabilistic signal for inquiry and contextual review	Quasi-diagnostic label or proof of fragility	Uncertainty and dignity are preserved or eroded
Human role	Meaningful review with authority to reject output	Rubber-stamping or automated consequence	Agency and procedural justice rise or fall
Intervention	Workload redesign, resource provision and voluntary confidential support	Closer monitoring, opportunity restriction or sanction	Resources are gained or lost
Feedback	Outcomes and unintended effects inform revision or withdrawal	Scores become self-validating and use expands	Learning can strengthen legitimacy; unchecked feedback can deepen strain

Source: Author synthesis using the Job Demands-Resources model, Conservation of Resources theory and the care-surveillance lens.

8.1 The supportive and preventive pathway

In a supportive pathway, the model is explicitly positioned as an early-warning aid for improving work rather than judging personal resilience. Employees are informed about its purpose and limits; data are proportionate; results undergo human review; and any response begins with a confidential, non-punitive conversation. Such a design can reduce the surprise and severity of crisis by making risk visible before exhaustion becomes entrenched. It may also communicate that the organisation accepts responsibility for the conditions under which work is performed.

The Job Demands-Resources model explains the first mechanism. A predictive signal becomes beneficial when it triggers a reduction in excessive demands or an increase in resources. Examples include redistributing workload, stabilising shifts, clarifying roles, increasing staffing, improving supervisory support, protecting recovery time or expanding discretion. The model itself does not create well-being; the organisational changes that follow from it do. Where those changes are credible, employees may experience greater control, competence, trust and capacity to recover (Bakker and Demerouti, 2017).

Conservation of Resources theory explains a second mechanism. Early support can protect valued resources before they are lost and can initiate a gain spiral. A reduction in workload may restore energy;

a respectful conversation may preserve status and dignity; access to support may strengthen coping capacity; and a clear explanation may reduce uncertainty. These gains can reinforce one another, improving psychological balance and the ability to engage meaningfully with work rather than merely reducing a symptom count.

There is also a symbolic mechanism. When the institution acts on team or system-level risk, it signals that burnout is not interpreted as private failure. This may reduce shame and make help-seeking safer. The signal is credible only when practice matches rhetoric. A well-being dashboard that produces no workload change, or an alert that is followed by closer scrutiny rather than assistance, will not generate the same meaning. Organisational care is demonstrated through resource-bearing action, not through the existence of analytics.

8.2 The surveillance, stigma and resource-loss pathway

In the opposing pathway, employees experience the system as an extension of organisational surveillance. They may not know which data are collected, how scores are generated or who has access to them. This uncertainty can create anticipatory anxiety: ordinary behaviours become possible signals, and the employee cannot reliably distinguish support from assessment. Opacity therefore functions as a psychological demand because it reduces predictability and control.

Privacy threat activates a resource-loss process. Employees may feel that control over personal information, professional reputation and self-presentation has shifted to an unseen system. These are valued resources under Conservation of Resources theory. The threat may be sufficient to produce strain even before any adverse decision occurs. Employees may withhold information, avoid support services or engage in defensive impression management, thereby reducing the quality of both organisational data and genuine care.

Stigma can deepen the loss. A high-risk classification may imply weakness, unreliability or limited capacity, despite the fact that burnout is strongly shaped by work conditions. Managers may unintentionally reduce developmental opportunities or autonomy in the name of protection. The employee may also internalise the classification. In this way, a probabilistic prediction can become socially performative: it changes expectations and behaviour and may contribute to the adverse outcome it was intended to anticipate (Leicht-Deobald et al., 2019; Giermindl et al., 2022).

Perceived unfairness further undermines well-being. Employees are more likely to distrust a model when they cannot understand or challenge its output, when errors appear to affect some groups more than others or when the decision process lacks meaningful human consideration. Research on algorithmic decision perceptions links fairness and transparency to trust and emotional response, even though much of the direct evidence comes from recruiting and broader algorithmic-management contexts rather than burnout prediction specifically (Lee, 2018; Ochmann et al., 2024).

Finally, punitive intervention converts risk estimation into an additional job demand. Linking scores to performance review, promotion, scheduling penalties or redundancy can make the disclosure of strain dangerous. Employees then have an incentive to conceal difficulty, managers have an incentive to shift responsibility towards individuals and the institution receives less valid evidence about the conditions producing burnout. The result is a self-defeating system: a technology introduced in the language of well-being erodes the trust and psychological safety required for prevention.

8.3 Interpretive safeguards

Transparency is the first interpretive safeguard. Employees should receive an intelligible account of the model's purpose, data categories, level of analysis, limitations, decision pathway and permitted consequences. Transparency does not require publishing proprietary code, but it does require enough information for affected employees to understand the nature of the inference and the institutional use made of it. Silence or vague assurances create an informational asymmetry that invites fear and speculation.

Privacy is both a legal-ethical requirement and a psychological resource. Data minimisation, purpose limitation, short retention periods, role-based access and separation from disciplinary records reduce the likelihood that employees experience the model as an invasion. Aggregate and team-level reporting should be preferred where it can achieve the preventive objective. Individual-level outputs require a substantially stronger justification and a protected route of access.

Predictive fairness requires scrutiny of the full decision pipeline. Organisations should test error rates and calibration across relevant groups, examine whether proxies encode historical disadvantage and assess whether the consequences of a score are distributed fairly. A technically similar error can have different human significance depending on what follows. Fairness assessment must therefore include intervention, opportunity and remedy rather than stopping at statistical parity.

Explainability requires a reasoned account of why a particular risk signal arose and what weight the institution gives it. Employees should not be told that 'the algorithm decided'. A human reviewer must be able to identify plausible contributing factors, acknowledge uncertainty and distinguish correlation from cause. Where an output cannot be explained sufficiently for a sensitive intervention, the appropriate response may be to use it only at aggregate level or not to use it at all (Burrell, 2016; Bankins, 2021).

Participation and redress strengthen legitimacy. Employee representatives, occupational-health specialists, data-protection expertise and relevant ethics functions should participate in design and review. Individuals need a confidential route to ask questions, correct inaccurate data and challenge inappropriate use without retaliation. Because consent in employment is constrained by power asymmetry, formal consent alone cannot substitute for institutional safeguards and collective voice.

The final and most decisive safeguard is the quality of intervention. A prediction system should be judged by what it authorises and what it prohibits after a signal is produced. Supportive responses include workload review, resource provision, voluntary access to confidential services and system-level redesign. Prohibited responses should include automated adverse action, diagnostic labelling and use in promotion, dismissal or performance scoring. The intervention protocol is the practical test of whether the institution is providing care or merely expanding control.

Table 7. Interpretive safeguards and likely employee experience

Safeguard	Operational requirement	Protected employee resource	Likely failure when absent
Transparency	Explain purpose, data categories, limitations, access and consequences	Predictability and informational control	Rumour, anticipatory anxiety and distrust
Privacy	Minimise data, restrict access, limit retention and prevent secondary use	Personal boundary, dignity and safety	Intrusion, defensive behaviour and reduced help-seeking
Predictive fairness	Test subgroup calibration, error and consequences	Equal treatment and professional opportunity	Disparate stigma, missed support or discriminatory impact

Safeguard	Operational requirement	Protected employee resource	Likely failure when absent
Explainability	Provide a comprehensible reason and acknowledge uncertainty	Voice, understanding and contestability	Helplessness and deference to opaque authority
Employee participation	Include representatives in design, review and change control	Collective voice and legitimacy	Technological imposition and low trust
Human review	Contextual judgement with authority to reject or pause output	Agency and procedural justice	Automated or ceremonial decision-making
Redress	Correction, challenge and confidential complaint without retaliation	Control over record and reputation	Persistent error and unaccountable harm
Supportive intervention	Use signals to improve work and offer voluntary support	Time, energy, autonomy, support and meaning	Prediction becomes monitoring, stigma or sanction

Source: Author governance synthesis.

9. Analytical findings

Because this is a theoretical-analytical study, the findings presented here are propositions derived from the reviewed literature and integrated framework rather than statistical results generated from original field data. Their purpose is to organise the available evidence, identify plausible mechanisms and define the conditions that a later empirical study should test. This distinction is essential in a field where technical demonstrations are sometimes interpreted as proof of organisational benefit without evidence about employees' lived experience.

The first finding is that the relationship between burnout prediction and psychological well-being is dual, conditional and non-deterministic. The literature supports plausible benefits from early detection, evidence-informed prevention and targeted resource allocation, but it also supports substantial risks from surveillance, opacity, privacy loss, stigma and unfair classification. Neither outcome follows automatically from the presence of machine learning. Organisational purpose and post-prediction practice determine which pathway becomes dominant (Tambe et al., 2019; Giermindl et al., 2022; Bankins, 2021).

The second finding is that actual organisational use, not the model as an abstract artefact, is the appropriate unit of analysis. A risk model includes the data regime, reporting level, threshold, access rights, human review and intervention. Studies that measure only attitudes towards AI or intentions to adopt HR analytics do not fully capture the independent construct examined here. A later field study should verify that a model is operational and document what decisions or support processes its outputs actually influence.

The third finding is that predictive accuracy is necessary but insufficient. A model with acceptable discrimination can still damage well-being if it is opaque, intrusive or attached to punitive consequences. Conversely, a modest model used cautiously at aggregate level may contribute to organisational learning when its output initiates human inquiry rather than automatic judgement. The relevant evaluation criterion is therefore sociotechnical effectiveness: technical validity, ethical legitimacy, intervention quality and employee outcome must be considered together.

The fourth finding is that the level of analysis materially shapes risk and value. Team- or unit-level prediction is often better aligned with the organisational causes of burnout and may reduce the danger of individual stigma. Individual-level prediction can support timely assistance in some circumstances, but it requires stronger evidence, confidentiality and safeguards. Organisations should adopt the least intrusive level capable of achieving the preventive purpose.

The fifth finding is that supportive intervention is the principal mechanism through which prediction can plausibly enhance psychological well-being. The model does not generate a resource merely by identifying risk. Benefit arises when the institution changes workload, staffing, autonomy, recovery, managerial behaviour or access to support. This finding shifts the design question from 'How accurately can we predict burnout?' to 'Which organisational action follows, and does it address the conditions that create risk?'

The sixth finding is methodological. Robust assessment requires multiple sources. Employee data are needed to measure psychological well-being and lived experience; organisational records are needed to verify deployment, governance and intervention; and model documentation is needed to assess validity and fairness. A single self-report survey risks conflating awareness, trust and actual use. A purely technical audit, by contrast, cannot reveal psychological effects.

The final synthesis is that burnout prediction should be treated as a conditional organisational capability. It can become a channel for care when it protects resources, improves work and respects agency. It becomes a source of additional demand when it expands surveillance or redistributes risk to employees. This conditionality is not a weakness in the argument; it is the central theoretical result and the basis for the governance framework proposed later in the paper.

10. Discussion

The central research question asked how the organisational use of machine-learning burnout prediction models may affect employees' psychological well-being. The analysis indicates that the effect is mediated in practice by the organisational meaning and consequences of prediction, even though the present core model retains a direct relationship for analytical economy. A score becomes psychologically consequential when employees understand, anticipate or experience what the institution does with it. Prediction is therefore both an information process and a relational event.

The first sub-question concerned the nature of the models. They are probabilistic organisational instruments, not diagnostic authorities. Their output reflects the target, data, context and threshold selected by designers. This clarification restricts epistemic overreach and helps define an appropriate role: they may support inquiry and prevention, but should not determine clinical status or substitute for qualified professional judgement.

The second sub-question concerned model construction and use. Development requires valid targets, proportionate data, contextual validation and lifecycle monitoring, but deployment requires an additional architecture of access, human review and intervention. The organisational system is incomplete until the institution specifies what happens after a signal. This explains why technical studies alone cannot establish whether the technology improves employee well-being.

The third sub-question concerned benefits and risks. Benefits are most plausible when models identify structural demands, prompt early work redesign and improve access to resources. Risks are most pronounced when data collection is intrusive, scores are individually consequential, use is opaque or the system is connected to managerial sanction. These opposing pathways align with the JD-R and COR frameworks: supportive use reduces demands and protects resources, whereas surveillance-oriented use adds demands and threatens valued resources.

The fourth sub-question concerned transparency, privacy, fairness, explainability, participation and intervention. These features should not be treated as decorative ethical principles. They shape how employees interpret the system and whether they retain control, dignity and trust. Although the present article does not add them as separate variables to the core statistical model, future research can test them as moderators or mediators once the direct relationship has been established empirically.

The fifth sub-question concerned methodology. A multi-organisation, multi-source quantitative design is appropriate for an initial explanatory test, provided that operational deployment is independently verified. Structural equation modelling may test the latent constructs, while organisational audits document the model and its use. Longitudinal and mixed-method research will ultimately be necessary to identify temporal change, causal mechanisms and unintended effects.

Three broad positions can be distinguished in the literature. An instrumental position emphasises the efficiency and anticipatory value of people analytics. A critical position emphasises surveillance, objectification and asymmetric power. A conditional institutional position, supported by this study, argues that both sets of effects are possible and that governance determines their relative weight. This third position has greater explanatory reach because it accommodates variation across sectors, organisations and deployment designs rather than attributing a fixed moral character to the technology.

The theoretical contribution is to connect occupational-stress theory with the governance of predictive HR analytics. JD-R and COR explain how model-enabled actions alter demands and resources, while the care-surveillance lens explains the institutional meaning attached to those actions. The paper thereby relocates the evaluation of AI from model performance alone to the psychological and organisational consequences of use.

The practical contribution is a decision principle: no burnout prediction model should be approved without a credible supportive-action pathway and a documented prohibition on adverse automated use. Recent research on AI and worker well-being likewise suggests that effects vary with how technology changes work, autonomy and institutional conditions rather than with AI adoption in the abstract (Valtonen et al., 2025; Giuntella et al., 2025). The burden of proof should therefore rest on the organisation to demonstrate that the system improves conditions rather than merely improving prediction.

11. Proposed human-centred burnout prediction and support framework

The analysis supports a human-centred framework in which machine learning is a limited component of an organisational prevention system. The framework does not assume that prediction is necessary in every organisation or that more data invariably produce better care. It establishes a sequence of purpose definition, data governance, model assurance, human review, supportive action and outcome evaluation. Progression through the sequence depends on evidence gates, and use should stop when the institution cannot justify proportionality, validity or benefit.

11.1 Purpose, scope and use boundary

The institution should begin with a written problem statement. It must identify the work-related harm to be reduced, explain why existing non-algorithmic methods are insufficient and specify the intended level of intervention. A narrowly framed purpose might be to identify organisational units experiencing sustained workload imbalance so that staffing and recovery arrangements can be reviewed. A vague purpose such as 'improving employee performance through well-being analytics' is too elastic to protect against function creep.

The use boundary should distinguish prevention from diagnosis and care from employment evaluation. The model should not determine medical status, fitness for work, promotion, remuneration, disciplinary action or termination. Individual scores should not enter general personnel files. Where an individual alert is permitted, it should be accessible only to a protected occupational-health or well-being function and should initiate voluntary, confidential support rather than line-management judgement.

The scope should be proportionate to evidence. Team-level indicators, periodic analysis and limited data categories are preferable starting points. Expansion to more granular or frequent monitoring should require a new assessment rather than being treated as a routine technical upgrade. The institution should publish the permitted and prohibited uses, the responsible owner and the criteria for suspension or withdrawal.

Table 8. Proposed operationalisation of the core constructs

Construct	Operational definition	Indicative dimensions or indicators	Primary source	Boundary condition
Actual organisational use of burnout prediction	Extent to which an operational machine-learning system produces burnout-related risk outputs and informs organisational action	Active model; documented target; periodic outputs; identified users; integration into a defined intervention pathway; operating record	Organisational audit and model documentation	Does not include intentions, prototypes without use or general beliefs about AI
Employees' psychological well-being	Positive psychological functioning and balance within the work context	Meaning; competence; interpersonal fit; recognition; engagement; adaptation and psychological comfort	Confidential employee survey using a validated work-related scale	Not reducible to job satisfaction, absence of distress or burnout score
Deployment experience (descriptive/contextual)	How employees encounter and understand the operational system	Awareness; perceived purpose; care/surveillance; clarity; privacy concern; perceived fairness; intervention experience	Employee survey and qualitative follow-up	Used initially for interpretation, not to replace verified organisational use

Construct	Operational definition	Indicative dimensions or indicators	Primary source	Boundary condition
Organisational intervention	Action taken after a risk output	Work redesign; workload reduction; staffing; manager support; voluntary referral; monitoring; adverse employment consequence	Organisational audit and employee report	Must distinguish supportive from punitive or purely symbolic action
Technical assurance	Evidence that the model is valid for its declared context	Calibration; sensitivity/specificity; subgroup error; drift; threshold rationale; interpretability	Model assurance record	Technical quality is necessary but not sufficient for human benefit

Source: Author operational framework. Final items should be validated through expert review, pilot testing and confirmatory factor analysis.

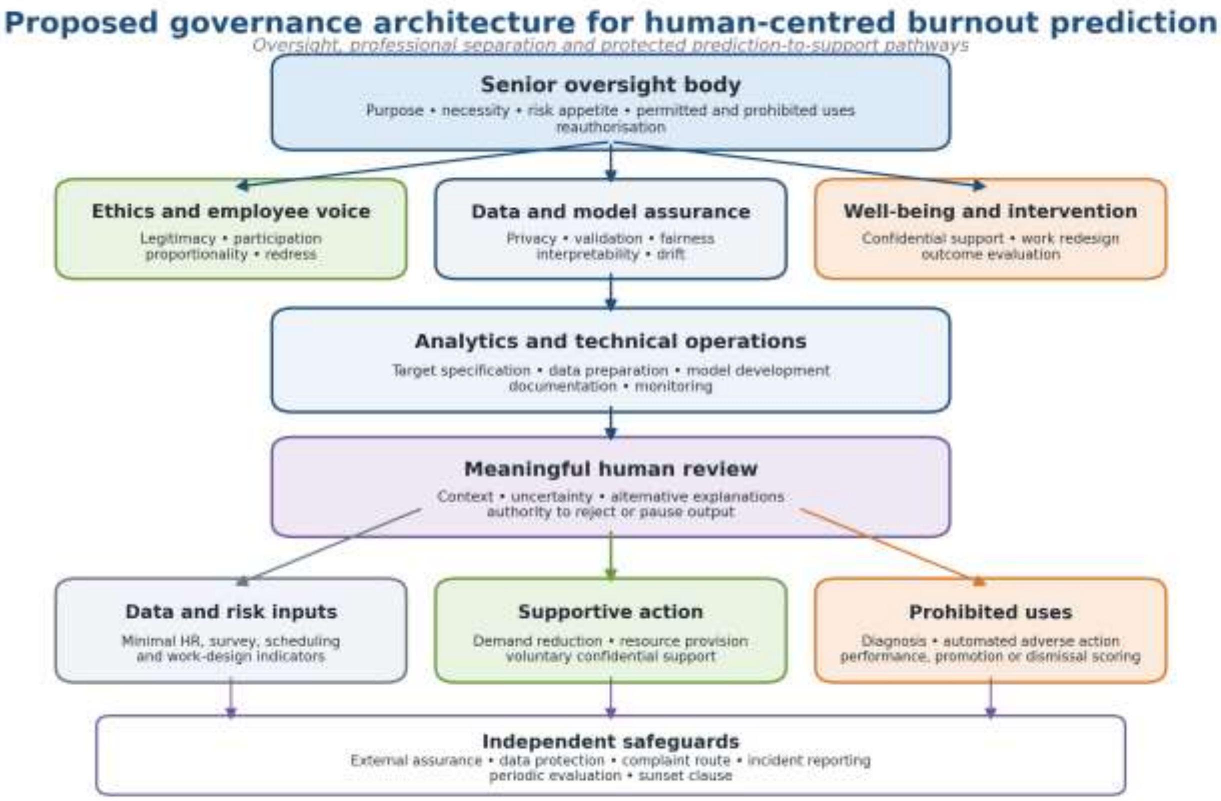
11.2 Data governance and model assurance

A data inventory should document source, purpose, lawful basis or organisational authority, sensitivity, retention, access and quality. Variables should be screened for necessity and proxy risk. Communication content, keystrokes, location or biometric data should not be used merely because they may improve prediction. De-identification and aggregation should be applied wherever the preventive objective does not require individual identification.

Model assurance should cover target validity, data quality, performance, calibration, subgroup error, interpretability, robustness and drift. Validation should be independent of the team responsible for deployment targets. A model card or equivalent record should state the intended population, excluded uses, training period, known limitations, threshold rationale and review schedule. Material changes to the target, data or intervention should trigger revalidation.

Governance roles should be separated. Senior oversight should approve purpose and risk appetite; an ethics and employee-voice function should test legitimacy and proportionality; data and model assurance should validate performance and fairness; and the well-being function should own the supportive intervention. The analytics team should not independently define the purpose, approve the model and judge its success. External review is appropriate for high-sensitivity deployments.

Figure 2. Proposed governance architecture for human-centred burnout prediction



Source: Author synthesis.

Table 9. Governance responsibilities and independence safeguards

Body or function	Core responsibility	Independence safeguard	Key accountability output
Senior oversight body	Approve purpose, risk appetite, deployment boundary and continuation	Includes non-operational members; records recusals and dissent	Annual purpose, necessity and proportionality statement
Ethics and employee-voice panel	Assess legitimacy, privacy, participation and distributional impact	Employee representation and direct escalation route	Pre-deployment impact assessment and change reviews
Data protection or privacy function	Review data inventory, authority, minimisation, access and retention	Independent reporting line and audit access	Data-protection assessment and access-exception report
Model assurance function	Validate target, performance, fairness, interpretability and drift	Separate from analytics delivery targets	Model card, validation opinion and monitoring dashboard

Body or function	Core responsibility	Independence safeguard	Key accountability output
Well-being and intervention function	Own supportive protocols and protect confidentiality	Separated from performance and disciplinary decision-making	Aggregate intervention and outcome report
Analytics team	Develop, document, monitor and technically maintain the model	Cannot approve its own production release or redefine purpose unilaterally	Technical documentation, version history and incident log
Human reviewers	Assess context, uncertainty and appropriate response	Authority to reject output; no volume-based incentive	Review record and rationale without diagnostic language
Independent audit or evaluation	Test controls, employee effects and compliance with prohibited uses	Unrestricted evidence access and periodic rotation	Assurance opinion and publicly summarised recommendations

Source: Author design informed by responsible-AI, HRM and algorithmic-governance literature.

11.3 Human review and supportive intervention

Human review must be meaningful rather than ceremonial. Reviewers should be trained to understand uncertainty, alternative explanations and the limits of feature importance. They should examine contextual evidence and should be authorised to reject or defer an output. A risk score should never be the sole basis for contacting an employee or changing work conditions. Review should also consider whether the appropriate response is at individual, team or organisational level.

Supportive intervention should follow a graduated protocol. System-level actions may include staffing review, job redesign, schedule stabilisation, recovery protection and manager development. Team-level actions may include workload dialogue and access to shared resources. Individual support, when justified, should be voluntary, confidential and separated from performance management. The intervention should not transfer responsibility for structurally produced burnout to the employee through generic resilience training alone.

Employees should have access to information, correction and redress. They should be able to learn whether an individual-level inference has been made, understand the main basis for it, correct inaccurate source data and challenge inappropriate consequences. A confidential complaint route should be independent of line management. Retaliation, adverse inference from non-participation and covert expansion of use should be expressly prohibited.

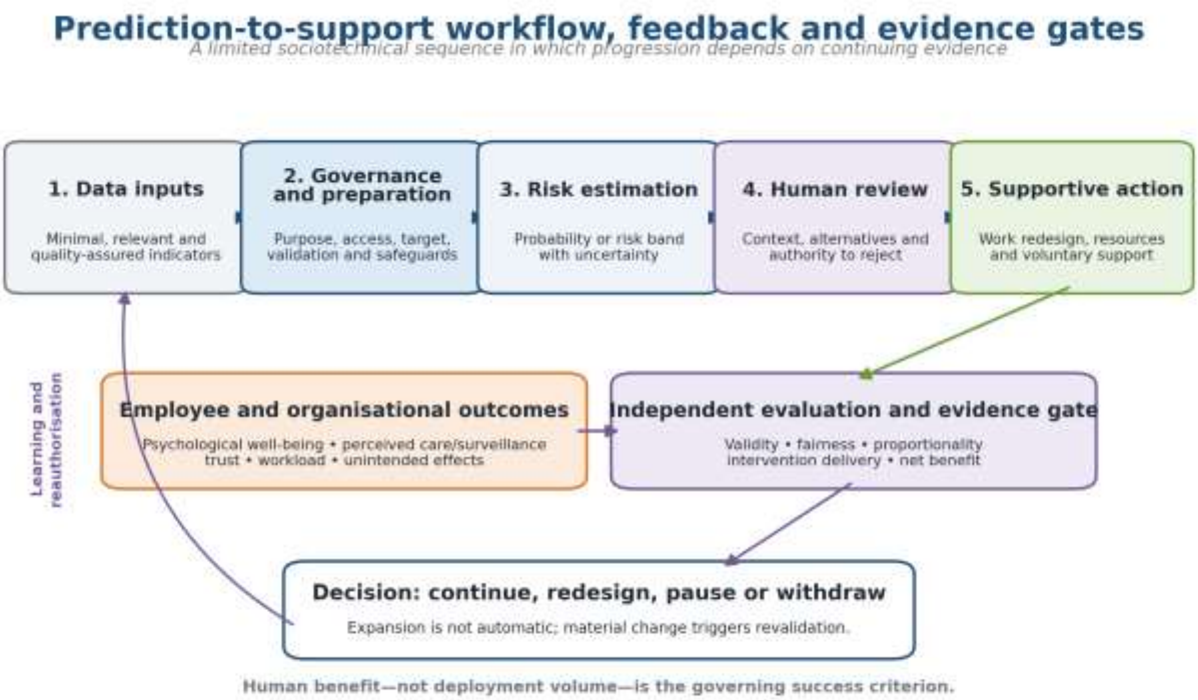
11.4 Performance, impact and accountability

Evaluation should integrate four domains: model performance, intervention delivery, employee well-being and governance integrity. Technical indicators may include calibration, sensitivity, false-positive rates and subgroup performance. Delivery indicators should show whether alerts were reviewed, whether supportive actions occurred and how quickly. Employee outcomes should include psychological well-being, perceived care, privacy concern, trust and unintended strain. Governance indicators should include access exceptions, complaints, prohibited-use incidents and remediation.

Outcome evaluation should distinguish correlation from contribution. An improvement in well-being after deployment cannot automatically be attributed to the model, particularly where other organisational changes occurred. Baselines, comparison units, phased introduction or matched designs can strengthen inference. Qualitative accounts are also important because a stable average may conceal that some employees experienced support while others experienced surveillance.

Public or workforce-facing reporting should be sufficiently transparent to sustain accountability without exposing individuals. It should include model purpose, use level, main data categories, validation status, intervention rates, aggregate outcomes, complaints and material changes. Unsuccessful or suspended deployments should be reported rather than removed from the organisational narrative. Trust depends on visible learning, not on success stories alone.

Figure 3. Prediction-to-support workflow, feedback and evidence gates



Source: Author synthesis.

Table 10. Performance, well-being and ethical assurance dashboard

Dimension	Core indicators	Interpretation	Frequency
Model validity	Calibration, sensitivity, specificity, false-positive/negative rates, drift	Tests whether risk estimates remain fit for the declared context	Monthly or quarterly; after material change

Dimension	Core indicators	Interpretation	Frequency
Distributional fairness	Subgroup calibration and error; differential intervention rates	Identifies unequal exposure to missed support or inappropriate classification	Quarterly and before expansion
Human review	Review rate, override rate, response time and documented rationale	Tests whether oversight is meaningful rather than ceremonial	Monthly
Supportive action	Work-design changes, resource provision, voluntary referrals and uptake	Shows whether prediction produces care-bearing intervention	Quarterly
Psychological well-being	Validated well-being score, perceived care, trust, privacy concern and psychological safety	Assesses intended benefit and unintended strain	Baseline; periodic; post-intervention
Data governance	Access exceptions, retention compliance, secondary-use requests and incidents	Tests adherence to purpose limitation and minimisation	Quarterly; immediate incident escalation
Employee voice and redress	Questions, corrections, complaints, resolution time and retaliation claims	Indicates contestability and institutional responsiveness	Quarterly with annual public summary
Organisational outcomes	Workload, absence, turnover, service quality and team-level risk patterns	Provides contextual outcomes without treating productivity as the sole purpose	Semi-annual or annual
Independent evaluation	Contribution analysis, comparison design, qualitative evidence and unintended effects	Distinguishes association from plausible organisational contribution	Pilot completion and every 2-3 years

Source: Author measurement framework.

11.5 Phased implementation and evidence gates

Implementation should proceed in stages. The first stage establishes necessity, legal and ethical authority and employee participation. The second develops governance, data controls and a non-production prototype. The third pilots a limited, preferably team-level use with supportive interventions. The fourth conducts independent evaluation against predefined technical, psychological and ethical criteria. Only then should controlled expansion be considered.

Each stage requires an evidence gate. A model should not proceed because development resources have already been spent. Failure to demonstrate valid targeting, proportionate data, fair performance, credible intervention or non-harm should lead to redesign or termination. Sunset clauses and periodic reauthorisation prevent temporary experimentation from becoming permanent surveillance by default.

Table 11. Phased implementation roadmap and evidence gates

Phase	Priority actions	Evidence gate before progression	Indicative decision
1. Necessity and authority	Define harm, compare alternatives, establish legal/ethical authority and involve employees	Documented need, legitimate purpose and proportionate scope	Proceed, redesign or stop before data expansion
2. Governance and prototype	Create oversight, data inventory, model card, prohibited uses and non-production prototype	Independent review confirms controls, target validity and testable benefit	Authorise limited pilot only
3. Limited supportive pilot	Use minimal data, preferably team-level outputs, trained review and predefined supportive actions	Technical performance, intervention delivery and no material privacy or fairness failure	Refine, pause or proceed to evaluation
4. Independent evaluation	Assess well-being, perceived care/surveillance, fairness, complaints and organisational outcomes	Evidence of plausible benefit without disproportionate harm	Expand, redesign or terminate
5. Controlled expansion	Add contexts or granularity only through new review and revalidation	Governance and support capacity grow at least as fast as analytical reach	Expand selectively, not automatically
6. Reauthorisation or sunset	Review necessity, alternatives, drift, function creep and accumulated effects	Continuing need and net benefit remain demonstrable	Renew for a fixed period or withdraw

Source: Author implementation framework.

12. Implications

The study has implications for theory, organisational practice, policy and the ethics of digital work. Its central message is that the value of burnout prediction lies neither in technological novelty nor in predictive accuracy alone. Value is produced when an institution can translate uncertain information into proportionate action that improves work while preserving rights and relationships.

12.1 Theoretical implications

First, the paper extends JD-R theory by treating an algorithmic system as a potential organisational resource or demand depending on its deployment. A supportive model can facilitate resource allocation and demand reduction; a surveillance-oriented model can add cognitive and emotional demands. The classification cannot be made at the level of technology alone.

Secondly, the paper extends Conservation of Resources theory to predictive employment systems. Privacy, informational control, professional reputation, autonomy and trust are resources that can be protected or threatened by algorithmic inference. The model's psychological effect therefore includes anticipatory resource loss, not merely the consequences of a completed managerial decision.

Thirdly, the care-surveillance lens clarifies why similar systems can generate divergent employee responses. Organisational care is not an intention declared by management; it is an interpretation grounded in transparency, voice and resource-bearing action. Surveillance is likewise experienced through uncertainty, asymmetry and consequential observation. This relational interpretation connects technical governance to psychological well-being.

12.2 Practical and policy implications

For HR leaders, the principal implication is to begin with work design rather than technology procurement. An organisation should ask which demands are producing strain, which resources are missing and which decisions could be improved by additional evidence. A predictive model is justified only where it offers a proportionate advantage over surveys, managerial dialogue, occupational-health assessment or direct workload monitoring.

For data and analytics teams, success criteria should extend beyond accuracy. Model documentation, subgroup testing, interpretability, drift monitoring and traceable human review are essential. Teams should be evaluated on data quality, safe use and organisational learning, not merely on the number of models moved into production.

For managers and occupational-health professionals, risk signals should prompt inquiry rather than judgement. Training should address uncertainty, confidentiality, supportive conversation and system-level intervention. Managers should not receive individual scores unless a compelling and protected use case exists, and they should never be asked to infer medical meaning from an organisational prediction.

For policymakers, regulators and professional bodies, burnout prediction sits at the intersection of employment, data protection, occupational health and AI governance. Guidance should clarify the distinction between prevention and diagnosis, set expectations for high-impact HR systems, protect employee voice and redress and require impact assessment before and after deployment. Voluntary ethics statements are unlikely to be sufficient where incentives favour monitoring and control.

12.3 Social and ethical implications

Human-centred use may help normalise organisational responsibility for burnout and direct attention towards preventable conditions. It may widen access to support for employees who would otherwise remain unseen, particularly in large or complex organisations. These benefits, however, depend on avoiding a model in which employees must surrender extensive privacy to qualify for care.

Unequal exposure requires particular attention. Workers in highly monitored, precarious or operational roles may generate more behavioural data and may have less power to contest its use than senior or professional employees. A system can therefore widen existing inequalities even when aggregate performance appears acceptable. Distributional analysis and collective participation are necessary components of ethical evaluation.

The broader social question concerns the boundary of legitimate organisational knowledge. Employment does not give an institution unlimited authority to infer psychological states from digital behaviour. Responsible innovation requires restraint: some predictions should not be made, some data should not be collected and some efficiencies should be rejected when they undermine dignity, agency or trust.

13. Limitations and future research

The study is theoretical and relies on a critical synthesis of published literature. It does not provide original empirical estimates of the relationship between burnout prediction and psychological well-being. The analytical propositions should therefore be tested rather than presented as settled causal findings. The direct evidence base on employees' experience of operational burnout-prediction systems remains limited. The literature spans occupational psychology, HR analytics, algorithmic management, AI ethics and decision perceptions. Some of the strongest evidence on fairness, transparency and trust comes from recruitment or platform-work settings rather than burnout prediction specifically. The transfer is theoretically justified by common mechanisms of probabilistic classification and power, but contextual differences may affect the magnitude and form of the relationship.

Future research should first conduct multi-organisation studies in settings where models are demonstrably operational. A sample of approximately 450-600 employees across six to eight organisations, as proposed in the methodology, would permit initial confirmatory measurement and direct-effect modelling, subject to a formal power analysis based on the final model and anticipated effect. Organisational verification should document target, level, data, output, intervention and duration of deployment.

Longitudinal research is needed because effects may change over time. Initial novelty or suspicion may diminish, while function creep or model drift may emerge later. Phased or stepped-wedge introductions, natural experiments and matched comparison designs could strengthen causal inference. Mixed methods would help explain why a similar score produces care in one unit and surveillance in another.

Further studies should examine cultural and institutional variation, employee participation, union or representative governance, occupational differences and alternative levels of analysis. Once the direct relationship is established, transparency, perceived organisational care, privacy concern, procedural justice, psychological safety and trust can be tested as mediators or moderators. Research should also compare predictive systems with less intrusive alternatives to determine whether machine learning adds genuine preventive value.

14. Conclusion

This study examined how the organisational deployment of machine-learning burnout prediction models may affect employees' psychological well-being. It argued that the relationship is neither inherently beneficial nor inherently harmful. Prediction becomes consequential through institutional design, data practice, human interpretation and the intervention that follows.

Used within a preventive and supportive architecture, a model may help identify harmful patterns, direct resources and prompt redesign before demands produce deeper exhaustion. In this pathway, the technology contributes indirectly by enabling the organisation to reduce demands and protect resources. Its legitimacy depends on privacy, proportionality, explanation, employee voice and the separation of support from adverse employment decisions.

Used within an opaque, individualising or punitive architecture, the same capability may become an additional source of strain. Surveillance, uncertainty, stigma and loss of informational control can erode trust and psychological safety. A technically accurate model can therefore fail the human purpose invoked to justify it.

The paper's principal contribution is to shift evaluation from the narrow question 'Does the model predict burnout?' to the more consequential question 'How does the organisation use prediction, and what does that use do to employees' experience of work?' The integrated JD-R and COR framework explains the demand-resource mechanisms, while the care-surveillance lens explains the relational meaning of deployment.

The appropriate institutional response is disciplined conditionality. Organisations should use the least intrusive method capable of achieving a legitimate preventive purpose, require meaningful human review, link any signal to supportive action, measure psychological and ethical outcomes and stop systems that cannot demonstrate benefit without disproportionate harm. The future of people analytics should be judged not only by what organisations can infer, but by whether they can act on inference without losing sight of the person.

References

1. Angrave, D., Charlwood, I., Kirkpatrick, I., Lawrence, M. and Stuart, M. (2016), "HR and analytics: why HR is set to fail the big data challenge", Human Resource Management Journal, Vol. 26 No. 1, pp. 1-11, doi: 10.1111/1748-8583.12090.
2. Bakker, A.B. and Demerouti, E. (2007), "The Job Demands-Resources model: state of the art", Journal of Managerial Psychology, Vol. 22 No. 3, pp. 309-328, doi: 10.1108/02683940710733115.
3. Bakker, A.B. and Demerouti, E. (2017), "Job Demands-Resources theory: taking stock and looking forward", Journal of Occupational Health Psychology, Vol. 22 No. 3, pp. 273-285, doi: 10.1037/ocp0000056.
4. Bakker, A.B., Demerouti, E. and Sanz-Vergel, A.I. (2014), "Burnout and work engagement: the JD-R approach", Annual Review of Organizational Psychology and Organizational Behavior, Vol. 1 No. 1, pp. 389-411, doi: 10.1146/annurev-orgpsych-031413-091235.
5. Bankins, S. (2021), "The ethical use of artificial intelligence in human resource management: a decision-making framework", Ethics and Information Technology, Vol. 23 No. 4, pp. 841-854, doi: 10.1007/s10676-021-09619-6.
6. Bar-Gil, O., Ron, T. and Czerniak, O. (2024), "AI for the people? Embedding AI ethics in HR and people analytics projects", Technology in Society, Vol. 77, 102527, doi: 10.1016/j.techsoc.2024.102527.
7. Barocas, S. and Selbst, A.D. (2016), "Big data's disparate impact", California Law Review, Vol. 104, pp. 671-732.
8. Bujold, A., Roberge-Maltais, I., Parent-Rocheleau, X., Boasen, J., Sénécal, S. and Léger, P.-M. (2024), "Responsible artificial intelligence in human resources management: a review of the empirical literature", AI and Ethics, Vol. 4 No. 4, pp. 1185-1200, doi: 10.1007/s43681-023-00325-1.
9. Burrell, J. (2016), "How the machine 'thinks': understanding opacity in machine learning algorithms", Big Data & Society, Vol. 3 No. 1, pp. 1-12, doi: 10.1177/2053951715622512.
10. Dagenais-Desmarais, V. and Savoie, A. (2012), "What is psychological well-being, really? A grassroots approach from the organizational sciences", Journal of Happiness Studies, Vol. 13 No. 4, pp. 659-684, doi: 10.1007/s10902-011-9285-3.
11. Demerouti, E., Bakker, A.B., Nachreiner, F. and Schaufeli, W.B. (2001), "The Job Demands-Resources model of burnout", Journal of Applied Psychology, Vol. 86 No. 3, pp. 499-512, doi: 10.1037/0021-9010.86.3.499.
12. Giermindl, L.M., Strich, L.M., Christ, O., Leicht-Deobald, U. and Redzepi, A. (2022), "The dark sides of people analytics: reviewing the perils for organisations and employees", European Journal of Information Systems, Vol. 31 No. 3, pp. 410-435, doi: 10.1080/0960085X.2021.1927213.

13. Giuntella, O., König, J. and Stella, L. (2025), "Artificial intelligence and the wellbeing of workers", Scientific Reports, Vol. 15 No. 1, 20087, doi: 10.1038/s41598-025-98241-3.
14. Grant, A.M., Christianson, M.K. and Price, R.H. (2007), "Happiness, health, or relationships? Managerial practices and employee well-being tradeoffs", Academy of Management Perspectives, Vol. 21 No. 3, pp. 51-63, doi: 10.5465/AMP.2007.26421238.
15. Grządzielewska, M. (2021), "Using machine learning in burnout prediction: a survey", Child and Adolescent Social Work Journal, Vol. 38 No. 2, pp. 175-180, doi: 10.1007/s10560-020-00733-w.
16. Guest, D.E. (2017), "Human resource management and employee well-being: towards a new analytic framework", Human Resource Management Journal, Vol. 27 No. 1, pp. 22-38, doi: 10.1111/1748-8583.12139.
17. Hobfoll, S.E. (1989), "Conservation of resources: a new attempt at conceptualizing stress", American Psychologist, Vol. 44 No. 3, pp. 513-524.
18. Hülter, S.M., Ertel, C. and Heidemann, A. (2024), "Exploring the individual adoption of human resource analytics: behavioural beliefs and the role of machine learning characteristics", Technological Forecasting and Social Change, Vol. 208, 123709, doi: 10.1016/j.techfore.2024.123709.
19. Johnson, T. and Shamroukh, S. (2024), "Predictive modeling of burnout based on organizational culture perceptions among health systems employees: a comparative study using correlation, decision tree, and Bayesian analyses", Scientific Reports, Vol. 14 No. 1, 6083, doi: 10.1038/s41598-024-56771-2.
20. Kinowska, H. and Sienkiewicz, Ł.J. (2023), "Influence of algorithmic management practices on workplace well-being: evidence from European organisations", Information Technology & People, Vol. 36 No. 8, pp. 21-42, doi: 10.1108/ITP-02-2022-0079.
21. Konuk, H., Ataman, G. and Kambur, E. (2023), "The effect of digitalized workplace on employees' psychological well-being: digital Taylorism approach", Technology in Society, Vol. 74, 102302, doi: 10.1016/j.techsoc.2023.102302.
22. Lee, M.K. (2018), "Understanding perception of algorithmic decisions: fairness, trust, and emotion in response to algorithmic management", Big Data & Society, Vol. 5 No. 1, pp. 1-16, doi: 10.1177/2053951718756684.
23. Leicht-Deobald, U., Busch, T., Schank, C., Weibel, A., Schafheitle, S., Wildhaber, I. and Kasper, G. (2019), "The challenges of algorithm-based HR decision-making for personal integrity", Journal of Business Ethics, Vol. 160 No. 2, pp. 377-392, doi: 10.1007/s10551-019-04204-w.
24. Marler, J.H. and Boudreau, J.W. (2017), "An evidence-based review of HR analytics", The International Journal of Human Resource Management, Vol. 28 No. 1, pp. 3-26, doi: 10.1080/09585192.2016.1244699.
25. Maslach, C. and Leiter, M.P. (2016), "Understanding the burnout experience: recent research and its implications for psychiatry", World Psychiatry, Vol. 15 No. 2, pp. 103-111, doi: 10.1002/wps.20311.
26. Maslach, C., Schaufeli, W.B. and Leiter, M.P. (2001), "Job burnout", Annual Review of Psychology, Vol. 52, pp. 397-422, doi: 10.1146/annurev.psych.52.1.397.
27. Meijerink, J. and Bondarouk, T. (2023), "The duality of algorithmic management: toward a research agenda on HRM algorithms, autonomy and value creation", Human Resource Management Review, Vol. 33 No. 1, 100876, doi: 10.1016/j.hrmr.2021.100876.
28. Meijerink, J., Boons, M., Keegan, A. and Marler, J. (2021), "Algorithmic human resource management: synthesizing developments and cross-disciplinary insights on digital HRM", The

International Journal of Human Resource Management, Vol. 32 No. 12, pp. 2545-2562, doi: 10.1080/09585192.2021.1925326.
29. Minbaeva, D.B. (2018), "Building credible human capital analytics for organizational competitive advantage", Human Resource Management, Vol. 57 No. 3, pp. 701-713, doi: 10.1002/hrm.21848.
30. Mittelstadt, B.D., Allo, P., Taddeo, M., Wachter, S. and Floridi, L. (2016), "The ethics of algorithms: mapping the debate", Big Data & Society, Vol. 3 No. 2, doi: 10.1177/2053951716679679.
31. Ochmann, J., Michels, L., Tiefenbeck, V., Maier, C. and Laumer, S. (2024), "Perceived algorithmic fairness: an empirical study of transparency and anthropomorphism in algorithmic recruiting", Information Systems Journal, Vol. 34 No. 2, pp. 384-414, doi: 10.1111/isj.12482.
32. Ryff, C.D. (1989), "Happiness is everything, or is it? Explorations on the meaning of psychological well-being", Journal of Personality and Social Psychology, Vol. 57 No. 6, pp. 1069-1081, doi: 10.1037/0022-3514.57.6.1069.
33. Ryff, C.D. (2014), "Psychological well-being revisited: advances in the science and practice of eudaimonia", Psychotherapy and Psychosomatics, Vol. 83 No. 1, pp. 10-28, doi: 10.1159/000353263.
34. Tambe, P., Cappelli, P. and Yakubovich, V. (2019), "Artificial intelligence in human resources management: challenges and a path forward", California Management Review, Vol. 61 No. 4, pp. 15-42, doi: 10.1177/0008125619867910.
35. Valtonen, A., Saunila, M., Ukko, J., Treves, L. and Ritala, P. (2025), "AI and employee wellbeing in the workplace: an empirical study", Journal of Business Research, Vol. 199, 115584, doi: 10.1016/j.jbusres.2025.115584.
36. van de Voorde, K., Paauwe, J. and van Veldhoven, M. (2012), "Employee well-being and the HRM-organizational performance relationship: a review of quantitative studies", International Journal of Management Reviews, Vol. 14 No. 4, pp. 391-407, doi: 10.1111/j.1468-2370.2011.00322.x.
37. Van Zyl-Cillié, M.M., Bührmann, J.H., Blignaut, A.J., Demirtas, D. and Coetzee, S.K. (2024), "A machine learning model to predict the risk factors causing feelings of burnout and emotional exhaustion amongst nursing staff in South Africa", BMC Health Services Research, Vol. 24 No. 1, 1665, doi: 10.1186/s12913-024-12184-5.
38. Zayid, H., Alzubi, H., Berberoğlu, A. and Khadem, K. (2024), "How do algorithmic management practices affect workforce well-being? A parallel moderated mediation model", Behavioral Sciences, Vol. 14 No. 12, 1123, doi: 10.3390/bs14121123.
39. Zhang, L. and Yencha, C. (2022), "Examining perceptions towards hiring algorithms", Technology in Society, Vol. 68, 101848, doi: 10.1016/j.techsoc.2021.101848.