Occupational Physicians Should Not Rely Solely on Artificial Intelligence in Fitness-for-Work Decisions

DOI: 10.52214/vib.v12i.14828

Noer Triyanto Rusli*

Keywords: Occupational Physicians, Artificial Intelligence, Fitness for Work, Autonomy, Dignity

Imagine a crane operator arriving for a routine shift and being told that an artificial intelligence system has classified him as "unfit for work" after combining recent medical notes, fatigue-risk indicators, wearable sleep data, and a pattern of short absences. The decision appears efficient from an administrative perspective, but its consequences are ethically significant. A fitness-for-work decision can affect employment, income, professional dignity, privacy, reputation among colleagues, access to reasonable accommodation, and the safety of coworkers and the public. Because such a decision sits at the intersection of health, livelihood, fairness, and public safety, it cannot be treated as a purely technical output. The core ethical question is therefore: should an occupational physician ever rely solely on AI to make a fitness-for-work decision?

AI can play a valuable role in occupational medicine by helping physicians identify patterns, organize complex information, and flag potential risks that might otherwise be missed. However, AI must remain a tool rather than the final decision-maker. These decisions are ethically and clinically complex because they require individualized assessment, knowledge of actual job demands, consideration of reasonable accommodations, protection of medical confidentiality, and professional accountability. The issue is not whether occupational physicians should use AI at all, but whether they may ethically outsource professional judgment to AI. Responsible AI use can support occupational physicians, but it cannot replace the human judgment required when a person's livelihood, dignity, privacy, and workplace safety are at stake.

A fitness-for-work assessment asks whether a worker can perform the essential functions of a particular job without creating an unacceptable risk to the worker, coworkers, patients, clients, or the public. A fitness-for-work assessment is an evaluation of whether a worker can perform tasks without risk to self or others, which requires knowledge of both health conditions and working conditions.[1] Such assessments may arise before employment, after illness or in-

[1] Consol Serra et al., Criteria and methods used for the assessment of fitness for work: a systematic review, Occupational and Environmental Medicine, https://pmc.ncbi.nlm.nih.gov/articles/PMC2092557/

* Noer Triyanto Rusli, MD, Occupational Medicine Specialist, Lecturer at Universitas Indo Global Mandiri

jury, during return-to-work evaluations, after changes in health status, or in hazardous and safety-sensitive jobs.[2] Fitness for work is a dynamic and context-dependent judgment that depends both on the worker's condition and on the work environment. A person may be unfit for one job, fit for another, or fit for the same job with restrictions, modified duties, assistive measures, or reasonable accommodations. This is why a simple algorithmic label of "fit" or "unfit" is ethically inadequate.

AI is attractive in occupational medicine because it can help teams manage complex information more efficiently and consistently. Properly designed systems assist with risk prediction, occupational health surveillance, documentation, pattern recognition, workload reduction, triage, ergonomic monitoring, exposure monitoring, smart personal protective equipment, and broader decision support. Potential uses of AI in health risk assessment include return-to-work prediction, injury severity estimation, injury management, workplace hazard detection, wearable monitoring, environmental sensing, and smart personal protective equipment.[3] Used responsibly, these tools may improve speed, consistency, and preventive action, especially where occupational health teams must evaluate large amounts of data under time pressure. AI or its potential value should not be rejected. AI can strengthen occupational medicine when it is transparent, validated, clinically supervised, and integrated into professional judgment. Yet these advantages do not justify sole reliance on AI in decisions that affect rights, safety, confidentiality, and professional duties.

Fitness-for-work decisions require contextual judgment that AI may not capture because the question is not simply whether a worker has a diagnosis, an abnormal data point, or an elevated risk score. The relevant question is whether this worker can safely perform this work in this workplace at this time. An occupational physician must understand the worker's tasks, physical and psychological job demands, workplace hazards, shift patterns, safety-sensitive responsibilities, available accommodations, clinical condition, and how they are integrated. AI systems often rely on available data, but available data are not always morally relevant. Important information may be missing, outdated, inaccurate, biased, or difficult to quantify, such as supportive supervision, recent symptom improvement, a modified task plan, or the practical feasibility of a gradual return to work. An algorithm may identify increased risk but may not understand whether a modified schedule, ergonomic adjustment, assistive device, or temporary restriction would allow safe work. Human clinical judgment is necessary to interpret data in context and distinguish genuine safety concerns from manageable limitations.

Respect for autonomy weighs against AI-only decision-making. Workers should not be treated as passive objects of algorithmic classification, especially when the outcome affects their livelihood and professional identity. They should understand how information about them is being used, have a meaningful opportunity to correct inaccurate or incomplete data, and be able to discuss their capacity, limitations, concerns, and preferences with a qualified professional. Sole reliance on AI weakens autonomy because the worker may not know what data was used, how the model reached its conclusion, or how to challenge the result. This concern is intensified in occupational medicine because workers may feel pressured to accept health monitoring, wearable tracking, or algorithmic evaluation when refusal

[2] Consol Serra et al., Criteria and methods used for the assessment of fitness for work: a systematic review, Occupational and Environmental Medicine, https://pmc.ncbi.nlm.nih.gov/articles/PMC2092557/. Neeladri Misra, Carlos Castro-Vasquez, and Daniel Keyes, Fitness for Duty and Return to Work, StatPearls, https://www.ncbi.nlm.nih.gov/books/NBK610688/.

[3] Mohamed El-Helaly, Artificial Intelligence and Occupational Health and Safety, Benefits and Drawbacks, La Medicina del Lavoro, https://pmc.ncbi.nlm.nih.gov/articles/PMC11181216/. Zaira S. Chaudhry and Avishek Choudhury, Clinical Applications of Artificial Intelligence in Occupational Health: A Systematic Literature Review, Journal of Occupational and Environmental Medicine, https://journals.lww.com/joem/abstract/2024/12000/clinical_applications_of_artificial_intelligence.1.aspx.

could be perceived as threatening employment. Ethical fitness-for-work practice requires meaningful communication and informed participation, not merely automated output.

Confidentiality and privacy are other reasons for caution. Occupational health data are sensitive because they can influence employment opportunities, workplace relationships, stigma, insurance, and managerial decisions. Occupational health ethics has long recognized that occupational health professionals occupy a position of trust that requires independence, confidentiality, and respect for workers' dignity.[4] In fitness-for-work practice, the employer needs information about functional capacity, restrictions, and accommodations, not unnecessary diagnostic detail. AI can intensify privacy risks because it can combine medical records, wearable data, absenteeism patterns, productivity indicators, location data, and workplace risk signals into predictions that appear objective. This creates the danger of function creep: data collected for health protection or safety may later be used for productivity monitoring, workforce selection, disciplinary action, or exclusion of workers perceived as costly or risky. Confidentiality obligations cannot be delegated to an algorithm, vendor, or dashboard. The occupational physician must remain responsible for deciding what information is relevant, necessary, and ethically permissible to disclose.

The principle of non-maleficence is equally important because erroneous fitness-for-work decisions can harm both workers and others. A false "fit" decision may expose a worker, coworkers, patients, clients, or the public to preventable harm. A false "unfit" decision may unjustly exclude a worker from employment, worsen financial insecurity, create stigma, or delay recovery and return to work. AI errors are ethically significant because model outputs may be affected by poor-quality data, model drift, lack of external validation, inappropriate deployment outside the population for which the system was developed, or opaque reasoning. AI has potential clinical utility but explainable models rigorously validated in real-world settings are still needed.[5] If generative AI is used, another risk is that the system may produce confident but inaccurate explanations. Occupational physicians must independently verify AI outputs before using them in consequential decisions.

Justice requires particular attention because fitness-for-work decisions must avoid unfair discrimination against workers with disabilities, chronic illnesses, mental health conditions, older workers, pregnant workers, and other vulnerable groups. AI systems can reproduce or amplify bias if trained on historical data shaped by inequitable employment practices, unequal access to healthcare, or biased documentation. The healthcare literature provides a cautionary example: a widely used population-health algorithm exhibited racial bias because it predicted healthcare costs rather than illness, and unequal access to care made cost a misleading proxy for need.[6] In occupational medicine, similar proxy problems may arise when models rely on absence patterns, healthcare utilization, medication use, productivity metrics, or wearable data that reflect social conditions rather than true inability to work safely. Responsible AI use requires bias audits, external validation, monitoring across worker groups, and individualized review. Neutrality is not the same as fairness.

[4] International Commission on Occupational Health, International Code of Ethics for Occupational Health Professionals, http://www.icohweb.org/site/code-of-ethics.asp

[5] Zaira S. Chaudhry and Avishek Choudhury, Clinical Applications of Artificial Intelligence in Occupational Health: A Systematic Literature Review, Journal of Occupational and Environmental Medicine, https://journals.lww.com/joem/abstract/2024/12000/clinical_applications_of_artificial_intelligence.1.aspx.

[6] Ziad Obermeyer et al., Dissecting racial bias in an algorithm used to manage the health of populations, Science, https://pubmed.ncbi.nlm.nih.gov/31649194/

Accountability is also central. Occupational physicians cannot avoid responsibility by pointing to an AI recommendation. Fitness-for-work decisions require reasons that can be explained to the worker, employer, regulator, tribunal, or court if necessary. The physician must be able to justify why a worker was judged fit, fit with restrictions, conditionally fit, or unfit. Sole reliance on AI creates accountability gaps because responsibility may become blurred among the physician, employer, software developer, data provider, occupational health service, insurer, and vendor. Healthcare professionals have already identified accountability, transparency, and bias as unresolved ethical concerns in AI-assisted clinical decision-making, particularly where "black box" models make it difficult to understand the basis for recommendations.[7] If no one can explain the decision, the worker has no meaningful path to challenge it. Professional accountability requires the physician to remain the responsible decision-maker, not merely a human signature on an algorithmic conclusion.

The strongest counterargument is that AI may be more objective and efficient than unaided human judgment. Supporters may argue that AI can reduce inconsistency, implicit bias, fatigue, and variation between physicians. They may also argue that AI can process larger datasets, identify early risk signals, reduce administrative burden, and improve workplace safety more efficiently than a physician working alone. These points should be taken seriously. Human judgment is imperfect, occupational medicine often operates under time pressure, and the evidence base for some fitness-for-work decisions is limited. There is no single validated method that provides unequivocal answers for all professions and circumstances.[8] These concerns strengthen the case for AI-assisted decision-making, but they do not justify AI-only decision-making.

Efficiency and apparent objectivity are not enough. AI may reduce some human errors while introducing new ones, including hidden bias, privacy intrusion, overreliance, lack of explainability, and misuse outside validated settings. A model's output may look precise even when it is uncertain, incomplete, or contextually inappropriate. The ethical answer is not to choose between humans and machines, but to define the proper relationship between them. AI should function as a second reader, warning system, triage support, documentation aid, ergonomic monitor, exposure signal, or risk-analysis tool. It should not become the sole authority over a person's capacity to work. The physician must integrate AI output with clinical assessment, job analysis, worker dialogue, legal standards, confidentiality duties, and accommodation options.

A practical framework for responsible use should be physician-led and rights-sensitive. AI should be used only when the system is validated for the relevant occupational context, transparent enough for clinical interpretation, regularly audited for bias and performance, and governed by strict privacy rules. The World Health Organization's guidance on AI for health emphasizes that ethics and human rights must be central to the design, deployment, and use of AI, and that governance should hold stakeholders accountable and responsive to the people affected by these technologies.[9] In fitness-for-work decisions, occupational physicians should document how AI was used, what data informed the recommendation, what limitations were recognized, whether the worker had a chance to correct information, what

[7] Saoudi C. E. Nouis, Victoria Uren, and Srushti Jariwala, Evaluating accountability, transparency, and bias in AI-assisted healthcare decision-making, BMC Medical Ethics, https://pmc.ncbi.nlm.nih.gov/articles/PMC12235780/

[8] Consol Serra et al., Criteria and methods used for the assessment of fitness for work: a systematic review, Occupational and Environmental Medicine, https://pmc.ncbi.nlm.nih.gov/articles/PMC2092557/

[9] World Health Organization, Ethics and governance of artificial intelligence for health, https://www.who.int/publications/i/item/9789240029200

accommodations were considered, and why the final decision was made. Documentation is not mere bureaucracy; it is part of ethical accountability.

Ethical requirement	Practical safeguard in AI-assisted fitness-for-work practice
Individualized assessment	Combine AI output with clinical examination, job analysis, and worker dialogue.
Confidentiality	Limit employer disclosure to work capacity, restrictions, and accommodations, rather than unnecessary medical detail.
Non-maleficence	Verify AI outputs and avoid using systems outside validated settings.
Justice	Audit performance across worker groups and review proxy variables for discriminatory effects.
Accountability	Record the physician's independent reasoning and the role AI played in the final decision.

Procedural fairness should be built into AI-informed occupational health practice. Workers affected by such decisions should receive understandable explanations of the decision, including the role AI played and the non-AI clinical reasons supporting the final conclusion. They should have a reasonable opportunity to provide additional information, correct inaccurate data, request review, or seek a second opinion. Guidance on healthcare algorithms emphasizes transparency, explainability, stakeholder engagement, equity, and ongoing monitoring; it also states that health decisions should not be made automatically by an algorithm and should involve human input.[10] In occupational medicine, these procedures protect both worker rights and workplace safety. They can improve trust, reduce the risk of hidden exclusion, and help ensure that AI serves prevention rather than surveillance or discrimination.

AI may help occupational physicians detect risks earlier, organize evidence more consistently, and support safer work. It may improve occupational health when used as a carefully governed form of decision support. But a fitness-for-work decision is not merely a prediction problem. It is a professional judgment about a particular person, a job, a set of risks, and a set of possible accommodations. Because the decision affects dignity, income, privacy, legal rights, and public safety, it requires a responsible professional who can listen, interpret, explain, and be held accountable. In fitness-for-work decisions, AI may inform the physician's judgment, but it must not become the judgment itself.

[10] Marshall H. Chin et al., Guiding Principles to Address the Impact of Algorithm Bias on Racial and Ethnic Disparities in Health and Health Care, JAMA Network Open, https://jamanetwork.com/journals/jamanetworkopen/fullarticle/2812958