SAFEGUARDS FOR A.I. IN HEALTH CARE

A.B. 2575 Assemblymember Liz Ortega

SUMMARY

This bill establishes commonsense safeguards for patients and health care workers when Artificial Intelligence (A.I.) is used in health care. These protections set clear rules for informing patients and workers when A.I. is being used, affirm clinicians' ability to override A.I.-driven decisions and to exercise professional judgment without retaliation, and ensure that developers and deployers of A.I. in health care settings can be held responsible for harm.

BACKGROUND

Hospitals, clinics, and other health care entities are incorporating A.I. into electronic health records, staffing systems, clinical decision-support tools, remote monitoring platforms, and administrative decision-making. These tools can generate patient acuity scores, treatment recommendations, insurance determinations, discharge plans, and nurse workload assignments.

As adoption of A.I. accelerates, nurses and other clinicians encounter A.I.-generated recommendations more frequently in routine care, while employers use automated systems to guide staffing, evaluate performance, and manage workflows. These changes are reshaping patient care and working conditions across the health care system.

Research shows that many health care A.I. tools can produce inaccurate recommendations, reflect bias in training data, or perform differently in real-world settings than in testing environments.[1] Studies have documented cases in which widely used health care algorithms produced racially discriminatory care recommendations and contributed to unequal access to services for Black patients.[2] Surveys of nurses and clinicians indicate growing concern that A.I.-driven tools do not always reflect clinical reality and can add pressure to follow automated outputs even when they conflict with professional judgment.[3]

Health care entities often provide patients and health care workers with little to no information about when they use A.I. systems, how they function, what data they rely on, or what risks and limitations they carry. At the same time, A.I. developers are actively seeking legal protections that would shield them from liability when their systems contribute to harmful outcomes.[4]

PROBLEM

As A.I. tools increasingly shape clinical decisions and working conditions, the absence of transparency, professional judgement override protections, and accountability for harm creates real risks for patients and undermines professional standards in care.

Lack of transparency. Health care entities fail to inform patients or health care workers when A.I. tools influence care decisions or workplace expectations. For example, A.I. systems may generate patient acuity scores, treatment prompts, or discharge recommendations without disclosing how those outputs were produced, what data they relied on, or their known limitations. Without this information, clinicians cannot fully evaluate the reliability of these tools, and patients cannot make informed decisions about technologies affecting their care.

Barriers to overriding A.I.-driven decisions. In practice, employer retaliation can discourage health care workers from disregarding automated recommendations. Staffing systems and productivity monitoring technologies put pressure on workers to follow algorithmic outputs even when the output conflicts with a nurse's professional judgment and assessment of a patient's needs. Workers may fear

A.B. 2575 continued »»

«« A.B. 2575 continued from front

discipline, performance penalties, or workload consequences for deviating from automated guidance, which can undermine professional judgment and place patients at risk.

Accountability gaps when harm occurs. Health care entities increasingly require health care workers to use opaque and error-prone A.I. systems in patient care. These A.I. systems generate recommendations that clinicians are expected to follow but cannot independently evaluate or easily override. This dynamic increases the risk of patient harm and exposes clinicians to liability for decisions imposed by technology the clinician did not design and cannot fully assess. When these systems contribute to injury or adverse patient outcomes, developers and deploying entities can attempt to evade liability by shifting responsibility on to a human health care worker in the loop. Developers and deployers of A.I. tools in health care should not be able to avoid responsibility when their tools contribute to adverse outcomes.

California must establish safeguards now to ensure these technologies support safe patient care rather than undermine transparency, professional judgement, and accountability.

SOLUTION

A.B. 2575 directly addresses gaps in transparency, worker protections, and accountability for the use of A.I. in health care.

Transparency. It requires hospitals, clinics, and other health care entities that use A.I. in patient care to disclose when they use these tools and provide information about the tool's developer, intended use, data inputs, performance, and known risks and limitations so clinicians and patients can make informed decisions about care.

Professional Judgement Protection. It affirms health care workers' right to exercise their professional judgement and override unsafe or inappropriate A.I. outputs and prohibits employer retaliation against workers for doing so.

Accountability. The bill also holds A.I. developers and deployers accountable for harm caused by their systems. Specifically, developers and deployers of A.I. systems cannot use a human health care worker in the loop as a defense against liability when those systems contribute to injury or adverse patient outcomes.

Together, these provisions create a practical framework for responsible use of A.I. that informs patients and workers, protects professional judgement in patient care, and holds developers and deployers of A.I. accountable for harm.

CONTACT

JP Hanna » California Nurses Association
jhanna@calnurses.org » 916-812-0118

SUPPORT

California Nurses Association (sponsor)

ENDNOTES

1 • Ross C. & Herman B. (Mar. 2023). Denied by A.I.: How Medicare Advantage Plans Use Algorithms to Cut Off Care for Seniors in Need. STAT. https://www.statnews.com/2023/03/13/medicare-advantage-plans-denial-artificial-intelligence

Rucker P. et al. (Mar. 25, 2023). How Cigna Saves Millions by Having Its Doctors Reject Claims Without Reading Them. ProPublica. https://www.propublica.org/article/cigna-pxdx-medical-health-insurance-rejection-claims

For further discussion algorithmic bias in health care, see National Nurses United (May 11, 2021). Comments to AHRQ, Use of Clinical Algorithms That Have the Potential to Introduce Racial/Ethnic Bias Into Healthcare Delivery. 86 Fed. Reg. 12,948 (Mar. 5, 2021). https://www.nationalnursesunited.org/sites/default/files/nnu/documents/NNU_Comments_AHRQ_Clinical_Algorithms_Racial_Ethnic_Bias.pdf

2 • Obermeyer Z. et al. (Oct. 2019). Dissecting Racial Bias in in Algorithm Used to Manage the Health of Populations. Science, 366(6464), 447-53. https://science.sciencemag.org/content/366/6464/447

3 • National Nurses United (May 2024). National Nurses United survey finds A.I. technology degrades and undermines patient safety. https://www.nationalnursesunited.org/press/national-nurses-united-survey-finds-ai-technology-undermines-patient-safety

4 • U.S. Senate. (2025). S. 2081 – Responsible Innovation and Safe Expertise (RISE) Act of 2025, 119th Congress. https://www.congress.gov/bill/119th-congress/senate-bill/2081