Int. j. adv. multidisc. res. stud. 2026; 6(6):896-898

Received: 30-09-2026
Accepted: 10-11-2026

International Journal of Advanced Multidisciplinary Research and Studies

ISSN: 2583-049X

Who Reports the AI Near Miss? Integrating Nursing into AI Safety Event Reporting and Organizational Learning

[1] Preeti Ramesh Mahadik, [2] Deepali Sudhakar More, [3] Elinor Sanjay Salve, [4] Pratik Ravindra Salve, [5] Neha Michael-Salve

[1] 2nd Year M.Sc. Nursing - Mental Health Nursing, Sancheti Institute of Nursing Education, Pune, India
[2] Clinical Instructor, M.Sc. Nursing, Sancheti Institute of Nursing Education, Pune, India
[3] Clinical Instructor, M.Sc. in Child Health Nursing, Pune, India
[4] Assistant Professor, M.Sc., MAEER's VishwaRaj Institute of Nursing, Pune, Maharashtra, India
[5] Clinical Instructor, M.Sc., Sadhu Vaswani College of Nursing, Pune, Maharashtra, India

Corresponding Author: **Pratik Ravindra Salve**

Abstract

Artificial intelligence (AI) is increasingly integrated into clinical practice through clinical decision support, predictive analytics, generative AI, automated documentation, and other technology-assisted processes. As AI becomes embedded in patient care, ensuring safety requires healthcare organizations to identify, report, investigate, and learn from situations in which AI does not perform as intended. An AI error should not automatically be considered an AI-caused adverse event, as safety events may arise from interactions among AI systems, clinicians, data, interfaces, workflows, and organizational factors. Near misses are particularly important because nurses may identify inaccurate recommendations, clinically significant omissions, or anomalous AI outputs before they affect patient care. Nurses should therefore be integrated into AI-related safety event reporting and organizational learning systems, without making them solely responsible for determining causation or managing AI-related risks. A "trust but verify" approach is essential, requiring nurses to compare AI-generated recommendations with the patient's clinical condition, available evidence, and professional judgment. When discrepancies are identified, they should be questioned, verified, and reported when appropriate. Integrating nursing observations into nonpunitive, multidisciplinary reporting and learning systems can support continuous monitoring, system improvement, and safer implementation of AI in healthcare.

Keywords: Artificial Intelligence, Patient Safety, Nursing, AI-related Safety Events, Organizational Learning

Introduction

As artificial intelligence (AI) is integrated into clinical practice, the question of patient safety evolves from can AI do this and that correctly under surveillance to how can we detect and learn from situations in which AI delivers incorrectly. Clinical decision support, predictive analytics, generative, documentation, and other AI applications can generate novel modes of failure that are not apparent at the time of introduction. Recent studies have identified health-related AI incidents in public repositories; current guidance promotes voluntary reporting of safety events [1-3]. The challenge is to make these events visible within clinical safety systems.

A crucially important distinction must be made at the very beginning: An AI error is not necessarily an AI-caused adverse event. An erroneous AI output might be caught before reaching the care environment; a clinician may inappropriately apply an otherwise correct AI output; an adverse event may occur through a complex set of interactions among AI, clinician, workflow, data, interface, and organization; or an AI system may be a contributing factor to an adverse event without causing it. Thus, reporting should begin with a potential AI-related safety event, with subsequent analysis determining whether and how an AI contributed.

This is a crucially important distinction because near misses can be safety events too. A nurse may recognize that the AI recommendation is wrong and take corrective action, or note a clinically significant omission from an AI-generated summary,

or see an anomalous output before it reaches the care environment. The patient may be harmed, or the patient may not be harmed, but either way, a learning opportunity is lost if the event is not reported. Patient-safety scholars have long recognized near misses as potential learning events, warning that the ability to identify latent hazards is diminished when only adverse events are subjected to analysis [4]. Likewise, a systematic review of AI and clinical risk management in 2024 concluded that incident reporting was a critical element for AI-related risk management, including the need for standardized terminology to describe reported events [5].

Healthcare organizations are beginning to respond to these needs. The Joint Commission and Coalition for Health AI encourages voluntary, blinded reporting of safety events, including unsafe recommendations from AI, near misses, performance issues after software updates, and biased outputs. The guidance promotes using existing patient safety reporting systems, where possible, rather than creating new ones [2]. It is a good start, but without a way for clinicians to know what to report and how to do it, the entire process can fail.

Nurses deserve special consideration for inclusion within incident-reporting systems. Nurses are at the intersection of patient care, documentation, handoffs, communication, and technology-assisted processes. Their roles offer multiple opportunities to act on discrepancies between what an AI system is asserting or predicting and what is actually happening with or to the patient. However, it would be incorrect to assume that nurses are the only potential safety detectors for AI-related events. This is an opportune moment to propose that nurses may comprise an important frontline for identifying AI-related safety events, and that their voices should be included in organizational learning and reporting systems.

Recent findings offer a compelling rationale for encouraging such inclusion. In a national cross-sectional survey of 1,189 US nurses conducted in June and July 2026, 424 nurses responded that they had encountered an AI-related patient-safety concern; of those, 83 (19.6%) indicated that they did not have a formal method for reporting that concern [6]. The study highlights disparities in the current state of AI-related incident reporting for US nurses, although its findings should not be extrapolated to nurses in other settings. It is also important to note that the study does not establish that nurses are uniquely positioned to identify AI incidents, only that they are currently unable to contribute meaningfully to incident learning in the absence of reporting systems that include them.

It is therefore fitting that the reporting infrastructure begins with the event, rather than the assumption of AI causation. A reporting system for AI-related events might ask whether AI was a factor, what kind of system was in use, what output or behaviour raised concern, whether the concern preceded patient exposure, whether clinical action was taken, what harm or potential for harm was present, and what other human, technical, data, or organizational factors were present. Such an approach would facilitate understanding of whether an AI output error or an AI-assisted clinical error occurred, whether automation bias or human factors contributed to the event, or whether there was an AI system failure or inappropriate use of a functioning system.

The medication error reporting literature offers some guidance on this subject. Medication-safety incidents are valuable learning events; clinicians and organizations can benefit from understanding what went wrong and why, with the ultimate goal of preventing future errors [4, 7]. The parallels with AI incidents are instructive, but medication incidents are fundamentally different from AI incidents in one essential way: Medication incidents represent failures within a specific medication use process, whereas AI incidents cut across a range of factors, including training data, model characteristics, software design, interface design, model versioning, clinical context, human factors, and organizational process. Thus, medication incident-reporting templates should not be used for AI incident reporting; rather, additional fields may be needed to characterize the role of AI within an incident.

A potential AI safety-learning pathway might begin with an AI output or behaviour and proceed through clinical use and verification, potential safety event, mitigation, incident/reporting, multidisciplinary review, socio-technical root-cause analysis, correction, and AI, workflow, and human-factor re-evaluation, culminating in organizational learning.

Notably, the pathway separates the act of identifying an event from the need to determine causation. A nurse who has observed an event should not be required to determine whether the cause was the AI model, the clinician, the data, the interface, the workflow, or the organization. The nurse's role is to serve as a safety signal; the multidisciplinary review process is necessary to identify the root cause.

This distinction is critical to ensuring that the adoption of AI does not produce an undue burden on clinicians. Nurses should not be the solution to the challenges of ensuring clinical safety with AI, and nurses should not bear the increased risk of litigation and disciplinary action that can accompany incident reporting. The reporting infrastructure should therefore be designed to be nonpunitive and promoting of learning. Evidence from medication incident reporting suggests that nurses' willingness to report medication errors is affected by concerns about negative consequences to themselves and their organizations [7]. Similar dynamics are likely to occur with AI incident reporting, although there is as yet little evidence on the subject.

The implications of these considerations extend beyond day-to-day nursing practice. Nursing education must prepare nurses to use, evaluate, and critique AI systems, including recognizing potential AI safety events and understanding the local options for learning and reporting. Nurse leaders can promote a safety culture that encourages report generation and ensures that frontline nurses' reports reach the appropriate patient-safety, informatics, quality improvement, and AI governance entities. Healthcare organizations must ensure that learning systems capture both clinical safety events and near misses, while informing the relevant personnel about the needed information for AI analysis (such as model version, intended use, and implementation context). Informatics personnel and AI developers, in turn, must be aware of frontline nurses' reports as post-deployment evidence for re-evaluation of the system.

The question of the moment therefore turns from whether AI can work to whether healthcare systems can learn from instances in which it does not work as intended. There is already evidence that health-related AI incidents are being reported, but the current state of affairs suggests that such

incidents are few and far between, both because of the existence of only a handful of public repositories and the low volume of incidents within those repositories [1]. A mature safety culture will ensure that near misses and adverse events alike are reported, understood, and acted upon to improve the safety of patients, clinicians, and the healthcare system as a whole.

Nurses should have a place in that culture—not as the sole judges of AI-related safety events, and certainly not as the sole safety net for flawed systems, but as contributors to learning and improvement. Enhancing the role of nurses within AI incident reporting and organizational learning may help healthcare organizations evolve from a model in which AI performance is mostly evaluated at the time of deployment to one in which safety is continuously learned and improved upon after deployment.

AI can be a valuable clinical support tool, but its outputs should not be accepted without professional judgment. Nurses should "trust but verify" AI-generated recommendations by comparing them with the patient's clinical condition, available evidence, and their own professional assessment [8]. When an AI output appears inaccurate, incomplete, or inconsistent with clinical reality, it should be questioned, verified, and reported when appropriate. This approach helps preserve human oversight and strengthens patient safety.

References

1. Denecke K, Rivera-Romero O, López-Campos G, Dorronzoro E, Gabarron E. Uncovering AI's hidden risks: An empirical analysis of health-related AI incidents and their ethical implications. AI Ethics. 2026; 6:169. Doi: 10.1007/s43681-026-01012-7
2. The Joint Commission. Coalition for Health AI. The Responsible Use of AI in Healthcare. Oakbrook Terrace, IL: The Joint Commission, 2025.
3. U.S. Food and Drug Administration. Artificial Intelligence-Enabled Device Software Functions: Lifecycle Management and Marketing Submission Recommendations. Draft Guidance for Industry and Food and Drug Administration Staff. Silver Spring, MD: FDA, 2025.
4. World Health Organization. Patient Safety Curriculum Guide: Multi-professional Edition. Geneva: World Health Organization.
5. Ferrara M, Bertozzi G, Di Fazio N, *et al*. Risk management and patient safety in the artificial intelligence era: A systematic review. Healthcare (Basel). 2024; 12(5):549. Doi: 10.3390/healthcare12050549
6. Carter GA. A national cross-sectional survey of artificial intelligence patient safety reporting infrastructure available to US nurses. J Nurs Care Qual, September 10, 2026. Doi: 10.1097/NCQ.0000000000001011
7. Braiki R, Douville F, Gagnon MP. Factors influencing the reporting of medication errors and near misses among nurses: A systematic mixed methods review. Int J Nurs Pract. 2024; 30(6):e13299. Doi: 10.1111/ijn.13299
8. Salve P, Ray S, Ray SP, Geddugol BJ, Nadaf H, Jabade M, *et al*. Trust and Reliance on Artificial Intelligence–Based Clinical Decision Support Systems. Implications for Nursing Practice. Investigacion y Educacion en Enfermeria. 2026; 44(2). Doi: 10.17533/udea.iee.v44n2e02