Submitted: 12 September, 2025 Accepted: 21 October, 2025 Published: 08 January, 2026 DOI:10.22514/sv.2026.006

Open Access
Signa Vitae

COMMENTARY

The lethal diamond of AI issue in critical care medicine: Great Power, Great Risk

Elena G. Bignami[1,*], Roberto Lanza[1], Michele Russo[1], Valentina Bellini[1]

[1]Department of Anesthesiology, Critical Care and Pain Medicine, University of Parma, 43126 Parma, Italy

***Correspondence**
elenagiovanna.bignami@unipr.it
(Elena G. Bignami)

Abstract

The integration of artificial intelligence (AI) into critical care promises transformative advancements in diagnostics, treatment, and resource allocation. However, the realization of AI's potential is hampered by critical gaps in data quality, ethical considerations, clinician education, and validation. This article introduces the "Lethal Diamond" framework, expanding on the traditional "Lethal Triad", to encompass these interconnected challenges that threaten AI's safe and effective deployment in high-stake critical care settings. It further proposes a shift from the reactive approach of "Garbage In, Garbage Out" (GIGO) to a proactive "Digging In, Diamonds Out" (DIDO) paradigm to cultivate excellence in AI implementation, ensuring AI systems enhance and augment, rather than undermine, healthcare delivery.

Keywords

AI; Data quality; Ethical; Validation; GIGO; Healthcare; Anesthesia; Critical care; Machine learning; Medical education

1. Introduction

Artificial intelligence (AI) is rapidly transforming healthcare, with the potential to revolutionize medical care [1]. Machine learning algorithms can now assist with tasks ranging from early disease identification to personalized treatment strategies and optimized resource allocation [2]. For example, AI-powered systems can analyze complex physiological data to predict patient deterioration, optimize ventilator settings, and personalize drug dosages [3]. However, realizing AI's promise in critical care medicine (CCM) requires confronting significant challenges, previously identified as the "Lethal Triad": poor data quality, ethical and legal ambiguities, and insufficient clinician education [4].

As AI systems evolve and their applications in critical care expand, a fourth critical pillar has emerged: the lack of robust external validation and real-world implementation frameworks [5]. Together, these challenges form what we term the "Lethal Diamond", an interdependent web of vulnerabilities that threatens AI's safe and effective deployment in high-stake CCM settings. These interconnected vulnerabilities are especially pertinent in acute care settings, where real-time decision-making is crucial, and the consequences of errors can be life-threatening [6].

This article traces the evolution from the Lethal Triad to the Lethal Diamond, analyzing progress and persistent failures across each domain. It further introduces the "Digging In, Diamonds Out" (DIDO) framework, a proactive paradigm shift from the reactive approach of "Garbage In, Garbage Out" (GIGO), to cultivate excellence in AI implementation for CCM. The goal is to raise awareness and provide workable starting points through the DIDO framework, aiding clinicians, policymakers, and developers in confronting the complexities of AI in critical care and unlocking its true potential to improve patient outcomes and enhance the efficiency of healthcare systems.

2. Data quality: from garbage in, garbage out to data-aware AI

The "Garbage In, Garbage Out" (GIGO) paradigm has long been a fundamental concern in AI development. While the technical components are clear, biased datasets, uncalibrated sensors, and non-standardized data formats inherently compromise model development, thus GIGO is not solely a technical flaw. It often stems from a critical human factor: the natural tendency to follow the path of least resistance in data handling and curation. However, this is frequently compounded by a human tendency toward resistance to change, a natural inclination to follow the path of least resistance, which we can translate as the famous: "it has always been done this way". In the context of data curation, this manifests as the avoidance of the labor-intensive, time-consuming tasks of data cleaning, rigorous annotation, and bias auditing. This is especially perilous in the era of large public datasets, where the accessibility of data can create an illusion of readiness, tempting researchers to bypass the essential, albeit tedious, groundwork. Therefore, overcoming GIGO requires a dual approach: addressing the technical vulnerabilities through robust pipelines and, crucially, fostering a cultural shift that values and rewards data

This is an open access article under the CC BY 4.0 license (https://creativecommons.org/licenses/by/4.0/).
 ©2026 The Author(s). Published by MRE Press.

diligence as a non-negotiable pillar of scientific rigor.

Algorithms trained on retrospective, noisy, biased, and unrepresentative datasets can produce misleading clinical outputs, undermining trust in AI's recommendations [7]. Many existing AI models in healthcare also lack adequate data labelling, and the heterogeneity of data collection limits their reproducibility [8]. However, improvements in data governance, bias auditing tools, and regulatory frameworks are paving the way for "Data-Aware AI" [9].

Drawing from insights in intelligent monitoring [6], data quality in CCM is vital. For example, for vital tracking of patients (*e.g.*, blood pressure, heart rate), accurate and calibrated sensors are needed to ensure reliable data. In the Intensive Care Unit (ICU) context, this extends to addressing motion artifacts in arterial line waveforms, standardizing data from diverse ventilator and monitor manufacturers, and handling the high frequency of missing data points common in critical care settings. Poor data quality can reduce model accuracy; therefore, it is a pillar of the Diamond, one of the barriers. Tools and techniques include standardizing data collection, cleaning methods, and robust data validation pipelines. High-quality data acquisition is the foundation of effective AI in critical care [7]. The diligence required for data curation cannot be overstated, especially with the increasing use of large public datasets like the Medical Information Mart for Intensive Care-IV (MIMIC-IV). These repositories must lead by example, providing exhaustive data quality reports and rigorous de-identification. Similarly, peer reviewers must apply the same scrutiny to studies using these datasets as they would to primary data collection, checking for documented data cleaning protocols, handling of missingness, and potential systemic biases inherent in the source data.

However, to avoid this data being a limitation to developing clinical decision support system (CDSS), we propose responsible sharing of large ICU datasets at all levels, implying finding the right balance between privacy protections and data usability.

DIDO: "Digging In" means rigorous data curation, annotation, and validation. DIDO: "Diamonds out" are well-structured, purpose-specific data that promotes reliability, reproducibility, and explainability.

3. Ethical and legal considerations: from legal vacuum to uneven regulation

Initially, the absence of AI-specific legal frameworks created a "grey zone" regarding legal liability, patient consent, and intellectual property ownership [10]. Algorithmic bias, privacy breaches, and "data colonialism" in underrepresented populations raised significant ethical concerns [11].

While legislation like the General Data Protection Regulation (GDPR) [9] and ethical frameworks from organizations like the World Health Organization (WHO) [12] have made inroads, new challenges have emerged. The rise of generative models and "black box" systems has heightened concerns around traceability, explainability, and misuse. Ethical considerations in post-operative care require doctors to integrate medical knowledge with the AI model's output in medical treatment [6].

In critical care, this becomes particularly acute when considering AI predictions of futility in comatose patients, or the legal liability surrounding a CDSS-recommended vasopressor dose that leads to ischemic complications. To avert AI-legitimized and AI-enabled further marginalization of those already disproportionately burdened by disease and societal inequities, regulatory guardrails are needed [13].

DIDO: Ethical oversight and fairness, and protection of patient rights and autonomy. DIDO: Transparency, accountability, and bias mitigation in AI algorithms and decision-making processes.

4. Cultivating digital intuition

Structural and evaluative deficiencies persist in AI medical education [14]. The fundamental challenge we face is not just teaching AI to clinicians, but rebuilding the bridge between two worlds that have grown too far apart: the art of medicine and the science of algorithms. How can we expect intensivists, already overwhelmed by the complexities of critical care, to simultaneously become data scientists? The answer lies not in creating clinician-programmers, but in forging clinician-interpreters.

We must start earlier. The AI conversation needs to enter the medical school classroom, where future doctors first learn to think clinically. Imagine medical students not just memorizing drug interactions, but also understanding how an algorithm might predict them; not just interpreting an electrocardiogram (ECG), but comprehending how machine learning could flag subtle patterns invisible to the human eye. This isn't about writing code, it's about developing a critical "digital intuition" that will allow them to be informed, skeptical, and ultimately, empowered users of AI at the bedside.

This foundational literacy is the prerequisite for the specialized training that follows in critical care and beyond. It enables the specific instruction on interpreting AI-generated "early warning" scores for sepsis or predicting weaning failure from mechanical ventilation within the complex ICU workflow [15]. Today, clinicians can choose from a diverse ecosystem of educational pathways tailored to different needs and career stages. For those seeking immersive clinical training, programs like the Clinical Fellow in Digital Health/Artificial Intelligence in Critical Care at the University of Oxford offer direct hands-on experience at the AI-ICU interface. For building fundamental computational thinking, certificate programs like Harvard's CS50 provide accessible entry points into computer science principles. At the pan-European level, the European Association for Artificial Intelligence (EurAI) not only sets standards, but also offers prestigious Fellowship recognition for excellence in AI research. Meanwhile, organizations like The Alan Turing Institute deliver flexible, online data science modules that accommodate the demanding schedules of practicing clinicians. This rich tapestry of opportunities makes advanced AI education increasingly accessible to healthcare professionals worldwide.

Without this foundational literacy, the complexity of AI algorithms will continue to breed distrust and impede adoption. Moving beyond *ad-hoc* learning, achieving true "AI literacy" means weaving digital competence into the very fabric of medical expertise. This transformed knowledge is a prerequisite for the high-quality data handling, critical evaluation of AI outputs, and the safe integration of these technologies into the high-stake environment of critical care. Educational initiatives must, therefore, focus on developing "AI literacy" among clinicians, enabling them to critically evaluate AI outputs, understand their limitations, and integrate them effectively into clinical decision-making [16]. DIDO: Robust training programs for effective interpretation and utilization of AI tools, understanding the limitations and biases. DIDO: Promote clinical expertise, AI integration in decision-making.

5. Validation and implementation in real-world deployment

The translation of AI from research to clinical practice faces a critical implementation gap. Models may excel in controlled trials but fail to undergo rigorous external validation across diverse care settings [17]. This validation deficit becomes particularly apparent when algorithms trained at academic centers underperform in community hospitals with different patient demographics and data acquisition protocols. For instance, a model trained on a general ICU population may fail when applied to a specialized cardiothoracic ICU with different monitoring protocols and patient pathologies. Such issues can be addressed with a combination of strong clinical and technical understanding [18]. Before an AI model is used, it must be tested to show that it improves health [19].

Beyond the lack of representativeness in training data, non-academic and resource-limited hospitals face profound Information Technology (IT) infrastructure challenges that hinder AI implementation. The technical barriers are particularly acute in critical care settings, where AI systems must process real-time, multimodal data streams from bedside monitors, infusion pumps, ventilators, and electronic health records. Unlike more structured domains like medical imaging, ICU data presents unique challenges, including heterogeneous formats, inconsistent sampling frequencies, significant data latency issues, and a fundamental lack of interoperability between systems (Health Level Seven/Fast Healthcare Interoperability Resources) [20]. These technical obstacles, combined with the high computational requirements for real-time processing and substantial maintenance costs, create a formidable digital divide. This disparity risks creating a two-tier healthcare system where advanced AI tools only benefit well-funded institutions, thereby exacerbating existing health inequities on a global scale. Any implementation strategy must, therefore, include comprehensive assessments of technical feasibility and sustainable plans for infrastructural support.

DIDO: AI models should be rigorously and fairly tested in real world. DIDO: A clear protocol needs to be implemented when integrating a new AI model for the benefit of both clinicians and patients.

6. The negative loop and the need for DIDO

The four components of the Lethal Diamond are not isolated; they form a self-reinforcing negative feedback loop, as illustrated in Fig. 1.

i. Data quality ↔ Validation: Poor data quality undermines external validation, rendering even robust models unreliable. If we train an AI on messy, low-quality data from just one hospital, it's like a student who only studies from a single textbook. They might ace the practice test, but they're completely unprepared for the final exam. In the AI world, that "final exam" is external validation—it's when we test the AI on fresh, unseen data from a completely different hospital. Without high-quality, diverse data, the model is guaranteed to fail this real-world test, rendering even a smart algorithm useless and unsafe elsewhere.

ii. Validation ↔ Educational Programs: Insufficient clinician training leads to improper use of AI tools, corrupting real-world validation through misuse or misinterpretation.

Conversely, the results from real-world validation studies provide critical, practical case studies that should be fed back into educational programs, making training more relevant and evidence-based.

iii. Educational Programs ↔ Ethics & Legal: Inadequate training fosters blind trust in AI, escalating ethical risks. Furthermore, comprehensive educational programs equip clinicians and researchers with the skills for meticulous data annotation and curation, directly improving data quality at its source through better documentation practices and bias awareness.

iv. Ethics & Legal ↔ Data quality: Ambiguous regulations paralyze data sharing, crippling multi-institutional validation and progress. Everyone agrees that to build better AI, we need to learn from diverse patients across many hospitals. But strict, often ambiguous, privacy rules make hospitals terrified to share data. This locks away the very information we need. We're not asking for reckless sharing, but for smart, secure ways, like a digital fortress, that lets researchers collaborate on data without ever exposing a single patient's identity. Without this secure data handshake, progress is paralyzed.

7. Comparative analysis of the four lethal diamond components

A comparative analysis of the progress and persistent gaps for each component of the Lethal Diamond, from past to present, is presented in Table 1.

8. The DIDO approach in action

AI is not merely a diagnostic tool; it is a roadmap for transformation in healthcare, particularly in critical care. The practical viability of the DIDO framework is demonstrated by several successful real-world implementations that have overcome the challenges of the Lethal Diamond:

- AI-powered ECG interpretation has proven capable of accelerating catheterization lab activation for subtle myocardial infarction cases, demonstrating how robust algorithms can enhance time-sensitive clinical workflows [21].
- Sepsis detection algorithms have not only improved adherence to treatment guidelines, but have also reduced mortality in some implementations, showcasing the potential for AI to directly impact patient outcomes when properly validated and integrated [22].

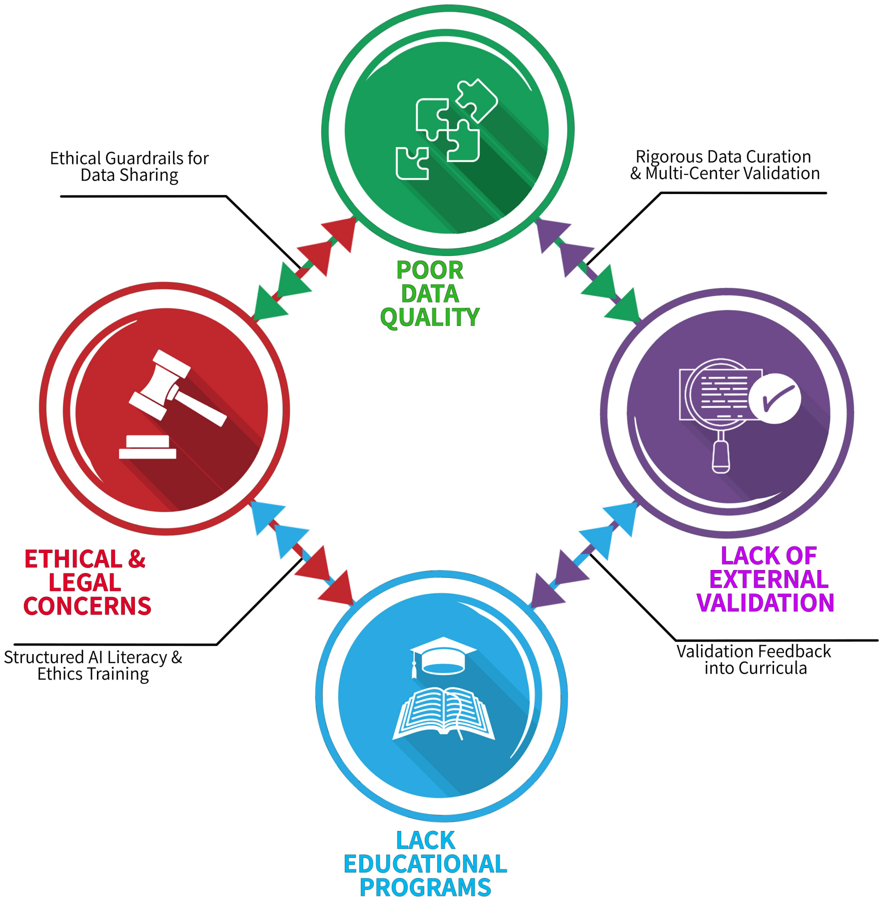

FIGURE 1. The "Lethal Diamond" of AI in anesthesia and ICU. This diagram illustrates the four major, interconnected barriers: Poor Data Quality, Ethical & Legal Concerns, Lack of External Validation, and Lack of Educational Programs, that hinder the safe, effective, and widespread implementation of artificial intelligence in anesthesia and critical care settings. The bidirectional arrows represent the negative feedback loops that create a self-reinforcing cycle of failure. Along each connection, we can find intervention points that can break these vicious cycles: (1) "Rigorous Data Curation & Multi-Center Validation" interrupts the link between poor data and failed validation. (2) "Validation Feedback into Curricula" ensures real-world performance informs training. (3) "Structured AI Literacy & Ethics Training" prevents educational gaps from exacerbating ethical risks. (4) "Ethical Guardrails for Data Sharing" enables responsible data sharing that improves quality and diversity.

TABLE 1. AI in healthcare: comparative analysis of the four lethal diamond components.

Component	2021 Challenges	2025 Progress	Remaining Gaps
Data Quality	GIGO principle ignored. Biases in datasets. Limited transparency.	Fair ML standards adopted liability. Bias audits (*e.g.*, EU AI Act).	Small hospitals lack resources. Retrospective data dominance.
Ethical/Legal	No AI-specific laws. Liability ambiguity.	GDPR updates for AI. WHO equity guidelines.	Algorithmic colonialism. Patent conflicts.
Education	UME-focused (39%). No frameworks/theories. 12% evaluated (Levels 1–2).	Specialty frameworks (*e.g.*, radiology). More PGME/CME programs. 32% evaluated (still no Levels 3–4).	No faculty training. No behavioral impact studies. Unstandardized accreditation.
External Validation/ Real-World	Rare external validation. Lab-only models.	FDA Pre-Cert Program. NHS AI Sandbox trials.	Generalizability failures. Costly longitudinal studies.

Diamond components, highlighting regulatory, educational, and technological advancements alongside unresolved barriers (Past vs. Present).
GIGO: Garbage In, Garbage Out; ML: Machine Learning; EU AI: Europe Artificial Intelligence; GDPR: General Data Protection Regulation; WHO: World Health Organization; UME: Undergraduate Medical Education; PGME: Postgraduate Medical Education; CME: Continuing Medical Education; FDA: Food and Drug Administration; NHS: National Health System.

- Ambient AI documentation systems, such as Kaiser Permanente's deployment, have generated millions of structured clinical notes while significantly reducing administrative burden, proving that AI can create operational efficiencies without disrupting clinical practice [23]. These successes share common DIDO principles: they target specific workflow bottlenecks, integrate with minimal friction, produce actionable outputs, and complement rather than replace clinical judgment. However, the path to implementation faces significant challenges. Machine learning (ML)—based artificial intelligence (AI) techniques, while capable of defining clinical states and predicting future events, often suffer from an inherent "black-box" nature. This opacity in many predictive algorithms and Clinical Decision Support Systems (CDSS) fundamentally undermines trustworthiness and acceptance by the medical community [24]. To address these barriers and systematically translate the DIDO framework from concept to practice, we developed a conceptual flowchart that maps abstract challenges to concrete solutions (Fig. 2). This visual guide demonstrates how to address each pillar of the Lethal Diamond through specific, actionable steps, with particular emphasis on creating transparent, interpretable systems that can earn clinician trust while delivering measurable clinical benefits.

9. Conclusions

To realize AI's transformative potential, a sustained multidisciplinary commitment is needed, uniting clinicians, engineers, ethicists, educators, and policymakers. A dynamic approach that involves regular assessment and refinement of AI technology is essential to align it with evolving healthcare needs and technological advancements. The DIDO paradigm offers a framework for responsible AI implementation. By institutionalizing DIDO's pillars, technical robustness, ethical foresight,

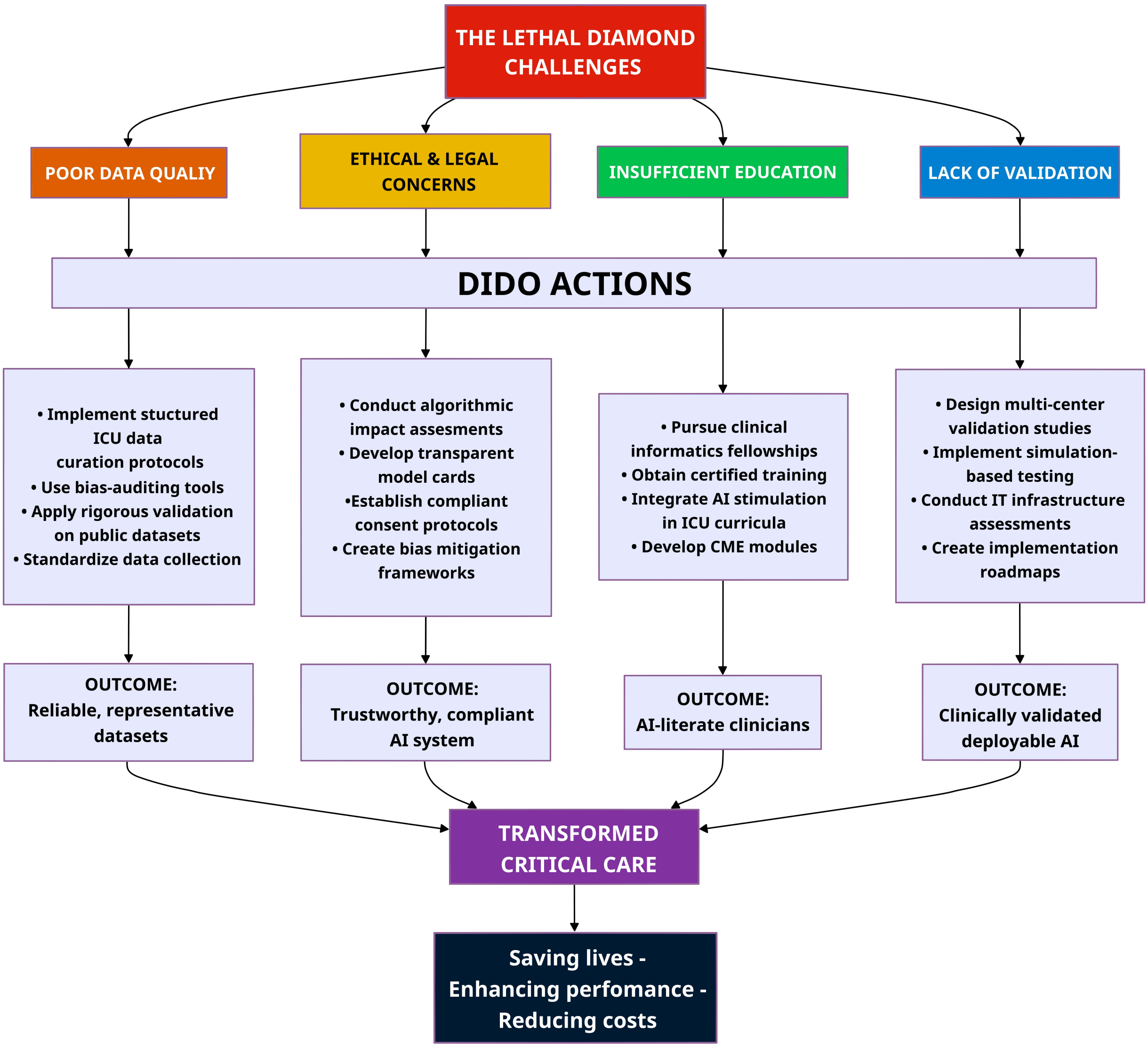

FIGURE 2. The DIDO implementation pathway: from abstract challenges to concrete solutions in critical care AI. This visual guide translates the four pillars of the Lethal Diamond into actionable strategies, providing clinicians and developers with a practical framework to systematically address data quality, ethical concerns, educational gaps, and validation challenges in healthcare AI. CME: Continuing Medical Education; ICU: Intensive Care Unit; IT: Information Technology; AI: Artificial Intelligence.

clinician empowerment, and equitable validation, critical care can transition from damage control to value creation. Implementing this vision requires targeted strategies for clinical adoption, rigorous model management, and foundational changes in medical education. It is through this multifaceted approach that AI will transition from a technological novelty to an indispensable tool for enhancing healthcare quality [2].

AVAILABILITY OF DATA AND MATERIALS

Not applicable.

AUTHOR CONTRIBUTIONS

EGB—Conceptualization, Writing-Original Draft. RL, MR, VB—Writing-Review & Editing.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

ACKNOWLEDGMENT

Not applicable.

FUNDING

This research received no external funding.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

REFERENCES

[1] Pinsky MR, Bedoya A, Bihorac A, Celi L, Churpek M, Economou-Zavlanos NJ, *et al*. Use of artificial intelligence in critical care: opportunities and obstacles. Critical Care. 2024; 28: 113.

[2] Cai X, Wang X, Zhu Y, Yao Y, Chen J. Advances in automated anesthesia: a comprehensive review. Anesthesiology and Perioperative Science. 2025; 3: 3.

[3] Porcellato E, Lanera C, Ocagli H, Danielis M. Exploring applications of artificial intelligence in critical care nursing: a systematic review. Nursing Reports. 2025; 15: 55.

[4] Bellini V, Montomoli J, Bignami E. Poor quality data, privacy, lack of certifications: the lethal triad of new technologies in intensive care. Intensive Care Medicine. 2021; 47: 1052–1053.

[5] Cecconi M, Greco M, Shickel B, Angus DC, Bailey H, Bignami E, *et al*. Implementing artificial intelligence in critical care medicine: a consensus of 22. Critical Care. 2025; 29: 290.

[6] Cao Y, Wang Y, Liu H, Wu L. Artificial intelligence revolutionizing anesthesia management: advances and prospects in intelligent anesthesia technology. Frontiers in Medicine. 2025; 12: 1571725.

[7] Price WN, Zhang A. Legal and policy barriers to fairness in AI. In Dubber MD, Pasquale F, Das S (eds.) The Oxford handbook of ethics of AI (pp. 123–145). 1st edn. Oxford University Press: New York, NY. 2020.

[8] Berkhout WEM, van Wijngaarden JJ, Workum JD, van de Sande D, Hilling DE, Jung C, *et al*. Operationalization of artificial intelligence applications in the intensive care unit: a systematic review. JAMA Network Open. 2025; 8: e2522866.

[9] Official Journal of the European Union. Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024. 2024. Available at: http://data.europa.eu/eli/reg/2024/1689/oj (Accessed: 03 July 2025).

[10] Pham T. Ethical and legal considerations in healthcare AI: innovation and policy for safe and fair use. Royal Society Open Science. 2025; 12: 241873.

[11] Mennella C, Maniscalco U, De Pietro G, Esposito M. Ethical and regulatory challenges of AI technologies in healthcare: a narrative review. Heliyon. 2024; 10: e26297.

[12] World Health Organization. Ethics and governance of artificial intelligence for health. 2021. Available at: https://www.who.int/publications/i/item/9789240029200 (Accessed: 05 July 2025).

[13] Montomoli J, Bitondo MM, Cascella M, Rezoagli E, Romeo L, Bellini V, *et al*. Algor-ethics: charting the ethical path for AI in critical care. Journal of Clinical Monitoring and Computing. 2024; 38: 931–939.

[14] Tolentino R, Baradaran A, Gore G, Pluye P, Abbasgholizadeh-Rahimi S. Curriculum frameworks and educational programs in AI for medical students, residents, and practicing physicians: scoping review. JMIR Medical Education. 2024; 10: e54793.

[15] Valikodath NG, Cole E, Ting DSW, Campbell JP, Pasquale LR, Chiang MF, *et al*.; American Academy of Ophthalmology Task Force on Artificial Intelligence. Impact of artificial intelligence on medical education in ophthalmology. Translational Vision Science & Technology. 2021; 10: 14.

[16] European Commission. AI talent, skills and literacy. 2024. Available at: https://digital-strategy.ec.europa.eu/en/policies/ai-talent-skills-and-literacy (Accessed: 15 January 2025).

[17] US Food and Drug Administration. Digital Health Software Precertification (Pre-Cert) Pilot Program. 2024. Available at: https://www.fda.gov/medical-devices/digital-health-center-excellence/digital-health-software-precertification-pre-cert-pilot-program (Accessed: 15 January 2025).

[18] UK Government. AI airlock: the regulatory sandbox for AIaMD. 2024. Available at: https://www.gov.uk/government/collections/ai-airlock-the-regulatory-sandbox-for-aiamd (Accessed: 15 January 2025).

[19] Elendu C, Amaechi DC, Elendu TC, Jingwa KA, Okoye OK, John Okah M, *et al*. Ethical implications of AI and robotics in healthcare: a review. Medicine. 2023; 102: e36671.

[20] Nauka PC, Kennedy JN, Brant EB, Komorowski M, Pirracchio R, Angus DC, *et al*. Challenges with reinforcement learning model transportability for sepsis treatment in emergency care. NPJ Digital Medicine. 2025; 8: 91.

[21] Lin C, Liu WT, Chang CH, Lee CC, Hsing SH, Fang WH, *et al*. Artificial intelligence–powered rapid identification of ST-elevation myocardial infarction via electrocardiogram (ARISE)—a pragmatic randomized controlled trial. NEJM AI. 2024; 1: AIoa2400190.

[22] Boussina A, Shashikumar SP, Malhotra A, Owens RL, El-Kareh R, Longhurst CA, *et al*. Impact of a deep learning sepsis prediction model on quality of care and survival. NPJ Digital Medicine. 2024; 7: 14.

[23] Tierney AA, Gayre G, Hoberman B, Mattern B, Ballesca M, Hannay SBW, *et al*. Ambient artificial intelligence scribes: learnings after 1 year and over 2.5 million uses. NEJM Catalyst Innovations in Care Delivery. 2025; 6: CAT.25.0040.

[24] European Data Protection Board. Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models. 2024. Available at: https://www.edpb.europa.eu/our-work-tools/our-documents/opinion-board-art-64/opinion-282024-certain-data-protection-aspects_en (Accessed: 15 January 2025).

How to cite this article: Elena G. Bignami, Roberto Lanza, Michele Russo, Valentina Bellini. The lethal diamond of AI issue in critical care medicine: Great Power, Great Risk. Signa Vitae. 2026; 22(1): 141-146. doi: 10.22514/sv.2026.006.