



## Review Article

# Artificial Intelligence for Personalized Critical Care



Moein Sabounchi<sup>1</sup>, Bomi Kim<sup>2</sup> and Ankit Sakhuja<sup>1,3,4,5\*</sup> 

<sup>1</sup>Windreich Department of Artificial Intelligence and Human Health, Icahn School of Medicine at Mount Sinai, New York, NY, USA; <sup>2</sup>Department of Nutritional Sciences, Cornell University, Ithaca, NY, USA; <sup>3</sup>Institute for Critical Care Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA; <sup>4</sup>Division of Data Driven and Digital Medicine, Department of Medicine, Icahn School of Medicine at Mount Sinai, New York, NY, USA; <sup>5</sup>Digital and Technology Partners, Mount Sinai Health System, New York, NY, USA

Received: September 16, 2025 | Revised: December 16, 2025 | Accepted: January 29, 2026 | Published online: June 15, 2026

## Abstract

Critical care medicine requires rapid, high-stakes decisions informed by dynamic and complex streams of patient data. Traditional predictive models have shown value in forecasting deterioration and identifying subphenotypes. However, this leaves a critical gap between anticipating adverse outcomes and guiding therapeutic interventions. Achieving true personalization demands moving beyond generalized protocols toward individualized strategies that account for patient heterogeneity and consequences of alternative clinical actions. Emerging methods in prescriptive artificial intelligence, particularly causal machine learning (causal ML) and reinforcement learning (RL), are beginning to bridge this gap. Causal ML enables estimation of individualized treatment effects by addressing confounding and enabling counterfactual reasoning, allowing clinicians to ask whether a specific intervention is likely to help or harm a given patient. RL can generate adaptive treatment policies that evolve with patient state. The objective of this review is to examine how critical care can progress from generalized prediction to true personalization through the development of prescriptive artificial intelligence. The review contributes by (1) surveying the achievements and limitations of current predictive models, (2) detailing how causal ML and RL can generate individualized treatment effects and sequential decision strategies, (3) identifying the major translational, technical, clinical, ethical, and regulatory barriers to implementation, and (4) outlining future pathways such as digital twins and clinician in the loop systems that may enable safe and actionable personalized decision support at the bedside.

## Introduction

Critical care is characterized by the need for rapid, specialized medical interventions informed by constantly evolving patient data. Decisions such as whether to administer additional intravenous fluids, adjust vasopressor dosing,<sup>1</sup> or modify positive end-expiratory pressure (PEEP) can determine whether an unstable patient moves toward recovery or deteriorates further.<sup>2</sup> Clinical decision-making in the intensive care unit (ICU) is characterized by high risk, minimal tolerance for error, and rapidly evolving patient conditions. Each patient presents a unique clinical trajectory; even within the same individual, physiological states can change significantly over short time intervals. Modern ICUs generate vast amounts of heterogeneous data, such as laboratory results, vital

signs, free-text clinical notes, high-frequency physiological waveforms, imaging, and device outputs. Artificial intelligence (AI) has created unprecedented opportunities to harness this data and potentially guide patient management in real time.<sup>3,4</sup>

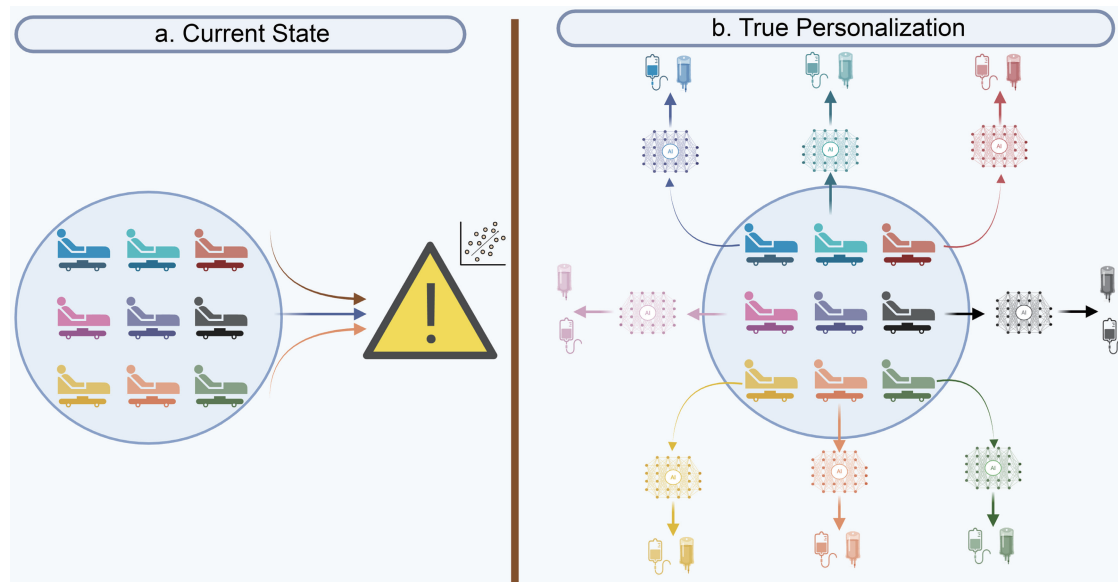
In this context, personalization goes far beyond broad risk prediction or standardized protocols.<sup>5</sup> It entails tailoring therapy to the unique physiological state, comorbidities, and evolving trajectory of each patient. For example, the optimal fluid resuscitation strategy for one patient with septic shock may differ dramatically from another,<sup>6</sup> depending on cardiac function, vascular tone, prior fluid balance, and even biomarker or genomic profiles. Similarly, ventilator management in acute respiratory distress syndrome (ARDS) is not a fixed recipe but a dynamic decision that balances oxygenation, lung protection, and hemodynamic stability, with careful consideration of patient-specific factors.<sup>7</sup> To achieve this level of nuance, clinicians need tools that can move from population-level averages to individualized estimates of treatment response tools that can anticipate not only the risks a patient faces but the consequences of clinical actions for that particular patient at that particular time.

AI provides the technical foundation for enabling such personalized care.<sup>5,8</sup> By leveraging multimodal ICU data streams, predictive models can forecast deterioration while causal machine learning

**Keywords:** Machine learning; Causal machine learning; Reinforcement learning; Digital twins; Personalization; Critical care.

\***Correspondence to:** Ankit Sakhuja, Windreich Department of Artificial Intelligence and Human Health, Icahn School of Medicine at Mount Sinai, New York, NY 10029, USA. ORCID: <https://orcid.org/0000-0002-2045-518X>. Tel: +1-212-241-6500, E-mail: [ankit.sakhuja@mssm.edu](mailto:ankit.sakhuja@mssm.edu)

**How to cite this article:** Sabounchi M, Kim B, Sakhuja A. Artificial Intelligence for Personalized Critical Care. *J Transl Crit Care Med* 2026;8(2):e00023. doi: 10.14218/JTCCM.2025.00023.



**Fig. 1. Artificial intelligence (AI)-driven transformation of critical care: From current state to true personalization.** (a) Predictive AI approaches that forecast outcomes using population-level or average treatment effects. (b) Prescriptive AI approaches that integrate causal inference and reinforcement learning to recommend personalized treatment strategies for each patient. The figure was created using BioRender.

(ML) approaches can estimate the likely impact of specific interventions.<sup>9,10</sup> Reinforcement learning (RL) extends this further by optimizing sequences of decisions across time to maximize long-term patient outcomes.<sup>11</sup> Together, these approaches enable AI systems to move beyond static prediction toward dynamic, adaptive decision support that accounts for patient heterogeneity, temporal evolution, and counterfactual reasoning.<sup>12</sup> In doing so, AI can complement established clinical decision-making by enhancing the ability of clinicians to deliver proactive and individualized therapy.<sup>8</sup>

This review examines how such a transformation might unfold, contrasting current predictive AI approaches that rely on population-level or average treatment effects with prescriptive AI methods that model the causal impact of interventions and their temporal sequencing to enable patient-specific decision making (Fig. 1). By integrating causal inference and reinforcement learning, prescriptive AI recommends patient-specific treatment strategies that adapt to evolving physiology, thus enabling true personalization in critical care. We begin by surveying the current state of AI in critical care, highlighting both the achievements and limitations of existing predictive models. Then, we turn to the question of what it means to achieve true personalization, focusing on RL and causal ML as complementary methods capable of estimating individualized treatment effects and optimizing sequential decision-making. Next, we consider the major challenges that arise in implementing these methods in real-world ICUs, spanning methodological, technical, clinical, ethical, and regulatory domains. Finally, we look ahead to future directions—digital twins,<sup>13</sup> causality-aware foundation models, and clinician-in-the-loop systems<sup>14</sup>—before concluding with a vision for how AI can help deliver on the long-standing promise of critical care of providing the right treatment, for the right patient, at the right time.

### Current state of AI use in critical care

In recent years, AI has begun to make inroads into the practice of critical care, although most applications to date have focused

on prediction rather than providing direct guidance in clinical action. The earliest and most widespread examples are early warning systems that forecast patient deterioration.<sup>15–17</sup> By analyzing temporal trends in vital signs, laboratory data, and clinical documentation, these models provide advance notice of events that might otherwise become apparent only after significant physiological decline. Parallel lines of work have focused on predicting specific outcomes—such as acute kidney injury (AKI),<sup>18</sup> sepsis,<sup>6</sup> or mortality<sup>15,16</sup>—thus offering risk stratification that can support clinical vigilance and guide the allocation of resources.

AI techniques have been utilized to stratify critically ill patients into subpopulations, thus revealing hidden patterns within syndromes traditionally treated as uniform. AI models have revealed subphenotypes of sepsis, AKI, and acute respiratory distress syndrome,<sup>7</sup> thus highlighting biologically and clinically distinct patient clusters. These discoveries are valuable both at the bedside, where they may help clinicians tailor therapies to individual physiology, and in the design and interpretation of clinical trials, where failure to account for heterogeneity has often obscured actual treatment effects.

The increasing sophistication of ML methods has also made it possible to integrate the ICU's diverse data sources into shared representations of patient state.<sup>19–22</sup> Rather than relying on a limited set of vital signs or laboratory values, contemporary models can integrate data obtained from structured electronic health records,<sup>23</sup> high-frequency physiological waveforms,<sup>24</sup> imaging studies, and even free-text notes. Time-series encoders—such as recurrent neural networks,<sup>25</sup> temporal convolutional networks,<sup>26</sup> and transformers<sup>27</sup>—have been applied to handle the irregular sampling and sparse dynamics of ICU data.<sup>22</sup> By leveraging clinical documentation, large language models (LLMs) can capture nuanced information regarding comorbidities, goals of care, and narrative context,<sup>28,29</sup> which are often missed in structured electronic health record data.<sup>30</sup> Meanwhile, vision-based networks trained on chest radiographs or computed tomography scans have been combined with clinical and physiological embeddings,<sup>31</sup> thus creating multi-

modal models that more closely approximate the manner in which intensivists synthesize information in practice.<sup>22,32,33</sup>

Despite this progress, most current systems remain predictive rather than prescriptive. While they estimate the likelihood of outcomes, they stop short of recommending specific interventions to alter those outcomes. A model may forecast that a patient has a high probability of requiring mechanical ventilation within six hours, but it does not advise whether to escalate noninvasive support, adjust sedation, or initiate prone positioning to change that trajectory. Similarly, a system that predicts impending hypotension cannot independently determine whether fluids, vasopressors, or both are most appropriate. This gap between prediction and action emphasizes a central limitation of current AI in critical care and points toward the need for approaches that explicitly model causality and sequential decision-making. Only by moving in this direction can AI evolve from a passive risk calculator into an active partner in delivering personalized critical care.

### Achieving true personalization

The promise of personalized critical care is to move from generalized prediction to individualized prescription. Rather than merely anticipating adverse outcomes, true personalization requires understanding how specific interventions will affect specific patients at specific times and then using this information to guide the sequences of clinical decisions as the illness evolves. Two complementary families of methods, causal ML and RL, provide the methodological foundation for this transition to prescriptive or navigational AI.

#### Causal ML

Causal ML estimates the heterogeneous effects of interventions rather than simple associations between covariates and outcomes. It does so by defining explicit causal estimands—such as the average treatment effect, the individualized treatment effect, and the value of an individualized treatment rule—which correspond to questions regarding how a particular intervention would change outcomes for a given patient or population.<sup>5,34,35</sup> In critical care, treatments are rarely randomly applied; clinicians select interventions based on evolving patient states. This introduces confounding, as the sickest patients may be more likely to receive a certain treatment, thus making it difficult to distinguish the effect of the intervention from the underlying severity of the illness. The credibility of these estimands relies on certain assumptions—such as conditional exchangeability, positivity, consistency, and appropriate model specification—all of which warrant explicit consideration in ICU data. In practice, these assumptions are commonly strained by unmeasured confounding, limited overlap produced by entrenched treatment patterns, time varying confounding as physiology evolves, coarse timing of interventions, and measurement error inherent in routine clinical documentation. Causal ML methods aim to correct this by explicitly modeling the data-generating process and estimating individualized treatment effects. Approaches such as propensity matching and inverse probability weighting estimation enable researchers to predict what outcomes would have occurred under different treatment strategies.<sup>34–37</sup> These challenges can be diagnosed or mitigated using overlap and balance assessments<sup>34</sup>, quantitative sensitivity analyses<sup>38</sup>; negative controls<sup>39</sup>; instrumental variable methods<sup>40</sup>; and longitudinal causal methods, such as marginal structural models or g-formula estimators to address time varying confounding.<sup>41,42</sup> Causal ML methods, such as causal forests and meta-learners,<sup>43,44</sup> have been utilized to identify heterogeneity in treatment response across pa-

tients, effectively moving from average treatment effects to conditional treatment effects. In addition, targeted learning frameworks and longitudinal g-methods also provide principled,<sup>45,46</sup> theory-based approaches for estimating individualized effects in complex, high-dimensional clinical settings. Applied to the ICU, these techniques can help answer questions such as whether additional fluid resuscitation is likely to help or harm a specific patient in septic shock or whether a higher PEEP strategy will improve oxygenation without worsening hemodynamics in a patient with ARDS. By grounding predictions in counterfactual reasoning, causal ML provides individualized estimates that go beyond traditional predictive models and support patient-specific treatment decisions. For example, while classical predictive models can forecast the likelihood of AKI or fluid overload,<sup>47,48</sup> they cannot tell us whether giving or withholding fluids will improve outcomes for a specific patient. In a recent study from our group, causal ML was used to tackle precisely this problem in septic patients with AKI.<sup>49</sup> By leveraging causal forests to estimate individualized treatment effects and applying a policy tree to make those effects interpretable, the study identified subgroups of patients most likely to benefit from a restrictive fluid strategy. In both development and external validation cohorts, those predicted to benefit and those who actually received restrictive fluids had higher rates of AKI reversal and fewer adverse kidney events. This work exemplifies how causal ML can move beyond population-level associations to individualized, counterfactual predictions that inform treatment strategies tailored to the patient in front of us. While promising, this analysis remains observational and, thus, susceptible to residual confounding, thus emphasizing the need for prospective evaluation before such treatment policies are applied in clinical practice.

#### RL

While causal ML addresses the effect of a single intervention, RL extends the framework to optimize sequences of interventions over time.<sup>50</sup> Critical care is inherently dynamic in situations where insulin infusions are titrated hourly, ventilator settings are adjusted as lung mechanics evolve, and fluid/vasopressor balance is revisited with every lab and vital sign update. Each decision influences both the immediate physiology and trajectory of the illness. RL formalizes this process as a Markov decision process,<sup>51</sup> comprising the following core elements:

#### States (s)

The set of all possible patient conditions at a given time, which may include demographics, comorbidities, vital signs, labs, mechanical ventilation parameters, and medications. As ICU data provide only a partial view of the underlying physiological state, representation learning plays a central role in offline RL. Deep sequence models—such as recurrent networks, temporal convolutional networks, and Transformers, as well as multimodal encoders that integrate labs, vitals, waveform data, imaging, and clinical text—can help recover latent patient trajectories from noisy, irregularly sampled observations.<sup>27,52–54</sup> Contrastive learning can further improve robustness to missingness and sensor dropout.<sup>55</sup> Once deployed, these learned state representations must be monitored for drift, for example by tracking shifts in embedding distributions, KL divergence or cosine similarity relative to training distributions, model performance on stability anchors, or abrupt changes in representation clustering that may reflect evolving practice patterns or patient mix.<sup>56</sup> Such monitoring is essential because changes in the representation space can invalidate the learned policy even when raw input features appear stable.

**Actions (a)**

This includes the set of all possible interventions available to the clinician—for example, giving intravenous fluids, vasopressors, insulin, or adjusting PEEP.

**Transitions (T)**

This includes the transition probabilities that map a given state and action to the next state, thus reflecting the patient's physiological response and random variability.

**Rewards (r)**

These are the immediate benefits or costs associated in moving from one state to the other as a result of a specific outcome. Reward functions may incorporate clinical outcomes or physiological targets.<sup>57</sup> In practice, clinically useful rewards are often multiobjective and must capture trade-offs—such as hemodynamic stability versus renal safety, oxygenation targets versus ventilator induced lung injury, and tight glycemic control versus hypoglycemia—with explicit safety constraints that prohibit clearly harmful actions regardless of short term reward.<sup>57–59</sup>

The goal of RL is to learn a sequence of actions (the treatment policy) for given states that can maximize the expected cumulative rewards. The logic of RL naturally resonates with the practice of intensive care medicine. Clinicians do not simply make one decision at the onset of illness—they orchestrate a series of decisions, each informed by prior responses and each shaping future possibilities. For example, a fluid bolus now may increase the probability of pulmonary edema later, which may influence the decision to intubate; in turn, this then changes the trajectory of weaning and sedation. What distinguishes RL from conventional prediction is precisely this temporal chaining, where RL recognizes that the best action is not always the one that maximizes immediate physiologic improvement but the one that sets the patient on the best overall path. In this sense, personalization emerges not only from accounting for individual baseline characteristics but also from dynamically adapting to how that individual responds to prior care.

However, in practice, learning such policies in medicine is constrained by the impossibility of trial-and-error experimentation on critically ill patients. Unlike games or robotics, where RL agents can interact with a simulated environment millions of times,<sup>60,61</sup> in health care, the environment is real patients, and the exploration of untested actions carries unacceptable risks such as adverse events and mortality. Consequently, almost all applications of RL in critical care adopt an offline paradigm. Offline RL learns policies from retrospective data collected under historical clinician practice,<sup>62</sup> without active experimentation. This makes it well suited to medicine, where vast repositories of electronic health records and ICU databases provide the observational trajectories of states, actions, and outcomes for training. The challenge then becomes to extract, from these imperfect and biased records of human practice, a policy that generalizes beyond what clinicians happened to do in certain situations. Offline RL provides a framework to estimate that policy.<sup>63</sup> As policies are learned without prospective exploration, careful off-policy evaluation is essential, typically combining importance-sampling-based estimators, doubly robust methods, and fitted Q evaluation to estimate policy value, along with high confidence bounds that quantify uncertainty before any bedside use.<sup>64</sup>

This framework has already been applied in early proof-of-concept studies. For example, RL has been used to identify optimal fluids and vasopressor doses in patients with sepsis and for ventilator management.<sup>59,65</sup> More recently, offline RL has been applied to improve glycemic control among critically ill patients

after cardiac surgery.<sup>57</sup> Importantly, this RL model underwent multiphase human validations, thus demonstrating that its recommendations were at least as safe, accurate, and acceptable as those of experienced clinicians. Nevertheless, these studies are largely retrospective, and prospective trials are needed to establish safety, effectiveness, and generalizability in real-world ICU practice. These studies highlight how RL can move beyond one-size-fits-all guidelines by tailoring sequences of interventions to the evolving characteristics of individual patients. RL can also be useful for many other high-impact decisions clinicians make daily in the ICU. In sedation and analgesia management, RL could help with titrating or switching sedative or analgesic agents, with guardrails informed by hemodynamic stability, respiratory drive, and delirium prevention. Decisions regarding antibiotic initiation and de-escalation—including choice, timing, and duration—are similarly sequential and context-dependent, constrained by hemodynamic instability, organ dysfunction, evolving microbiologic data, and stewardship considerations, thus making them another domain where RL could be helpful. RL could also inform many additional complex decision processes, such as transfusion of blood products, initiation and adjustment of anticoagulants, delivery and titration of nutritional support, initiation and dosing of dialysis modalities (including continuous renal replacement therapy), and the titration of extracorporeal membrane oxygenation and other forms of mechanical circulatory or respiratory support. These high-stakes decisions share a common structure in that they require balancing competing physiological priorities under uncertainty, adapting actions over time as a patient's condition evolves, and respecting explicit safety constraints—all situations that are well suited to sequential decision-making frameworks such as offline RL.

Across both causal ML and RL, existing studies should be interpreted as hypothesis-generating ones, with prospective evaluation as a prerequisite for clinical deployment.

**Challenges with the implementation of causal ML and RL**

While causal ML and RL models have enormous potential to provide personalized solutions for a wide variety of medical applications, the leap from algorithmic development to clinical implementation involves significant challenges,<sup>50,66</sup> which are explained below.

**Data quality and integration**

One of the first barriers to implementing causal ML and RL in the ICU is the quality and structure of the underlying data. Electronic health records are riddled with missingness, delayed documentation, inconsistencies in units, and artifacts from devices.<sup>23,24</sup> Waveform data may be stored at high frequency but fragmented across vendors; medication administration records often lack precise timestamps or infusion-rate adjustments. For causal ML, this undermines the reliability of confounder adjustment; for RL, it erodes the fidelity of state representations. Therefore, implementation requires investment in robust data engineering pipelines, real-time ingestion, harmonization across systems, and validation of physiological plausibility before these models can even begin to run in clinical practice. An additional challenge is the lack of interoperability across ICU information systems, where heterogeneous data schemas, vendor-specific formats, and limited adherence to standards (such as HL7 FHIR) impede the reliable integration of multimodal datasets.<sup>67,68</sup> When institutions utilize incompatible documentation workflows or nonstandardized device interfaces, even simple features—such as vasopressor dose or ventilator set-

tings—may be represented differently, thereby complicating model training and deployment. Therefore, effective implementation of these AI models requires robust data engineering pipelines, real time data ingestion, harmonization across disparate systems, and physiological validation layers that ensure data plausibility before these models can begin to operate in clinical practice.

### ***Interpretability and clinician trust***

Another challenge is interpretability.<sup>69</sup> Causal ML may estimate individualized treatment effects and RL may recommend sequences of actions, but unless the rationale underlying these recommendations can be explained, clinicians are unlikely to trust or adopt them. In practice, intensivists need to know not only what the model suggests but also the “why.” Which clinical features are driving the recommendation, which counterfactual scenarios were considered, and which uncertainties remain. Translating complex algorithms into intuitive explanations is essential. Without this, models risk being perceived as “black boxes”,<sup>70</sup> thus leading to skepticism or rejection at the bedside. Recent work in explainable AI, counterfactual reasoning, and human-centered design provides tools to bridge this gap, but these approaches remain under explored in high acuity settings such as the ICU.<sup>71</sup>

### ***Workflow integration and human factors***

Critical care workflows are fast-paced and team-based, with decisions often made under severe time pressure. Implementing causal ML or RL systems requires more than algorithmic accuracy. It requires seamless integration into these workflows. A system that issues alerts or recommendations at inconvenient times or in formats disconnected from the workflow of end-users risks adding cognitive burden rather than alleviating it.<sup>72</sup> Furthermore, ICUs function through multidisciplinary collaboration. Thus, a recommendation made to a bedside nurse, a fellow, or an attending must fit into the communication patterns of the team. Therefore, effective implementation requires codesign with clinicians to ensure that the recommendations are timely, context-aware, and aligned with existing decision pathways, rather than disruptive. Thus, it is essential to incorporate human-centered design and ensure usability testing during the development and deployment of AI models. After deployment, the recommendations should surface within the natural flow of team-based ICU activities rather than through intrusive pop-up alerts. Moreover, clinicians must retain full authority over treatment decisions, with AI systems providing suggestions or risk estimates that can be accepted, modified, or overridden. Override mechanisms should be straightforward, encouraged, and automatically logged to create an auditable record that supports transparency and iterative model refinement. To minimize alert fatigue, thresholds for when to display recommendations should be carefully tuned and routinely monitored using metrics such as alert frequency, acceptance rates, and downstream clinical actions. Explanations should remain concise and clinically meaningful, highlighting the patient features and tradeoffs that drove a recommendation, to ensure that clinicians can rapidly judge appropriateness during time-based decision-making.

### ***Technical integration, privacy, and security***

These systems will need to interface with electronic health records and monitoring systems to ingest data and return outputs in real time. Standards such as FHIR and SMART on FHIR provide a practical basis for interoperable integration of real-time clinical data and AI-driven recommendations into the bedside record.<sup>68</sup> Privacy and security safeguards—including strong authentication,

role-based access control, and audit logging—are critical given the sensitivity of ICU data. Where possible, data processing should adhere to data-minimization principles and institutional policies should specify how logs, overrides, and model outputs are stored and accessed.

### ***Prospective validation and evaluation***

Prospective evaluation remains one of the most significant barriers to translating causal ML and RL into real-world critical care. Similar to traditional risk prediction models, RL and causal ML cannot be entirely validated by retrospective accuracy metrics alone. In RL, various off-policy evaluation techniques—such as fitted Q-evaluation (FQE),<sup>73</sup> weighted importance sampling (WIS),<sup>74</sup> and, more recently, DICE<sup>75</sup>—have been developed to approximate how a learned policy might perform in practice.

Although these methods provide essential safeguards and enable the identification of unsafe or unstable policies during development, they cannot fully anticipate the behavior of models deployed in dynamic clinical environments. Therefore, prospective validation is critical, but designing such studies poses ethical and methodological challenges. Randomized controlled trials are resource-intensive and may be difficult to justify when a model proposes actions that diverge from accepted clinical practice. Emerging strategies—such as silent deployment, where recommendations are generated but withheld from clinicians—provide a lower risk pathway for assessing reliability, stability, and usability of policies before they are actively integrated. Further, simulation-based evaluation environments, such as digital twins, should be explored as a means of stress testing policies under controlled conditions without exposing patients to harm.<sup>76</sup> However, these approaches remain technically demanding and have not yet been implemented at scale. Consequently, the field still lacks standardized prospective evaluation frameworks capable of ensuring that prescriptive AI systems can be safely deployed in high-acuity settings.

### ***Documentation, transparency, and continuous monitoring***

Transparent documentation is essential for safe deployment of policies. Model cards and similar standardized summaries can specify the model’s intended use, provide training data, measure performance across subgroups, work with known limitations, and ensure appropriate monitoring.<sup>77</sup> Moreover, data provenance and versioning should be tracked so every model prediction can be linked to its underlying code, parameters, and data snapshots. Continuous monitoring for dataset shifts—using changes in input distributions, calibration, or outcome frequencies—can identify when retraining or recalibration is needed.<sup>78</sup>

### ***Governance, regulation, and liability***

Bringing RL and causal ML into the ICU also raises issues of liability and governance.<sup>79</sup> Unlike static models, RL-based policies may evolve over time as more data are ingested, thereby complicating regulatory oversight. It is important for hospitals and regulators to determine the level of autonomy these systems can have, who bears the responsibility for adverse outcomes, and how updates to models are controlled. Further, causal ML introduces additional questions, such as, if individualized treatment effect estimates differ from guideline-recommended care, who decides whether to follow the model or the guideline? As these systems mature, collaboration with regulatory bodies will be essential to ensuring safety and public trust. Regulatory and governance frameworks, such as Good Machine Learning Practice and guidance for software as a medical device (SaMD), emphasize clear intended use, data qual-

ity, risk management, and a prespecified change control plan for adaptive models.<sup>80</sup> Under both Food and Drug Administration and European Union Medical Device Regulation approaches to SaMD, sponsors are expected to define how models will be updated, how performance will be monitored post-market, and how evidence will be generated when substantial changes are introduced.<sup>80–82</sup> Who will assume responsibility when model recommendations diverge from guidelines should be explicitly specified in institutional governance, with clear documentation regarding how recommendations are generated, when they may be safely ignored, and how conflicts are resolved. Clear governance frameworks, change-control protocols, and legal clarity must be the prerequisites for implementation.

### **Fairness and equity in real-world deployment**

Finally, there is the issue of fairness.<sup>83,84</sup> Both causal ML and RL are only as good as the data they are trained on, and historical ICU data often reflect inequities in care delivery across race, sex, geography, and socioeconomic status. A model trained on such data may learn policies that inadvertently perpetuate disparities—for example, recommending fewer interventions in patients from groups that historically received less aggressive care. In research settings, subgroup analyses can identify such patterns, but in implementation, continuous auditing and fairness-aware retraining will be required. Without explicit attention to equity, deployment risks the widening of gaps rather than their narrowing in critical care outcomes.

### **Future directions**

The next stage in the development of AI for personalized critical care lies not in further demonstration of feasibility but in building systems that can be trusted, validated, and safely deployed at the bedside. For causal ML, future research will need to move beyond exploratory analyses of heterogeneity toward methods that produce clinically reliable treatment effect estimates. One important direction is the use of target trial emulation frameworks,<sup>85</sup> where observational ICU data are explicitly structured as though they were randomized trials. This approach strengthens causal validity and provides estimates that are more easily aligned with clinical reasoning. Advances in methods that address time-varying confounding will also be crucial, as patient physiology and treatment decisions interact in feedback loops over hours and days. Further, high-dimensional extensions of g-methods and doubly robust estimators could enable more faithful estimations of individualized treatment effects in these longitudinal settings.<sup>86,87</sup> In addition, causal ML studies should prespecify estimands,<sup>88</sup> articulate assumptions using directed acyclic graphs,<sup>89</sup> report positivity and overlap diagnostics, and conduct quantitative sensitivity analyses for unmeasured confounding. In addition, external validation together with explicit assessment of transportability across hospitals and health systems should become standard practice. Ensuring that causal ML tools generate outputs that clinicians can interpret and apply is equally important. Instead of abstract effect estimates, these tools will need to provide clear narratives, such as explaining that “given this patient’s fluid balance, urine output, and hemodynamic profile, a restrictive fluid strategy is likely to improve renal recovery.” Our group’s recent work on fluid management in septic patients with AKI illustrates this trajectory, where individualized treatment effects estimated using causal ML were externally validated and shown to identify those subgroups that would be the most likely to benefit from a restrictive approach.<sup>49</sup> Expanding this

paradigm to other decisions within critical care represents an important future path.

In RL, offline RL methods will remain central, since exploration in actual patients is not feasible. In this setting, the learned policy is constrained by the behavior policy that generated the data; thus, adequate coverage of clinically relevant actions is essential to avoid extrapolation to areas of the state action space that were rarely or never visited in practice.<sup>90</sup> Directly estimating the behavior policy and examining action frequency and overlap across patient subgroups provide practical diagnostics for coverage. Further, algorithms such as conservative Q-learning,<sup>91</sup> which penalize unsafe deviations from clinician behavior, and distributional RL, which models the entire distribution of possible outcomes rather than merely averages, will be particularly important for safe policies. Additional conservative offline approaches—such as batch-constrained deep Q learning, behavior regularized actor critic, and implicit Q learning—further limit unsupported extrapolation; thus, providing safeguards against high-risk actions can improve stability in high-stakes clinical settings.<sup>92–94</sup> Another critical direction is reward design. Mortality and duration of hospital stay are too sparse to guide useful policies on their own; thus, multiobjective rewards that balance competing priorities—such as hemodynamic stability versus renal safety, glucose control versus hypoglycemia avoidance—will bring RL policies closer to the trade-offs intensivists make daily. However, the greatest challenge will be prospective validation. To ensure credibility, RL studies should explicitly estimate and report the behavior policy, assess state–action coverage to avoid unsupported extrapolation, and apply multiple off-policy evaluation methods, including importance sampling, doubly robust estimators, fitted Q evaluation, and DICE.<sup>95–99</sup> Ablation studies on state representations should test robustness to missing or noisy modalities. Silent deployment phases (where policies generate recommendations but clinicians are blinded to them), simulation studies, and, ultimately, pragmatic trials will be required to ensure safety and benefit. A practical translational path begins with retrospective model development using prespecified diagnostics and internal validation, which must be followed by rigorous off-policy evaluation with uncertainty bounds and predefined safety thresholds. Thereafter, high-fidelity simulators and digital twins of the ICU environment can then be used to stress test proposed policies and explore clinically important scenarios.<sup>13,100,101</sup> Subsequent silent prospective evaluation would enable teams to compare model recommendations with actual clinician actions and outcomes without influencing care, after which clinician in the loop pilots can introduce recommendations into the workflow with explicit override options, logging, and auditing. It is also important to ensure alignment with Good Machine Learning Practice and FDA guidance for adaptive software.<sup>81</sup> It is only once safety, usability, and fidelity to clinical priorities are demonstrated in these stages that pragmatic trials should be considered to assess impact on patient and system-level outcomes.

A third strand of progress will likely originate from the utilization of digital twins, which are virtual patient simulators that blend mechanistic physiology with real-time data to create individualized models of disease evolution.<sup>13</sup> Digital twins can provide safe and controlled environments to test RL policies, stress-test treatment strategies, derived from causal ML, and explore counterfactual scenarios before implementing interventions in real patients. By synchronizing with live ICU data, a digital twin could project likely trajectories under different interventions, thus providing clinicians both a predictive forecast and an adaptive decision-support tool. Over time, this technology could enable RL systems

to “practice” policies *in silico* before recommending them at the bedside while also providing clinicians a virtual sandbox to query counterfactual scenarios, such as, “what if I reduce PEEP instead of increasing vasopressors?” As both causal ML and RL mature, digital twins may emerge as the bridge that enables safe prospective evaluation and, ultimately, real-world deployment.

Further, to ensure that causal ML and RL studies in critical care are transparent, reproducible, and clinically interpretable, investigators should follow established guidelines. For causal ML, target trial emulation checklists, prespecified estimands, DAGs and adjustment sets, positivity diagnostics, sensitivity analyses, and external validation with transportability assessment should be routinely reported. For RL, behavior-policy estimation, coverage diagnostics, off-policy evaluation with uncertainty, ablations on state representation, and preregistered evaluation protocols should be standard. For prediction components that feed into causal or RL workflows, TRIPOD-AI and PROBAST-AI provide updated guidance for transparent reporting and bias assessment.<sup>102,103</sup> Moreover, early-stage clinical evaluations should be reported in line with DECIDE-AI, while randomized trials involving AI decision support should follow SPIRIT-AI and CONSORT-AI.<sup>104,105</sup> Accordingly, a list of common pitfalls for use of AI in critical care is presented in Box 1.

#### **BOX 1. Common pitfalls for use of AI in critical care**

- Confusing correlation with causation
  - Trusting models without external validation
  - Designing RL rewards that do not reflect clinical reasoning
  - Skipping human-factors testing, override logging, and alert-fatigue monitoring
  - Failing to monitor calibration, drift, and fairness over time
- AI, artificial intelligence; RL, reinforcement learning.

With the careful development and integration of these technologies, the ICU of the future will look very different from that of today. Instead of relying on generic protocols or population-based guidelines, clinicians will likely have access to individualized treatment effect estimates that clarify which interventions are most likely to help or harm the patient in front of them. RL-based systems will provide adaptive recommendations that evolve hour by hour as physiology changes, balancing competing priorities with a view toward long-term recovery. Digital twins will likely run silently alongside patients, projecting trajectories under alternative strategies and providing clinicians a safe environment to test and refine decisions. In such a world, intensivists will not be replaced by algorithms but will be augmented by them, supported by tools that continuously synthesize data, reason across counterfactuals, and anticipate consequences, thereby ensuring that the right decision can be made for each patient at the right time.

#### **Limitations**

This review has several limitations. We conducted a narrative rather than a systematic review, so our selection of studies reflects editorial judgment and may not capture all relevant work or quantify the strength of the underlying evidence. The literature we surveyed on causal ML and RL in critical care remains largely retrospective, observational, and proof-of-concept, and we identified no prospective trials demonstrating that prescriptive AI improves patient outcomes at the bedside. Finally, because this is a rapidly evolving field, the methods, tools, and regulatory frameworks we describe

may change quickly, and we did not formally grade the certainty of the evidence we cite.

#### **Conclusions**

AI has opened the door to a new era in critical care, where decision support can move beyond generic prediction and toward individualized prescription. Causal ML can estimate individualized treatment effects, while RL can optimize sequences of actions over time. Together, these approaches move from population averages towards patient-specific guidance. However, a few significant challenges remain, such as data quality, interpretability, workflow integration, validation, governance, and equity; however, advances such as digital twins offer pathways for safe testing and deployment. Importantly, these systems are intended as decision-support tools, not as autonomous decision-makers. They are designed to inform and contextualize clinical judgment, with clinicians retaining ultimate authority and responsibility for patient care. If responsibly developed, these tools may augment clinical judgment, thereby enabling clinicians to deliver on the core promise of intensive care—the right treatment, for the right patient, at the right time.

#### **Acknowledgments**

None.

#### **Funding**

This work was supported by the National Institutes of Health (NIH) (Grant No. 5K08DK131286 to AS).

#### **Conflict of interest**

AS is a consultant for Roche Diagnostics Corporation. Other authors have no conflicts of interest.

#### **Author contributions**

Conceptualization (MS, AS); data curation (MS); methodology (MS, AS); software (MS); resources (AS); writing—original draft (MS, BK); writing—review & editing (AS); visualization (MS, AS); supervision (AS); project administration (AS); funding acquisition (AS). All authors have approved the final version and publication of the manuscript.

#### **References**

- [1] Overgaard CB, Dzavik V. Inotropes and vasopressors: review of physiology and clinical use in cardiovascular disease. *Circulation* 2008; 118(10):1047–1056. doi:10.1161/CIRCULATIONAHA.107.728840, PMID:18765387.
- [2] Luecke T, Pelosi P. Clinical review: Positive end-expiratory pressure and cardiac output. *Crit Care* 2005;9(6):607–621. doi:10.1186/cc3877, PMID:16356246.
- [3] Hadweh P, Niset A, Salvagno M, Al Barajraji M, El Hadwe S, Taccone FS, *et al*. Machine Learning and Artificial Intelligence in Intensive Care Medicine: Critical Recalibrations from Rule-Based Systems to Frontier Models. *J Clin Med* 2025;14(12):4026. doi:10.3390/jcm14124026, PMID:40565770.
- [4] Choi RY, Coyner AS, Kalpathy-Cramer J, Chiang MF, Campbell JP. Introduction to Machine Learning, Neural Networks, and Deep Learning. *Transl Vis Sci Technol* 2020;9(2):14. doi:10.1167/tvst.9.2.14,

- PMID:32704420.
- [5] Kent DM, Steyerberg E, van Klaveren D. Personalized evidence based medicine: predictive approaches to heterogeneous treatment effects. *BMJ* 2018;363:k4245. doi:10.1136/bmj.k4245, PMID:30530757.
  - [6] Singer M, Deutschman CS, Seymour CW, Shankar-Hari M, Annane D, Bauer M, *et al*. The Third International Consensus Definitions for Sepsis and Septic Shock (Sepsis-3). *JAMA* 2016;315(8):801–810. doi:10.1001/jama.2016.0287, PMID:26903338.
  - [7] Matthay MA, Zemans RL, Zimmerman GA, Arabi YM, Beitler JR, Mercat A, *et al*. Acute respiratory distress syndrome. *Nat Rev Dis Primers* 2019;5(1):18. doi:10.1038/s41572-019-0069-0, PMID:30872586.
  - [8] Desai RJ, Glynn RJ, Solomon SD, Claggett B, Wang SV, Vaduganathan M. Individualized Treatment Effect Prediction with Machine Learning - Salient Considerations. *NEJM Evid* 2024;3(4):EVIDo2300041. doi:10.1056/EVIDo2300041, PMID:38776640.
  - [9] Yang E, Vasishtha R, Dad LQ, Kachnic LA, Hope A, Wang E, *et al*. CAST: Time-Varying Treatment Effects with Application to Chemotherapy and Radiotherapy on Head and Neck Squamous Cell Carcinoma. *arXiv* 2025. doi:10.48550/arXiv.2505.06367.
  - [10] Kaddour J, Lynch A, Liu Q, Kusner MJ, Silva R. Causal machine learning: a survey and open problems. *Found Trends Optim* 2025;9(1-2):1–247. doi:10.1561/24000000052.
  - [11] Liu S, See KC, Ngiam KY, Celi LA, Sun X, Feng M. Reinforcement Learning for Clinical Decision Support in Critical Care: Comprehensive Review. *J Med Internet Res* 2020;22(7):e18477. doi:10.2196/18477, PMID:32706670.
  - [12] Verma S, Boonsanong V, Hoang M, Hines KE, Dickerson JP, Shah C. Counterfactual explanations and algorithmic recourses for machine learning: a review. *ACM Comput Surv* 2024;56(12):1–42. doi:10.1145/3677119.
  - [13] Tao F, Zhang H, Liu A, Nee AYC. Digital twin in industry: state-of-the-art. *IEEE Trans Ind Inform* 2019;15(4):2405–2415. doi:10.1109/TII.2018.2873186.
  - [14] Tang S, Modi A, Sjoding MW, Wiens J. Clinician-in-the-loop decision making: reinforcement learning with near-optimal set-valued policies. *Proc Mach Learn Res* 2020;119:9387–9396.
  - [15] Yu Z, Ashrafi N, Li H, Alaei K, Pishgar M. Prediction of 30-day mortality for ICU patients with Sepsis-3. *BMC Med Inform Decis Mak* 2024;24(1):223. doi:10.1186/s12911-024-02629-6, PMID:39118128.
  - [16] Hou N, Li M, He L, Xie B, Wang L, Zhang R, *et al*. Predicting 30-days mortality for MIMIC-III patients with sepsis-3: a machine learning approach using XGboost. *J Transl Med* 2020;18(1):462. doi:10.1186/s12967-020-02620-5, PMID:33287854.
  - [17] Adams R, Henry KE, Sridharan A, Soleimani H, Zhan A, Rawat N, *et al*. Prospective, multi-site study of patient outcomes after implementation of the TREWS machine learning-based early warning system for sepsis. *Nat Med* 2022;28(7):1455–1460. doi:10.1038/s41591-022-01894-0, PMID:35864252.
  - [18] Bellomo R, Kellum JA, Ronco C. Acute kidney injury. *Lancet* 2012;380(9843):756–766. doi:10.1016/S0140-6736(11)61454-2, PMID:22617274.
  - [19] Zadeh A, Chen M, Poria S, Cambria E, Morency LP. Tensor fusion network for multimodal sentiment analysis. In: Palmer M, Hwa R, Riedel S (eds). *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*; 2017 Sep; Copenhagen, Denmark. Stroudsburg (PA): Association for Computational Linguistics; 2017:1103–1114. doi:10.18653/v1/D17-1115.
  - [20] Guo Z, Li X, Huang H, Guo N, Li Q. Medical image segmentation based on multi-modal convolutional neural network: study on image fusion schemes. 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018); 2018 Apr 4-7; Washington, DC, USA. Piscataway (NJ): IEEE; 2018:903–907. doi:10.1109/ISBI.2018.8363717.
  - [21] Du J, Li W, Lu K, Xiao B. An overview of multi-modal medical image fusion. *Neurocomputing* 2016;215:3–20. doi:10.1016/j.neucom.2015.07.160.
  - [22] Khader F, Kather JN, Müller-Franzes G, Wang T, Han T, Tayebi Arasteh S, *et al*. Medical transformer for multimodal survival prediction in intensive care: integration of imaging and non-imaging data. *Sci Rep* 2023;13(1):10666. doi:10.1038/s41598-023-37835-1, PMID:37393383.
  - [23] Evans RS. Electronic Health Records: Then, Now, and in the Future. *Yearb Med Inform* 2016;25(Suppl 1):S48–S61. doi:10.15265/IYS-2016-s006, PMID:27199197.
  - [24] Goodwin AJ, Eytan D, Greer RW, Mazwi M, Thommandram A, Goodfellow SD, *et al*. A practical approach to storage and retrieval of high-frequency physiological signals. *Physiol Meas* 2020;41(3):035008. doi:10.1088/1361-6579/ab7cb5, PMID:32131060.
  - [25] Medsker LR, Jain LC, editors. *Recurrent Neural Networks: Design and Applications*. Boca Raton (FL): CRC Press; 2001. doi:10.1201/9781003040620.
  - [26] Lea C, Vidal R, Reiter A, Hager GD. Temporal convolutional networks: a unified approach to action segmentation. *Lect Notes Comput Sci* 2016;9915:47–54. doi:10.1007/978-3-319-49409-8\_7.
  - [27] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez AN, *et al*. Attention is all you need. *Adv Neural Inf Process Syst* 2017;30:5998–6008.
  - [28] Thirunavukarasu AJ, Ting DSI, Elangovan K, Gutierrez L, Tan TF, Ting DSW. Large language models in medicine. *Nat Med* 2023;29(8):1930–1940. doi:10.1038/s41591-023-02448-8, PMID:37460753.
  - [29] Yang R, Tan TF, Lu W, Thirunavukarasu AJ, Ting DSW, Liu N. Large language models in health care: Development, applications, and challenges. *Health Care Sci* 2023;2(4):255–263. doi:10.1002/hcs2.61, PMID:38939520.
  - [30] Agaronnik ND, Davis J, Manz CR, Tulsy JA, Lindvall C. Feasibility Study for Using Large Language Models to Identify Goals-of-Care Documentation at Scale in Patients With Advanced Cancer. *JCO Oncol Pract* 2026;22(2):294–305. doi:10.1200/OP-24-00992, PMID:40209123.
  - [31] Bukhari SUK, Khalid SS, Syed A, Shah SSH. The evaluation of convolutional neural network (CNN) for the assessment of chest X-ray of COVID-19 patients. *Ann Clin Anal Med* 2020;11(6):639–642. doi:10.4328/ACAM.20175.
  - [32] Ngiam J, Khosla A, Kim M, Nam J, Lee H, Ng AY. Multimodal deep learning. *Proceedings of the 28th International Conference on Machine Learning (ICML 2011)*; 2011 Jun 28-Jul 2; Bellevue, WA, USA. Madison (WI): Omnipress; 2011:689–696.
  - [33] Cheng J, Sollee J, Hsieh C, Yue H, Vandal N, Shanahan J, *et al*. COVID-19 mortality prediction in the intensive care unit with deep learning based on longitudinal chest X-rays and clinical data. *Eur Radiol* 2022;32(7):4446–4456. doi:10.1007/s00330-022-08588-8, PMID:35184218.
  - [34] Stuart EA. Matching methods for causal inference: A review and a look forward. *Stat Sci* 2010;25(1):1–21. doi:10.1214/09-STS313, PMID:20871802.
  - [35] Rosenbaum PR, Rubin DB. The central role of the propensity score in observational studies for causal effects. *Biometrika* 1983;70(1):41–55. doi:10.1093/biomet/70.1.41.
  - [36] Caliendo M, Kopeinig S. Some practical guidance for the implementation of propensity score matching. *J Econ Surv* 2008;22(1):31–72. doi:10.1111/j.1467-6419.2007.00527.x.
  - [37] Mansournia MA, Altman DG. Inverse probability weighting. *BMJ* 2016;352:i189. doi:10.1136/bmj.i189, PMID:26773001.
  - [38] VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. *Ann Intern Med* 2017;167(4):268–274. doi:10.7326/M16-2607, PMID:28693043.
  - [39] Lipsitch M, Tchetgen Tchetgen E, Cohen T. Negative controls: a tool for detecting confounding and bias in observational studies. *Epidemiology* 2010;21(3):383–388. doi:10.1097/EDE.0b013e3181d61eeb, PMID:20335814.
  - [40] Baiocchi M, Cheng J, Small DS. Instrumental variable methods for causal inference. *Stat Med* 2014;33(13):2297–2340. doi:10.1002/sim.6128, PMID:24599889.
  - [41] Robins JM, Hernán MA, Brumback B. Marginal structural models and causal inference in epidemiology. *Epidemiology* 2000;11(5):550–560. doi:10.1097/00001648-200009000-00011, PMID:10955408.
  - [42] Daniel RM, De Stavola BL, Cousens SN. gformula: estimating causal effects in the presence of time-varying confounding or mediation using the g-computation formula. *Stata J* 2011;11(4):479–517. doi:10.1177/1536867X1201100401.
  - [43] Sverdrup E, Petukhova M, Wager S. Estimating Treatment Effect Heterogeneity in Psychiatry: A Review and Tutorial With Causal Forests. *Int J Methods Psychiatr Res* 2025;34(2):e70015. doi:10.1002/mpr.70015, PMID:40178041.

- [44] Vilalta R, Drissi Y. A perspective view and survey of meta-learning. *Artif Intell Rev* 2002;18(2):77–95. doi:10.1023/A:1019956318069.
- [45] Pang M, Schuster T, Filion KB, Eberg M, Platt RW. Targeted Maximum Likelihood Estimation for Pharmacoepidemiologic Research. *Epidemiology* 2016;27(4):570–577. doi:10.1097/EDE.0000000000000487, PMID:27031037.
- [46] Hernán MA, Robins JM. *Causal Inference: What If*. Boca Raton (FL): Chapman & Hall/CRC; 2020. doi:10.1201/9781315374932.
- [47] Yue S, Li S, Huang X, Liu J, Hou X, Zhao Y, *et al*. Machine learning for the prediction of acute kidney injury in patients with sepsis. *J Transl Med* 2022;20(1):215. doi:10.1186/s12967-022-03364-0, PMID:35562803.
- [48] Zhang Y, Xu D, Gao J, Wang R, Yan K, Liang H, *et al*. Development and validation of a real-time prediction model for acute kidney injury in hospitalized patients. *Nat Commun* 2025;16(1):68. doi:10.1038/s41467-024-55629-5, PMID:39747882.
- [49] Oh W, Takkavatakarn K, Al-Taie Z, Kittrell H, Shawwa K, Gomez H, *et al*. Personalized Fluid Management in Patients With Sepsis and Acute Kidney Injury: A casual Machine Learning Approach. *Crit Care Explor* 2025;7(12):e1354. doi:10.1097/CCE.0000000000001354, PMID:41357343.
- [50] Jayaraman P, Desman J, Sabouchi M, Nadkarni GN, Sakhujia A. A Primer on Reinforcement Learning in Medicine for Clinicians. *NPJ Digit Med* 2024;7(1):337. doi:10.1038/s41746-024-01316-0, PMID:39592855.
- [51] Puterman ML. Markov decision processes. *Handb Oper Res Manag Sci* 1990;2:331–434. doi:10.1016/S0927-0507(05)80172-0.
- [52] Hansen ER, Sagi T, Hose K. Multimodal representation learning for medical analytics - a systematic literature review. *Health Informatics J* 2024;30(4):14604582241290474. doi:10.1177/14604582241290474, PMID:39511041.
- [53] Oh W, Veshtaj M, Sawant A, Agrawal P, Gomez H, Suarez-Farinas M, *et al*. ORAKLE: Optimal Risk prediction for mAKE30 in patients with sepsis associated AKI using deep Learning. *Crit Care* 2025;29(1):212. doi:10.1186/s13054-025-05457-w, PMID:40420108.
- [54] Jangda M, Patel J, Vaid A, Gill J, McCarthy P, Desman J, *et al*. NutriSight: interpretable transformer model for dynamic prediction of underfeeding enteral nutrition in mechanically ventilated patients. *Nat Commun* 2025;16(1):11189. doi:10.1038/s41467-025-66200-1, PMID:41408053.
- [55] Yang Y, Lin Q, Li Z, Wang Y, Liang S, Zhang S, *et al*. View-aware contrastive learning for incomplete tabular data with low-label regimes. *Appl Sci* 2025;15(11):6001. doi:10.3390/app15116001.
- [56] Cheng JY, Goh H, Dogrusoz K, Tuzel O, Azemi E. Subject-aware contrastive learning for biosignals. *arXiv* 2020. doi:10.48550/arXiv.2007.04871.
- [57] Desman JM, Hong ZW, Sabouchi M, Sawant AS, Gill J, Costa AC, *et al*. A distributional reinforcement learning model for optimal glucose control after cardiac surgery. *NPJ Digit Med* 2025;8(1):313. doi:10.1038/s41746-025-01709-9, PMID:40425725.
- [58] Sabouchi M, Desman JM, Amit IS, Oh W, Capone C, Jayaraman P, *et al*. Personalized hemodynamic management using reinforcement learning to prevent persistent acute kidney injury after cardiac surgery. *medRxiv* 2025. doi:10.1101/2025.10.23.25338698, PMID:41282947.
- [59] Komorowski M, Celi LA, Badawi O, Gordon AC, Faisal AA. The Artificial Intelligence Clinician learns optimal treatment strategies for sepsis in intensive care. *Nat Med* 2018;24(11):1716–1720. doi:10.1038/s41591-018-0213-5, PMID:30349085.
- [60] Mnih V, Kavukcuoglu K, Silver D, Graves A, Antonoglou I, Wierstra D, *et al*. Playing Atari with deep reinforcement learning. *arXiv* 2013. doi:10.48550/arXiv.1312.5602.
- [61] Rudin N, Hoeller D, Reist P, Hutter M. Learning to walk in minutes using massively parallel deep reinforcement learning. *Proc Mach Learn Res* 2022;164:91–100.
- [62] Levine S, Kumar A, Tucker G, Fu J. Offline reinforcement learning: tutorial, review, and perspectives on open problems. *arXiv* 2020. doi:10.48550/arXiv.2005.01643.
- [63] Tang S, Wiens J. Model Selection for Offline Reinforcement Learning: Practical Considerations for Healthcare Settings. *Proc Mach Learn Res* 2021;149:2–35. PMID:35702420.
- [64] Uehara M, Shi C, Kallus N. A review of off-policy evaluation in reinforcement learning. *arXiv* 2022. doi:10.48550/arXiv.2212.06355.
- [65] Peine A, Hallawa A, Bickenbach J, Dartmann G, Fazlic LB, Schmeink A, *et al*. Development and validation of a reinforcement learning algorithm to dynamically optimize mechanical ventilation in critical care. *NPJ Digit Med* 2021;4(1):32. doi:10.1038/s41746-021-00388-6, PMID:33608661.
- [66] Sanchez P, Voisey JP, Xia T, Watson HI, O'Neil AQ, Tsafaris SA. Causal machine learning for healthcare and precision medicine. *R Soc Open Sci* 2022;9(8):220638. doi:10.1098/rsos.220638, PMID:35950198.
- [67] Bender D, Sartipi K. HL7 FHIR: an agile and RESTful approach to healthcare information exchange. 2013 IEEE 26th International Symposium on Computer-Based Medical Systems. 2013;326–331. doi:10.1109/CBMS.2013.6627810.
- [68] Mandel JC, Kreda DA, Mandl KD, Kohane IS, Ramoni RB. SMART on FHIR: a standards-based, interoperable apps platform for electronic health records. *J Am Med Inform Assoc* 2016;23(5):899–908. doi:10.1093/jamia/ocv189, PMID:26911829.
- [69] Esser-Skala W, Fortelny N. Reliable interpretability of biology-inspired deep neural networks. *NPJ Syst Biol Appl* 2023;9(1):50. doi:10.1038/s41540-023-00310-8, PMID:37816807.
- [70] Lipton ZC. The mythos of model interpretability: in machine learning, the concept of interpretability is both important and slippery. *Queue* 2018;16(3):31–57. doi:10.1145/3236386.3241340.
- [71] Reddy S. Explainability and artificial intelligence in medicine. *Lancet Digit Health* 2022;4(4):e214–e215. doi:10.1016/S2589-7500(22)00029-2, PMID:35337639.
- [72] Ancker JS, Edwards A, Nosal S, Hauser D, Mauer E, Kaushal R; with the HITEC Investigators. Effects of workload, work complexity, and repeated alerts on alert fatigue in a clinical decision support system. *BMC Med Inform Decis Mak* 2017;17(1):36. doi:10.1186/s12911-017-0430-8, PMID:28395667.
- [73] Le HM, Voloshin C, Yue Y. Batch policy learning under constraints. *Proc Mach Learn Res* 2019;97:3703–3712.
- [74] Thomas P, Theocharous G, Ghavamzadeh M. High-confidence off-policy evaluation. *Proc AAAI Conf Artif Intell* 2015;29(1):3000–3006. doi:10.1609/aaai.v29i1.9541.
- [75] Yang M, Nachum O, Dai B, Li L, Schuurmans D. Off-policy evaluation via the regularized Lagrangian. *Adv Neural Inf Process Syst* 2020;33:6551–6561.
- [76] Zhang K, Zhou HY, Baptista-Hon DT, Gao Y, Liu X, Oermann E, *et al*. Concepts and applications of digital twins in healthcare and medicine. *Patterns (N Y)* 2024;5(8):101028. doi:10.1016/j.patter.2024.101028, PMID:39233690.
- [77] Mitchell M, Wu S, Zaldivar A, Barnes P, Vasserman L, Hutchinson B, *et al*. Model cards for model reporting. *Proceedings of the Conference on Fairness, Accountability, and Transparency*; 2019 Jan 29–31; Atlanta, GA. New York: Association for Computing Machinery; 2019:220–229. doi:10.1145/3287560.3287596.
- [78] Quiñero-Candela J, Sugiyama M, Schwaighofer A, Lawrence ND, editors. *Dataset Shift in Machine Learning*. Cambridge (MA): MIT Press; 2009. doi:10.7551/mitpress/9780262170055.001.0001.
- [79] Price WN II, Gerke S, Cohen IG. Liability for use of artificial intelligence in medicine. In: Solaiman B, Cohen IG (eds). *Research Handbook on Health, AI and the Law*. Cheltenham, UK: Edward Elgar Publishing; 2024:150–166. doi:10.4337/9781802205657.ch09, PMID:40245228.
- [80] U.S. Food and Drug Administration. Proposed regulatory framework for modifications to Artificial Intelligence/Machine Learning (AI/ML)-based Software as a Medical Device (SaMD): discussion paper and request for feedback. Silver Spring (MD): U.S. Food and Drug Administration; 2019. Available from: <https://www.fda.gov/files/medical%20devices/published/US-FDA-Artificial-Intelligence-and-Machine-Learning-Discussion-Paper.pdf>.
- [81] U.S. Food and Drug Administration. *Good Machine Learning Practice for Medical Device Development: Guiding Principles*. Silver Spring (MD): U.S. Food and Drug Administration; 2025. Available from: <https://www.fda.gov/medical-devices/software-medical-device-samd/good-machine-learning-practice-medical-device-development-guiding-principles>.
- [82] European Parliament, Council of the European Union. Regulation (EU) 2017/745 of the European Parliament and of the Council of 5

- April 2017 on medical devices. Official Journal of the European Union. 2017. Available from: <https://eur-lex.europa.eu/eli/reg/2017/745/oj>.
- [83] Barocas S, Hardt M, Narayanan A. *Fairness and Machine Learning: Limitations and Opportunities*. Cambridge (MA): MIT Press; 2023.
- [84] Ueda D, Kakinuma T, Fujita S, Kamagata K, Fushimi Y, Ito R, *et al*. Fairness of artificial intelligence in healthcare: review and recommendations. *Jpn J Radiol* 2024;42(1):3–15. doi:10.1007/s11604-023-01474-3, PMID:37540463.
- [85] Hernán MA, Wang W, Leaf DE. Target Trial Emulation: A Framework for Causal Inference From Observational Data. *JAMA* 2022;328(24):2446–2447. doi:10.1001/jama.2022.21383, PMID:36508210.
- [86] Funk MJ, Westreich D, Wiesen C, Stürmer T, Brookhart MA, Davidian M. Doubly robust estimation of causal effects. *Am J Epidemiol* 2011;173(7):761–767. doi:10.1093/aje/kwq439, PMID:21385832.
- [87] Bang H, Robins JM. Doubly robust estimation in missing data and causal inference models. *Biometrics* 2005;61(4):962–973. doi:10.1111/j.1541-0420.2005.00377.x, PMID:16401269.
- [88] Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *J Educ Psychol* 1974;66(5):688–701. doi:10.1037/h0037350.
- [89] Pearl J. *Causality: Models, Reasoning, and Inference*. 2nd ed. Cambridge: Cambridge University Press; 2009. doi:10.1017/CBO9780511803161.
- [90] Gottesman O, Johansson F, Komorowski M, Faisal A, Sontag D, Doshi-Velez F, *et al*. Guidelines for reinforcement learning in healthcare. *Nat Med* 2019;25(1):16–18. doi:10.1038/s41591-018-0310-5, PMID:30617332.
- [91] Kumar A, Zhou A, Tucker G, Levine S. Conservative Q-learning for offline reinforcement learning. *Adv Neural Inf Process Syst* 2020; 33:1179–1191.
- [92] Fujimoto S, Meger D, Precup D. Off-policy deep reinforcement learning without exploration. *Proc Mach Learn Res* 2019;97:2052–2062.
- [93] Wu Y, Tucker G, Nachum O. Behavior regularized offline reinforcement learning. *arXiv* 2019. doi:10.48550/arXiv.1911.11361.
- [94] Kostrikov I, Nair A, Levine S. Offline reinforcement learning with implicit Q-learning. *arXiv* 2021. doi:10.48550/arXiv.2110.06169.
- [95] Precup D, Sutton RS, Singh S. Eligibility traces for off-policy policy evaluation. In: Langley P (ed). *Proceedings of the Seventeenth International Conference on Machine Learning (ICML 2000)*; 2000 Jun 29-Jul 2; Stanford, CA, USA. San Francisco: Morgan Kaufmann; 2000:759–766.
- [96] Sutton RS, Barto AG. *Reinforcement Learning: An Introduction*. 2nd ed. Cambridge (MA): MIT Press; 2018.
- [97] Jiang N, Li L. Doubly robust off-policy value evaluation for reinforcement learning. *Proc Mach Learn Res* 2016;48:652–661.
- [98] Nachum O, Chow Y, Dai B, Li L. DualDICE: behavior-agnostic estimation of discounted stationary distribution corrections. *Adv Neural Inf Process Syst* 2019;32:2315–2325.
- [99] Antos A, Szepesvári C, Munos R. Fitted Q-iteration in continuous action-space MDPs. *Adv Neural Inf Process Syst* 2007;20:9–16.
- [100] Nadeem M, Kostic S, Dornhöfer M, Weber C, Fathi M. A comprehensive review of digital twin in healthcare in the scope of simulative health-monitoring. *Digit Health* 2025;11:20552076241304078. doi:10.1177/20552076241304078, PMID:39777066.
- [101] Halpern GA, Nemet M, Gowda DM, Kilickaya O, Lal A. Advances and utility of digital twins in critical care and acute care medicine: a narrative review. *J Yeungnam Med Sci* 2025;42:9. doi:10.12701/jyms.2024.01053, PMID:39587767.
- [102] Moons KGM, Damen JAA, Kaul T, Hooft L, Andaur Navarro C, Dhiman P, *et al*. PROBAST+AI: an updated quality, risk of bias, and applicability assessment tool for prediction models using regression or artificial intelligence methods. *BMJ* 2025;388:e082505. doi:10.1136/bmj-2024-082505, PMID:40127903.
- [103] Collins GS, Moons KGM, Dhiman P, Riley RD, Beam AL, Van Calster B, *et al*. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. *BMJ* 2024;385:e078378. doi:10.1136/bmj-2023-078378, PMID:38626948.
- [104] Rivera SC, Liu X, Chan AW, Denniston AK, Calvert MJ, SPIRIT-AI and CONSORT-AI Working Group. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI Extension. *BMJ* 2020;370:m3210. doi:10.1136/bmj.m3210, PMID:32907797.
- [105] Liu X, Cruz Rivera S, Moher D, Calvert MJ, Denniston AK, SPIRIT-AI and CONSORT-AI Working Group. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nat Med* 2020;26(9):1364–1374. doi:10.1038/s41591-020-1034-x, PMID:32908283.