Beyond Baseline Severity: Dynamic Prognostication and Modifiable Determinants of Survival in Invasively Ventilated Critically Ill Patients

Toqa Khaled Baiomy El-Sayed[1], Adel S. Hassan [1], Ashraf E. Sileem [1] , Waleed Mansour [1], Ahmed M. Awadalla [2]

1 Chest Diseases Department, Faculty of Medicine, Zagazig University

2 Anaesthesiology Department, Faculty of Medicine, Zagazig University

Corresponding author: Toqa Khaled Baiomy El-Sayed

Received: 28 October 2024, **Accepted:** 17 November 2024, **Published:** 20 November 2024

Abstract

Background: Mortality in patients receiving invasive mechanical ventilation emerges from interactions among host vulnerability, acute organ dysfunction, ventilator-related exposures, and evolving responses to treatment. Admission severity scores quantify initial risk but may inadequately represent changes occurring during intensive care. Objective: To critically synthesize evidence linking baseline and time-varying determinants of survival, assess potentially modifiable critical care practices, and outline principles for individualized prognostication. Content: Established tools including APACHE II, SAPS, SOFA, and mortality probability models support case-mix adjustment and initial risk assessment, although discrimination, calibration, and portability vary between medical and surgical intensive care populations. Serial changes in circulatory support, oxygenation, lung mechanics, renal function, and neurological responsiveness may refine estimates but require careful timing and interpretation. Randomized and guideline evidence supports lung-protective ventilation, prone positioning in severe acute respiratory distress syndrome, selected escalation therapies, coordinated sedation management, infection prevention, and structured liberation. Observational relationships between mortality and vasopressor exposure, neuromuscular blockade, tracheostomy, or ventilation duration require particular caution because treatment selection and survival time influence apparent associations. Contemporary prediction methods should be evaluated for temporal leakage, missingness, calibration, equity, and clinical impact rather than solely area under the curve. Conclusion: The most defensible approach combines baseline risk assessment with repeated physiological reassessment and evidence-based, individualized treatment. Prospective studies should test whether dynamic risk information improves clinical decisions, recovery, and patient-centered outcomes.

Keywords

acute kidney injury; critical illness; hypomagnesemia; intensive care unit; serum magnesium

Introduction

Invasive mechanical ventilation is indispensable when gas exchange, airway protection, or respiratory muscle performance becomes insufficient to sustain life. Yet a ventilator does not reverse the biological process responsible for critical illness. Survival reflects a sequence of interactions among the precipitating insult, pre-existing physiological reserve, evolving organ dysfunction, ventilator exposure, and response to treatment. A score obtained at admission describes only one point on that sequence. The central clinical challenge is to identify when an initially plausible pathway toward recovery is changing into persistent or irreversible multisystem failure. [1, 2]

Traditional ICU prognostic systems were developed to quantify disease severity, compare populations, and support risk-adjusted evaluation of care. APACHE II, SAPS II, SAPS 3, and later models made critically ill cohorts more comparable, but their predicted risks represent average outcomes in development populations rather than immutable individual destinies. Changes in admission practices, treatments, case mix, and goals of care can alter model performance. Prognostic reasoning is therefore strongest when baseline estimates are updated using information obtained after resuscitation, early ventilation, and reassessment of organ support. [3, 4]

The distinction between a predictor of death and a cause of death is especially important. Advanced age, impaired consciousness, high ventilatory pressure, prolonged ventilation, and vasopressor use may all correlate with mortality; their causal interpretations differ. Some describe vulnerability, some are consequences of illness, and others are exposures that can be modified. Confounding by indication can even make an effective treatment appear harmful because it is preferentially administered to patients with the worst prognosis. The present review organizes evidence around that distinction and relates prognostic models to actions at the bedside. [5, 6]

This narrative review examines host vulnerability, respiratory and extrapulmonary trajectories, ventilator-induced injury, hemodynamics, infection, sedation, liberation, and the methodological requirements for clinically useful dynamic prediction. Particular attention is given to mixed medical and surgical ICU populations, where a shared

dependence on invasive ventilation conceals markedly different pathways to deterioration. The objective is to develop a clinically actionable framework rather than promote a single universal mortality score.

Clinical context and the heterogeneity of ventilated populations

Outcome comparisons across ICUs are difficult because intubation is a treatment decision, not a single diagnosis. An elective postoperative patient may undergo brief ventilation after an otherwise controlled procedure, whereas a patient with pneumonia, septic shock, and acute kidney injury can require simultaneous respiratory, cardiovascular, and renal support. Even within ARDS, pulmonary injury arises from heterogeneous direct and indirect insults. Mortality should therefore be interpreted against indication, admission source, disease severity, baseline function, and the timing of escalation rather than as a property of ventilation itself. [7, 8]

The thesis motivating this review observed high mortality in a cohort drawn from medical and surgical ICUs and differing associations between survival and patient characteristics, organ dysfunction, and established severity scores. These observations are hypothesis-generating; their local rates should not be extrapolated to all ventilated populations. Comparable variation in critical care outcomes has been reported between settings with differing admission thresholds and resources. Apparent differences in crude mortality may reflect referral delay, burden of comorbidity, case mix, discharge policy, or access to rescue therapies rather than differences in treatment quality alone. [9, 10]

Choice of outcome further changes conclusions. ICU mortality excludes deaths occurring after ICU transfer, whereas hospital mortality may be influenced by discharge pathways and available post-acute facilities. Twenty-eight-day and ninety-day mortality provide fixed assessment windows but capture different fractions of late deaths. Ventilator-free days combine survival and duration of ventilation, yet require prespecified handling of death and competing events. Patient-centered outcomes also include cognition, mobility, symptom burden, and quality of life after discharge. Good prognostic research must define its endpoint before choosing variables or fitting a model. [4, 6]

Early mortality following intubation typically stems from extreme disease severity, unmanageable shock, or permanent brain injury. In contrast, later deaths are often driven by hospital-acquired infections, prolonged muscle and tissue breakdown, progressive organ failure, or decisions to withdraw life support though these pathways frequently overlap [2, 6].

Non-modifiable vulnerability and baseline biological reserve

Chronological age is reproducibly associated with adverse critical care outcomes, but age alone incompletely captures biological vulnerability. Two patients of similar age may differ substantially in frailty, exercise capacity, cognition, nutritional reserve, independence, and burden of chronic disease. [1, 5]

Chronic pulmonary disorders influence both the reason for intubation and the hazards of ventilation. In obstructive disease, expiratory flow limitation produces dynamic hyperinflation, intrinsic PEEP, and increased inspiratory threshold load; excessive respiratory rate or inadequate expiratory time may worsen hemodynamics. Restrictive disorders, chronic neuromuscular weakness, and obesity-related hypoventilation create different mechanical and liberation challenges. These phenotypes should not be compressed into one category of chronic respiratory disease when interpreting risk or tailoring ventilator settings. [11, 12]

Cardiac, renal, hepatic, and neurologic comorbidities alter tolerance of acute injury. Congestive heart failure makes liberation vulnerable to increased left ventricular afterload and pulmonary edema; chronic kidney disease reduces capacity to manage fluid, electrolytes, and metabolic changes; liver disease impairs immune competence and coagulation. The relevant issue is often the interaction of chronic and acute organ dysfunction. Baseline creatinine without its earlier value, for example, cannot distinguish chronic impairment from new kidney injury and may obscure a clinically important trajectory. [2, 6]

Sex-associated differences in ICU admission and outcomes have been documented in pooled evidence, but sex-related mortality patterns are not uniform across diagnoses or health systems. Biological mechanisms, selection into intensive care, treatment allocation, and comorbidity can contribute in different proportions. Associations observed in one mixed ICU sample require adequate adjustment and replication before being interpreted as causal. Similarly, residence, occupation, exposure to biomass fuel, or ability to reach tertiary services may function as markers of social disadvantage and delayed referral rather than biological determinants that can be assigned a fixed prognostic weight. [13, 10]

The most useful purpose of baseline stratification is to anticipate problems that may emerge during treatment. Pre-existing frailty suggests the need for rehabilitation planning and careful goals-of-care discussions; difficult airway anatomy changes the hazards of reintubation; advanced obstructive disease requires attention to air trapping and extubation support. These features are relatively fixed, but the consequences of failing to recognize them are potentially modifiable. Predictive precision improves when patient vulnerability is connected to a clinical plan rather than recorded merely as a list of admission comorbidities. [9, 12]

Dynamic physiological deterioration and organ dysfunction

Physiological change following initial treatment is often more informative than the admission reading itself. A

hypotensive patient whose perfusion recovers promptly differs from a patient needing escalating vasopressors despite appropriate source control and resuscitation. Similarly, persistent hypoxemia despite improved lung mechanics has different implications from transient oxygenation failure that resolves with recruitment or removal of excess fluid. Serial trends should be interpreted as responses to known interventions, not as uncontextualized numbers. An improving score may reflect genuine recovery, transient treatment effects, or altered measurement conditions. [2, 6]

The SOFA framework provides a practical language for tracking respiratory, cardiovascular, hepatic, coagulation, renal, and neurologic dysfunction. Sequential measurements can describe whether organ support is stabilizing or escalating, while modified versions may facilitate use where laboratory availability is limited. However, scoring instruments depend on measurement quality and assumptions: sedation confounds neurological assessment, vasopressor dose depends on the care protocol, and changes in creatinine lag behind injury. Applying modified SOFA thresholds developed in one resource environment without validation elsewhere risks systematic misclassification. [2, 14]

Oxygenation ratios, respiratory-system compliance, and ventilatory requirements should be read together. PaO_2/FiO_2 depends on PEEP, inspiratory oxygen, lung recruitment, and sampling conditions. A numerically improving ratio obtained after substantial increases in ventilatory support may not indicate biological improvement. Airway plateau pressure, driving pressure, compliance, minute ventilation, and respiratory effort help contextualize oxygenation. For ARDS, the Berlin classification provides a standardized starting framework; ongoing physiology remains more nuanced than a categorical severity group assigned at one moment. [8, 15]

Hemodynamic measurements also demand integration. Mean arterial pressure alone does not establish adequate tissue perfusion; urine output, mental status, skin findings, lactate trends, and the need for increasing catecholamines contribute information. A falling lactate may be reassuring, but impaired hepatic clearance, adrenergic stimulation, and other mechanisms complicate interpretation. Persistent renal dysfunction, thrombocytopenia, or altered mentation should prompt active search for new or uncontrolled causes. Prognostic value comes from combining consistent signals rather than responding reflexively to an isolated laboratory abnormality. [16, 17]

Clinical prediction faces a technical hazard when dynamic variables are included: information leakage. A model predicting death within seven days cannot legitimately use a complication first detected on day eight, and a decision-support system designed for intubation cannot incorporate information collected after days of treatment. Longitudinal models should specify landmark times, account for discharge and death as competing outcomes when relevant, and update forecasts only as new information becomes available. Without these safeguards, apparent high accuracy may be a consequence of design rather than clinically meaningful prediction. [18, 19]

Time is an overlooked modifier of physiological prognosis. Early improvements may represent treatment-responsive physiology rather than resolution of injury, and some interventions change the measured components of a score without changing the underlying disease. Sedation reduces observable consciousness; dialysis modifies creatinine and biochemical abnormalities; and escalating oxygen or PEEP can transiently improve the arterial oxygen ratio. A responsible interpretation therefore annotates serial scores with the treatment being delivered, the indication for changing that treatment, and whether the same direction of improvement persists when support is cautiously reduced. [2, 6]

Modifiable ventilation-related determinants of outcome

Protective ventilation is a cornerstone of management when the functional lung is reduced and susceptible to stress amplification. The ARDS Network trial established a survival benefit from lower rather than conventional tidal volumes in acute lung injury and ARDS. Current guidance favors approximately 4–8 mL/kg predicted body weight with plateau pressure below 30 cm H_2O in ARDS, recognizing that settings must be adapted to gas exchange, synchrony, respiratory effort, and pressure limits. Predicted rather than actual body weight is used because lung size relates more closely to height and sex than to adipose mass. [11, 20]

Driving pressure offers a physiologically appealing way to integrate tidal volume and respiratory-system compliance. Analyses have demonstrated associations between higher driving pressure and mortality in ARDS, but an observational association is not equivalent to proof that protocolized reductions in driving pressure independently improve survival. Changes in chest-wall mechanics can also influence the relationship between airway pressure and lung stress. The most defensible clinical strategy is to use driving pressure as one signal among tidal volume, plateau pressure, oxygenation, patient effort, and hemodynamic tolerance. [15, 12]

Positive end-expiratory pressure preserves alveolar patency yet can increase overdistension and impede venous return. The balance depends on recruitability, lung heterogeneity, chest-wall pressure, and circulatory reserve. Higher PEEP may be appropriate in moderate-to-severe ARDS under contemporary recommendations, but prolonged recruitment maneuvers are discouraged because they can cause instability and have not demonstrated an acceptable benefit–harm balance. A single PEEP target is unlikely to suit every phenotype. Responses should be checked for improvements in oxygenation and mechanics without deterioration in blood pressure or evidence of overdistension. [20, 21]

The ART trial illustrates why a physiologically plausible intervention may not improve patient outcomes. Aggressive recruitment combined with PEEP titration produced harm in a randomized population with moderate-to-severe ARDS, despite an intervention designed to reopen collapsed units. Transient improvement in an intermediate measure does not guarantee better survival. The broader lesson for ventilatory prognostication is to give precedence to patient-important randomized evidence over mechanistic intuition or changes in surrogate endpoints. The presence of severe hypoxemia should trigger structured assessment and appropriate escalation rather than indiscriminate application of maximal pressures. [22, 20]

Prone positioning alters the distribution of transpulmonary stress, improves ventilation–perfusion relationships, and recruits dependent lung regions while reducing regional overdistension. In severe ARDS, prolonged prone sessions confer a survival benefit when combined with protective ventilation and coordinated care. The magnitude of oxygenation improvement during a session is not, by itself, a validated criterion for selecting survivors or determining treatment continuation. Proning is labor-intensive and requires attention to airway security, pressure injury, vascular access, and hemodynamics, but the evidence supports its use in appropriately selected patients rather than waiting until all other strategies have failed. [23, 24]

Neuromuscular blockade may reduce injurious dyssynchrony and facilitate protective settings in selected patients with severe ARDS. The ACURASYS and ROSE trials differed in sedation strategy, co-interventions, and results, illustrating why routine deep sedation with continuous paralysis should not be assumed beneficial for all ventilated patients. The 2024 ATS guidance offers a conditional recommendation in early severe ARDS, requiring case-by-case assessment. Continuous assessment of ventilator synchrony, patient comfort, depth of sedation, and ability to transition back to spontaneous activity is integral to balancing short-term lung protection against diaphragm disuse and other harms. [25, 26]

Extracorporeal membrane oxygenation can provide rescue gas exchange for selected patients with exceptionally severe, potentially reversible respiratory failure. Its value depends on selection, timing, expertise, complications, and the ability to continue lung-protective ventilation. Evidence from severe ARDS trials and updated professional guidance supports careful referral to experienced centers rather than indiscriminate deployment. Importantly, the need for ECMO is a marker of severe illness; comparing mortality in treated and untreated patients without accounting for eligibility and treatment timing invites severe confounding by indication. [27, 20]

Mechanical ventilation also interacts with circulation. High intrathoracic pressures can decrease venous return and increase right ventricular afterload, particularly when lung overdistension coexists with pulmonary vascular disease or shock. Ventilator adjustments should therefore be assessed through their integrated respiratory and hemodynamic effects. In obstructive lung disease, insufficient expiratory time and intrinsic PEEP may cause preventable cardiovascular compromise even when oxygen saturation initially looks adequate. Management requires physiological interpretation of waveforms and bedside examination, rather than routine escalation of support whenever an arterial blood gas appears abnormal. [12, 21]

Interventions that protect the lung should also be evaluated by their implementation fidelity. Documenting a target tidal volume is insufficient if delivered volumes increase during spontaneous effort or dyssynchrony; an intended plateau-pressure limit may be misleading when measurements are made under inconsistent conditions. Routine review of ventilator waveforms, predicted body weight, and patient–ventilator interaction can reveal avoidable deviations from a prescribed approach. Yet quality improvement should report adherence and clinical outcomes separately, since better documentation does not automatically establish a survival effect. [11, 12]

Sepsis, shock, infection, and systemic treatment exposures

Sepsis frequently drives the interaction between respiratory and extrapulmonary failure. The Sepsis-3 framework identifies life-threatening organ dysfunction attributable to a dysregulated host response to infection and distinguishes septic shock by a more profound circulatory and metabolic disturbance. In an intubated patient, escalating oxygen requirements may result from pneumonia or ARDS, but concomitant acute kidney injury, altered mental status, and circulatory collapse often determine near-term survival. Prediction models that assess only ventilator variables consequently miss key pathways to death. [16, 17]

Rapid recognition of infection, collection of appropriate microbiological specimens, antimicrobial treatment when indicated, source control, and individualized hemodynamic resuscitation are potentially actionable elements of care. Timing studies support the importance of prompt treatment in sepsis, while randomized evidence informs fluid composition and blood pressure targets. At the same time, treatment is not synonymous with increasing every support variable: excess fluid may aggravate lung edema, and high vasopressor targets may increase adverse effects without universal benefit. Clinical judgment must balance perfusion goals against respiratory mechanics and comorbidity. [28, 29]

Vasopressor dose, fluid balance, and renal replacement therapy are especially difficult to interpret as mortality predictors. They may represent the consequence of disease severity, a treatment decision, or both. In studies of critically ill patients, the apparent prognostic association of treatment intensity must be adjusted for why the intervention was started and how the clinical state changed beforehand. Time-varying confounding can make

straightforward regression misleading. These factors should still be recorded because they describe resource use and illness trajectory, but they should not automatically be labeled independent harmful exposures. [17, 30]

Ventilator-associated pneumonia adds infectious burden and often extends the duration of mechanical ventilation. Prevention combines minimizing avoidable ventilation, appropriate oral care and airway management, safe positioning, sedation practices that permit liberation, and attention to infection prevention. Diagnostic uncertainty is substantial: radiographic infiltrates, secretions, fever, and inflammatory markers overlap with noninfectious processes. Treatment should follow credible clinical and microbiological evaluation and local antimicrobial resistance patterns. Unnecessary broad-spectrum therapy may increase toxicity and resistance, whereas delayed active treatment in severely infected patients may have serious consequences. [31, 32]

Clinical studies should separate incident complications from illnesses already present at intubation. An infection arising after prolonged ventilation may predict later death, but patients who die very early have had insufficient time to develop that complication. This creates immortal-time and competing-risk issues if exposure is defined as ever versus never infected. The clinically relevant question is often how prognosis changes once a complication is detected, rather than whether patients who eventually develop it have higher crude mortality. [31, 32]

Sedation, delirium, mobility, and respiratory muscle dysfunction

Sedation and analgesia are necessary for comfort and safe ventilation, but their depth and duration influence neurological assessment, spontaneous effort, delirium, mobility, and time to liberation. Overly deep sedation may conceal neurological deterioration and reduce opportunities to reassess ventilatory independence; inadequate sedation may create distress and severe dyssynchrony. Accordingly, medication exposure should be examined in the context of specific patient needs, including refractory hypoxemia, raised intracranial pressure, or required neuromuscular blockade, rather than judged using a universal target divorced from clinical circumstances. [33, 34]

Paired awakening and breathing assessments established an influential model of coordinated daily reassessment. Their aim is not to withdraw necessary support indiscriminately, but to determine whether ongoing sedation and ventilation remain required. Randomized studies comparing sedative approaches have also demonstrated that the choice of medication affects delirium or sedation-related outcomes, yet advantages are dependent on setting, comparator, and population. An apparent association between deeper sedation and death can be confounded because sicker patients commonly require more intensive treatment. [33, 35]

Guidelines addressing pain, agitation, delirium, immobility, and sleep disruption encourage structured assessment and nonpharmacological strategies rather than a sedation-only approach. Delirium is not equivalent to sedation, and its detection depends on whether the patient is sufficiently arousable to be assessed. Prolonged immobility compounds muscle weakness, pressure injury, and later functional dependency. Early rehabilitation has shown meaningful effects on functional outcomes in selected patients, but feasibility and safety depend on hemodynamic stability, oxygenation, staff resources, and secure devices. [34, 36]

Diaphragm unloading can be both therapeutic and hazardous. Reducing excessive respiratory effort may prevent fatigue and patient self-inflicted lung injury; prolonged near-complete inactivity may accelerate diaphragmatic atrophy. Studies of diaphragm thickness during ventilation have linked pronounced reductions to adverse clinical outcomes, but such associations do not yet establish an optimal universal level of inspiratory effort. Clinicians must consider lung stress, diaphragm activity, gas exchange, and comfort together. Future personalized ventilation may require simultaneous protection of the lung and respiratory muscles rather than optimizing one organ at the expense of the other. [37, 38]

The potential for modifying outcomes through sedation and mobilization depends on systems of care. Reliable analgesia assessments, daily multidisciplinary rounds, protocolized evaluation of eligibility for activity, and documentation of the reasons for not performing liberation assessments are practical process targets. Changes in these processes can be audited more credibly than claims that a particular sedative has independently prevented death in an uncontrolled before-and-after cohort. [39, 34]

Weaning, extubation, and prolonged ventilation trajectories

Liberation from invasive ventilation is a distinct clinical phase requiring recovery of the precipitating illness, adequate respiratory drive and mechanics, manageable secretions, airway protection, and stable circulation. Failure to separate from the ventilator may result from cardiac dysfunction, ongoing infection, neurological impairment, weakness, excessive respiratory load, or an improperly timed attempt. A prolonged duration of ventilation is therefore simultaneously a marker of unresolved disease and a contributor to complications. It should not automatically be interpreted as a modifiable causal determinant of mortality in observational regression models. [40, 41]

The international WEAN SAFE cohort documented substantial variability in weaning practice and outcome across a diverse set of ICUs. Its findings reinforce the need to identify opportunities for earlier structured assessment without assuming that every delay is avoidable. Earlier prospective studies established practical categories of weaning difficulty and demonstrated differences between trials of pressure support, spontaneous breathing, and other approaches. These data support organized clinical decisions while making clear that liberation strategy must reflect

the respiratory load, strength, cardiovascular reserve, and risk of extubation failure in the individual patient. [41, 42]

The 2024 AARC guideline suggests that the rapid shallow breathing index need not be calculated to decide readiness for a spontaneous breathing trial, supports trials with or without pressure support, and encourages a standardized daily assessment. This approach reduces unnecessary barriers to testing readiness and avoids presenting any single bedside index as determinative. Passing an SBT is necessary but may be insufficient: cough effectiveness, airway patency, neurological status, fluid balance, and secretion burden remain important before removal of the endotracheal tube. [43, 44]

Extubation failure exposes patients to complications of delayed reintubation and repeated airway manipulation. Preventive noninvasive ventilation in appropriately selected high-risk patients and high-flow nasal oxygen in selected populations may reduce postextubation respiratory failure or reintubation. These strategies are not interchangeable substitutes for recognizing and treating a failing airway or rapidly progressive disease. Prognostic studies should report planned extubation, unplanned extubation, reintubation, and postextubation rescue separately because the risk mechanisms and clinical implications differ substantially. [45, 46] Preventive support should also be differentiated from rescue support started after respiratory failure has become established, as timing can determine benefit and risk. [47]

Tracheostomy is often considered when prolonged ventilation is anticipated, but studies comparing tracheostomy timing are vulnerable to selection and survivorship bias in observational settings. Patients must remain alive long enough to undergo a later procedure. A lower observed mortality among tracheostomized patients therefore cannot be used as proof that the procedure reduces death. The timing decision should instead incorporate airway needs, anticipated course, comfort, secretion clearance, mobility plans, and shared decisions with the patient or family. [40, 44]

Risk scores: discrimination, calibration, transportability, and clinical use

APACHE II translates acute physiological derangements, age, and chronic health information into an interpretable severity classification. Later APACHE versions, SAPS models, and mortality probability tools expanded or re-estimated variables using larger and more recent populations. SOFA addresses a related but different task by describing the burden of organ dysfunction; its use as a stand-alone mortality estimator requires separate evaluation. In mixed medical and surgical ICUs, performance may change because identical laboratory or physiological values can arise in different contexts and because the probability of death differs across patient groups. [1, 3]

Three questions should be asked before accepting a reported area under the receiver-operating-characteristic curve. First, can the model distinguish patients who experience the outcome from those who do not (discrimination) ? Second, do predicted probabilities correspond to observed risks(calibration)? Third, will the model perform in a hospital with different patients, measurements, treatments, and admission policies(transportability)? A high AUC alone does not demonstrate that counseling, triage, or allocation decisions based on the model are safe. Calibration can deteriorate markedly even when rank ordering remains reasonably stable. [18, 48]

The underlying thesis compared APACHE II, modified SOFA, MPM II, and APPS and found that selected tools performed differently in medical and surgical ICU strata. This is an important local observation but does not demonstrate that each unit should adopt a particular score as its sole prognostic standard. Small subgroup samples and limited death-event counts create instability in adjusted models; comparing AUCs without confidence intervals or formal tests can exaggerate apparent superiority. Local recalibration and external validation should precede widespread deployment of any unit-specific cutoff. [14, 9]

The APPS score was designed around age, plateau pressure, and oxygenation severity in ARDS rather than the full spectrum of neurologic, septic, postoperative, and obstructive indications for intubation. It would be unsurprising if its discrimination changed markedly when applied to a mixed ICU cohort. A domain-specific model can be valuable within its intended population and misleading outside it. Before comparing APPS with general ICU scores, investigators should explicitly define eligibility, calculation time, missing variables, and how the score relates to the studied outcome. [8, 15]

Scores also have ethical limitations. Probability estimates should support risk communication and transparent allocation policy, not serve as automatic instructions to withhold life-sustaining therapy. A prediction reflects outcomes under a particular system of care and may embed disparities in access or treatment limitation. Repeated estimates are particularly vulnerable to self-fulfilling outcomes when they influence treatment withdrawal. Prognostic discussions should disclose uncertainty, separate estimated risk from recommendations, and give clinical goals and patient preferences a defined role. [4, 19]

For bedside deployment, calibration should be shown graphically across the clinically relevant risk range and checked in subgroups defined before analysis, including medical and surgical admissions. Net benefit requires specifying the decision threshold and the action the prediction is expected to trigger. Model updating may be preferable to fitting an entirely new algorithm when local risk differs mainly because of a change in baseline mortality. Prediction intervals, missing-data handling, and subgroup sample sizes should be reported so apparent precision is not confused with certainty. [18, 48]

Toward dynamic, individualized, and trustworthy prognostication

A clinically useful dynamic model should begin with information available at a defined landmark, such as immediately before intubation or during the first ICU day and update its predictions at intervals that can alter care. Inputs might include trajectories of cardiovascular support, oxygenation and respiratory mechanics, renal function, neurological responsiveness, and emerging complications. The time of measurement is as important as the value. For example, a decreasing vasopressor requirement during the first day may indicate recovery, while the same dose reached after progressive escalation suggests a markedly different situation. [6, 18]

Machine-learning approaches can integrate interactions and nonlinear trends that are inconvenient to specify in traditional regression, but complexity alone does not ensure clinical benefit. Missing observations, practice-dependent sampling frequency, treatment confounding, class imbalance, and poor external transportability can all create impressive internal accuracy without bedside usefulness. Transparent feature definitions, temporally correct splits, realistic validation cohorts, and reporting of calibration are necessary whether the method is logistic regression, boosting, or a neural network. Claims of improvement should be tested against strong simple baselines and assessed for decision consequences. [19, 49]

Prediction should be linked to an intervention with a plausible mechanism. A model that alerts clinicians to rising risk could prompt a diagnostic reassessment, review of ventilator synchrony, source-control check, hemodynamic ultrasound, medication review, or discussion of goals of care. Yet alerting is not inherently beneficial: false positives cause alarm fatigue and unnecessary interventions, while false negatives create false reassurance. Decision-curve analysis, workflow studies, and ultimately prospective impact evaluations are needed before algorithmic predictions can be considered improvements in patient outcomes. [18, 50]

Research in resource-constrained ICUs should prioritize measurements that are reliably available and actions that are realistically deployable. A bedside score dependent on costly tests may be less effective than a calibrated model using basic vital signs, urine output, routinely available laboratory results, and serial physical examination. At the same time, parsimony cannot justify omission of clinically essential domains. Local data quality, coding consistency, referral patterns, and treatment limitations must be described because they influence the meaning and portability of every modeled variable. [14, 10]

The next generation of clinical studies should compare a baseline-only model with prespecified serial updates, evaluate calibration at clinically meaningful time horizons, and externally validate across medical and surgical ICUs. Outcomes should incorporate survival and post-ICU function while accounting for competing events. The highest-quality evidence would evaluate whether acting on a dynamic risk estimate improves treatment delivery, patient experience, or recovery rather than simply demonstrating a higher AUC. These design choices transform prediction from statistical description into a testable clinical strategy. [48, 19]

A final priority is to test whether prognostic communication changes practice safely. If an updated prediction leads to earlier rescue escalation, the appropriate endpoint includes both timely intervention and patient outcomes. If it informs discussions with families, investigators should assess comprehension, decisional conflict, and concordance with patient values, while documenting any treatment limitations that may influence observed survival. This is a higher standard than retrospective association and would better define when dynamic scoring adds value beyond experienced multidisciplinary clinical review. [19, 50]

Clinical implications, research priorities, and limitations of the evidence

A practical multidisciplinary approach begins with standardized documentation of indication for intubation, baseline function, major comorbidities, and early organ dysfunction. During each subsequent assessment, the team should revisit pulmonary mechanics, perfusion, infection control, neurological status, medication effects, and readiness for rehabilitation or ventilator liberation. The purpose is not to add another score to an already complex workflow, but to identify mismatches between the current clinical trajectory and the care plan. The same framework can support audit of time-sensitive protective ventilation, prone positioning when indicated, and avoidance of unnecessary sedation or ventilation. [20, 17]

Several limitations shape the current evidence. Many risk associations come from heterogeneous observational cohorts with inconsistent definitions of complications, treatment timing, and mortality. Randomized evidence is strong for some respiratory interventions but cannot resolve every question about combined organ support or baseline prognostic features. Dynamic prediction studies frequently report discrimination without sufficient assessment of calibration and clinical impact. Future multicenter projects should use common outcome definitions, avoid temporal leakage, examine treatment-related confounding, include resource-diverse hospitals, and report performance separately in medical and surgical populations before claiming generalizability. [41, 50]

Conclusion

Survival during invasive mechanical ventilation is shaped by baseline vulnerability and, more importantly for ongoing clinical decisions, by the changing interaction of organ dysfunction, respiratory mechanics, infection, hemodynamics, sedation, and treatment response. Established severity scores remain useful for population description and initial

risk assessment, but none can replace serial examination or demonstrate that a modifiable exposure is causal. A clinically credible approach combines individualized lung- and diaphragm-protective care, timely treatment of systemic complications, coordinated liberation, and repeatedly reassessed prognosis. Future prediction systems should be externally validated and judged by whether they improve decisions and patient-centered outcomes, not by discrimination alone.

References

1. Knaus WA, Draper EA, Wagner DP, Zimmerman JE. APACHE II: a severity of disease classification system. Crit Care Med. 1985;13(10):818-829.
2. Vincent JL, Moreno R, Takala J, et al. The SOFA (Sepsis-related Organ Failure Assessment) score to describe organ dysfunction/failure. Intensive Care Med. 1996;22(7):707-710.
3. Le Gall JR, Lemeshow S, Saulnier F. A new Simplified Acute Physiology Score (SAPS II) based on a European/North American multicenter study. JAMA. 1993;270(24):2957-2963.
4. Moreno RP, Metnitz PGH, Almeida E, et al. SAPS 3 From evaluation of the patient to evaluation of the intensive care unit. Part 2: Development of a prognostic model for hospital mortality at ICU admission. Intensive Care Med. 2005;31:1345-1355.
5. Zimmerman JE, Kramer AA, McNair DS, Malila FM. Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. Crit Care Med. 2006;34(5):1297-1310.
6. Lambden S, Laterre PF, Levy MM, Francois B. The SOFA score development, utility and challenges of accurate assessment in clinical trials. Crit Care. 2019;23:374.
7. Gajic O, Dabbagh O, Park PK, et al. Early identification of patients at risk of acute lung injury: evaluation of lung injury prediction score in a multicenter cohort study. Am J Respir Crit Care Med. 2011;183(4):462-470.
8. Ranieri VM, Rubenfeld GD, Thompson BT, et al. Acute respiratory distress syndrome: the Berlin Definition. JAMA. 2012;307(23):2526-2533.
9. Keegan MT, Gajic O, Afessa B. Comparison of APACHE III, APACHE IV, SAPS 3, and MPM0III and influence of resuscitation status on model performance. Chest. 2012;142(4):851-858.
10. Abate SM, Assen S, Yinges M, Basu B. Survival and predictors of mortality among patients admitted to the intensive care units in southern Ethiopia: a multi-center cohort study. Ann Med Surg (Lond). 2021;65:102318.
11. The ARDS Network. Ventilation with lower tidal volumes as compared with traditional tidal volumes for acute lung injury and the acute respiratory distress syndrome. N Engl J Med. 2000;342(18):1301-1308.
12. Fan E, Del Sorbo L, Goligher EC, et al. An official ATS/ESICM/SCCM clinical practice guideline: mechanical ventilation in adult patients with acute respiratory distress syndrome. Am J Respir Crit Care Med. 2017;195(9):1253-1263.
13. Modra L, Higgins A, Vithanage R, Abeygunawardana V, Bailey M, Bellomo R. Sex differences in illness severity and mortality among adult intensive care patients: a systematic review and meta-analysis. J Crit Care. 2021;65:116-123.
14. Grissom CK, Brown SM, Kuttler KG, et al. A modified sequential organ failure assessment score for critical care triage. Disaster Med Public Health Prep. 2010;4(4):277-284.
15. Amato MBP, Meade MO, Slutsky AS, et al. Driving pressure and survival in the acute respiratory distress syndrome. N Engl J Med. 2015;372(8):747-755.
16. Singer M, Deutschman CS, Seymour CW, et al. The third international consensus definitions for sepsis and septic shock (Sepsis-3). JAMA. 2016;315(8):801-810.
17. Evans L, Rhodes A, Alhazzani W, et al. Surviving Sepsis Campaign: international guidelines for management of sepsis and septic shock 2021. Intensive Care Med. 2021;47:1181-1247.
18. Steyerberg EW, Vickers AJ, Cook NR, et al. Assessing the performance of prediction models: a framework for traditional and novel measures. Epidemiology. 2010;21(1):128-138.
19. Wolff RF, Moons KGM, Riley RD, et al. PROBAST: a tool to assess the risk of bias and applicability of prediction model studies. Ann Intern Med. 2019;170(1):51-58.
20. Qadir N, Sahetya S, Munshi L, et al. An update on management of adult patients with acute respiratory distress syndrome: an official American Thoracic Society clinical practice guideline. Am J Respir Crit Care Med. 2024;209(1):24-36.
21. Brower RG, Lanken PN, MacIntyre N, et al. Higher versus lower positive end-expiratory pressures in patients with the acute respiratory distress syndrome. N Engl J Med. 2004;351(4):327-336.
22. Cavalcanti AB, Suzumura ÉA, Laranjeira LN, et al. Effect of lung recruitment and titrated positive end-expiratory pressure (PEEP) vs low PEEP on mortality in patients with acute respiratory distress syndrome: a randomized clinical trial. JAMA. 2017;318(14):1335-1345.
23. Guérin C, Reignier J, Richard JC, et al. Prone positioning in severe acute respiratory distress syndrome. N Engl J Med. 2013;368(23):2159-2168.
24. Beitler JR, Shaefi S, Montesi SB, et al. Prone positioning reduces mortality from acute respiratory distress syndrome in the low tidal volume era: a meta-analysis. Intensive Care Med. 2014;40:332-341.
25. Papazian L, Forel JM, Gacouin A, et al. Neuromuscular blockers in early acute respiratory distress syndrome. N Engl J Med. 2010;363(12):1107-1116.
26. National Heart, Lung, and Blood Institute PETAL Clinical Trials Network. Early neuromuscular blockade in the acute respiratory distress syndrome. N Engl J Med. 2019;380(21):1997-2008.
27. Combes A, Hajage D, Capellier G, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome. N Engl J Med. 2018;378(21):1965-1975.
28. Seymour CW, Gesten F, Prescott HC, et al. Time to treatment and mortality during mandated emergency care for sepsis. N Engl J Med.

2017;376(23):2235-2244.
29. Semler MW, Self WH, Wanderer JP, et al. Balanced crystalloids versus saline in critically ill adults. N Engl J Med. 2018;378(9):829-839.
30. Asfar P, Meziani F, Hamel JF, et al. High versus low blood-pressure target in patients with septic shock. N Engl J Med. 2014;370(17):1583-1593.
31. Kalil AC, Metersky ML, Klompas M, et al. Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 clinical practice guidelines by the IDSA and ATS. Clin Infect Dis. 2016;63(5):e61-e111.
32. Klompas M, Branson R, Cawcutt K, et al. Strategies to prevent ventilator-associated pneumonia, ventilator-associated events, and nonventilator hospital-acquired pneumonia in acute-care hospitals: 2022 update. Infect Control Hosp Epidemiol. 2022;43(6):687-713.
33. Girard TD, Kress JP, Fuchs BD, et al. Efficacy and safety of a paired sedation and ventilator weaning protocol for mechanically ventilated patients in intensive care (Awakening and Breathing Controlled trial): a randomised controlled trial. Lancet. 2008;371(9607):126-134.
34. Devlin JW, Skrobik Y, Gélinas C, et al. Clinical practice guidelines for the prevention and management of pain, agitation/sedation, delirium, immobility, and sleep disruption in adult ICU patients. Crit Care Med. 2018;46(9):e825-e873.
35. Pandharipande PP, Pun BT, Herr DL, et al. Effect of sedation with dexmedetomidine vs lorazepam on acute brain dysfunction in mechanically ventilated patients: the MENDS randomized controlled trial. JAMA. 2007;298(22):2644-2653.
36. Schweickert WD, Pohlman MC, Pohlman AS, et al. Early physical and occupational therapy in mechanically ventilated, critically ill patients: a randomised controlled trial. Lancet. 2009;373(9678):1874-1882.
37. Jaber S, Jung B, Matecki S, Petrof BJ. Clinical review: ventilator-induced diaphragmatic dysfunction human studies confirm animal model findings! Crit Care. 2011;15:206.
38. Goligher EC, Dres M, Fan E, et al. Mechanical ventilation-induced diaphragm atrophy strongly impacts clinical outcomes. Am J Respir Crit Care Med. 2018;197(2):204-213.
39. Shehabi Y, Howe BD, Bellomo R, et al. Early sedation with dexmedetomidine in critically ill patients. N Engl J Med. 2019;380(26):2506-2517.
40. Beduneau G, Pham T, Schortgen F, et al. Epidemiology of weaning outcome according to a new definition. The WIND study. Am J Respir Crit Care Med. 2017;195(6):772-783.
41. Pham T, Heunks L, Bellani G, et al. Weaning from mechanical ventilation in intensive care units across 50 countries (WEAN SAFE): a multicentre, prospective, observational cohort study. Lancet Respir Med. 2023;11(5):465-476.
42. Esteban A, Frutos F, Tobin MJ, et al. A comparison of four methods of weaning patients from mechanical ventilation. N Engl J Med. 1995;332(6):345-350.
43. Roberts KJ, Goodfellow LT, Battey-Muse CM, et al. AARC clinical practice guideline: spontaneous breathing trials for liberation from adult mechanical ventilation. Respir Care. 2024;69(7):891-901.
44. Ouellette DR, Patel S, Girard TD, et al. Liberation from mechanical ventilation in critically ill adults: an official ATS/ACCP clinical practice guideline. Chest. 2017;151(1):166-180.
45. Hernández G, Vaquero C, González P, et al. Effect of postextubation high-flow nasal cannula vs conventional oxygen therapy on reintubation in low-risk patients: a randomized clinical trial. JAMA. 2016;315(13):1354-1361.
46. Ferrer M, Valencia M, Nicolas JM, Bernadich O, Badia JR, Torres A. Early noninvasive ventilation averts extubation failure in patients at risk: a randomized trial. Am J Respir Crit Care Med. 2006;173(2):164-170.
47. Rubenfeld GD, Caldwell E, Peabody E, et al. Incidence and outcomes of acute lung injury. N Engl J Med. 2005;353(16):1685-1693.
48. Collins GS, Reitsma JB, Altman DG, Moons KGM. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. Ann Intern Med. 2015;162(1):55-63.
49. Komorowski M. Artificial intelligence in intensive care: are we there yet? Intensive Care Med. 2019;45:1298-1300.
50. Collins GS, Moons KGM, Dhiman P, et al. TRIPOD+AI statement: updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ. 2024;385:e078378.