Review Article

# APPLICATION OF MACHINE LEARNING PRINCIPLES IN THE MANAGEMENT OF HEMORRHAGIC SHOCK

***GIORGI KUCHAVA**, PhD;* [1,2] ***IOSEB KARTVELISHVILI**, PhD;* [2] ***MAKA MANTSKAVA**, PhD* [1]

[1] *Ivane Beritashvili Center of Experimental Biomedicine, Tbilisi, Georgia*

[2] *Georgian Technical University, Tbilisi, Georgia*

## ABSTRACT

**Background:** Hemorrhagic shock remains one of the leading causes of preventable mortality in trauma, surgery, and obstetric emergencies. Early recognition, accurate prediction of clinical deterioration, and timely intervention are critical for improving patient outcomes. Advances in artificial intelligence (AI) and machine learning (ML), particularly gradient-boosted ensembles, recurrent neural architectures for physiological time series, and multimodal deep learning, have created new opportunities to support clinical decision-making by continuously analyzing high-dimensional physiological, laboratory, and imaging data streams. Objective: To review, from a technical and methodological standpoint, the machine learning architectures, feature engineering strategies, and validation frameworks applied to diagnosis, risk stratification, monitoring, and management of hemorrhagic shock, with emphasis on model design choices that affect clinical reliability.

**Methods:** A narrative review of contemporary literature was conducted, focusing on supervised, ensemble, and deep learning algorithms applied to trauma care, perioperative medicine, emergency medicine, and critical care. Studies evaluating predictive models based on electronic health records (EHR), continuous vital-sign telemetry, laboratory biomarkers, point-of-care ultrasound/imaging, and waveform-derived hemodynamic features were examined, with attention to model architecture, input feature space, validation methodology, and reported discrimination/calibration metrics.

**Results:** Ensemble tree methods (random forest, XGBoost, LightGBM) and recurrent/temporal architectures (LSTM, GRU, temporal convolutional networks, and transformer-based sequence models) applied to heart rate, blood pressure, lactate trajectory, hemoglobin/hematocrit trends, shock index and its derivatives, base excess, and photoplethysmography or arterial-waveform-derived features have demonstrated AUROC values frequently in the 0.80-0.93 range for predicting need for massive transfusion or hemorrhagic decompensation, generally outperforming threshold-based scores (e.g., ABC score, shock index alone) in retrospective

CITATION: *Kuchava G, Kartvelishvili I, Mantskava M. Application of machine learning principles in the management of hemorrhagic shock. Medical Times. Georgian-German Reproduction Center; 2026;(5):87–92. https://doi.org/10.71419/mtggrc.2026.44*

cohorts. AI-assisted closed-loop and decision-support systems have been explored for titrating fluid resuscitation, blood product ratios, and vasopressor dosing, and for optimizing timing of surgical or endovascular hemorrhage control. Model interpretability (via SHAP, attention weights, or integrated gradients), external validation across heterogeneous populations, class imbalance in rare massive-hemorrhage events, and prospective clinical-workflow integration remain the principal technical barriers to deployment.

**Conclusions:** Machine learning represents a technically mature but clinically unproven adjunct to conventional assessment in hemorrhagic shock. Realizing its benefit requires rigorous external validation, post-deployment calibration monitoring, interpretable model outputs, and prospective multicenter trials demonstrating outcome benefit – not merely discrimination – before routine implementation in emergency and critical care workflows.

**Keywords:** machine learning; deep learning; artificial intelligence; hemorrhagic shock; clinical decision support; predictive analytics; trauma; critical care; massive transfusion; precision medicine; emergency medicine; time-series modeling; model interpretability.

## Introduction

Hemorrhagic shock is a state of inadequate tissue perfusion resulting from acute blood volume loss, and it remains a leading cause of preventable death in trauma, postpartum hemorrhage, and major surgical or gastrointestinal bleeding. Conventional clinical assessment relies on discrete, often delayed indicators heart rate, systolic blood pressure, urine output, and serial hemoglobin measurements that may fail to capture the nonlinear, multivariate physiological trajectory preceding decompensation. Compensatory mechanisms (tachycardia, peripheral vasoconstriction) can mask significant volume loss until a critical threshold is crossed, at which point cardiovascular collapse can occur rapidly. Machine learning offers a fundamentally different paradigm: rather than applying static thresholds to individual variables, ML models can learn joint, time-varying, nonlinear relationships among dozens to hundreds of correlated physiological and laboratory features, continuously updating risk estimates as new data arrive. This review focuses specifically on the technical architecture of these systems the input feature space, model classes, training/validation methodology, and performance characterization rather than solely their clinical framing.[1-5] Pathophysiological Rationale for a Data-Driven Approach: Hemorrhagic shock progresses through compensated, decompensated, and irreversible phases, each characterized by distinct hemodynamic and metabolic signatures: rising shock index (heart rate / systolic BP), narrowing pulse pressure, rising lactate and falling base excess (reflecting anaerobic metabolism), falling central venous oxygen saturation, and coagulopathy (reflected in prolonged PT/INR, falling fibrinogen, and thromboelastography/thromboelastometry parameters such as clot amplitude and lysis index). Because these signals interact – e.g., lactate clearance kinetics are influenced by resuscitation volume, base deficit, and hepatic perfusion simultaneously – the underlying prediction problem is inherently multivariate and time-dependent, favoring ML approaches that model interaction effects and temporal dynamics over univariate threshold rules.[6-9]

*Table 1.* ML Architect and Hemorrhagic shock.

| Model class | Typical use case | Technical notes |
|---|---|---|
| Logistic regression / regularized GLM (LASSO, ridge) | Baseline risk scores, interpretable comparators | Serves as the reference model against which more complex architectures are benchmarked; coefficients are directly interpretable. |
| Random forest / Extra Trees | Massive transfusion prediction, mortality risk | Robust to feature scaling and missingness; provides native feature-importance ranking (mean decrease in impurity) |
| Gradient boosting (XGBoost, LightGBM, CatBoost) | Structured EHR + lab data, tabular risk models | Frequently the top-performing tabular-data architecture; requires careful regularization and early stopping to avoid overfitting on rare-event outcomes |
| Recurrent networks (LSTM, GRU) | Continuous vital-sign streams, early-warning systems | Captures temporal dependencies in sequential vitals; sensitive to irregular sampling intervals, often requires imputation or time-aware embeddings |
| Temporal convolutional networks / transformers | ICU waveform and multivariate time-series forecasting | Attention mechanisms allow post-hoc interpretation of which time points drove a prediction; more data-hungry than RNNs |
| Convolutional neural networks | Point-of-care ultrasound (FAST exam) interpretation, imaging-based bleeding source localization | Applied to identify free fluid, active extravasation, or organ injury from ultrasound/CT |
| Reinforcement learning | Dynamic resuscitation policy optimization (fluid/ vasopressor titration) | Largely investigational; raises significant safety and off-policy evaluation concerns before bedside deployment |
| Source: Spahn DR, Bouillon B, Cerny V, et al. The European guideline on management of major bleeding and coagulopathy following trauma. Sixth edition. Crit Care. 2023;27:80. | | |

## Clinical Application Domains

### Early detection and risk stratification

Models trained on rolling-window vital-sign and lab trends can flag patients trending toward decompensation before conventional early-warning scores (e.g., NEWS2, MEWS) trigger an alert, by learning nonlinear combinations of features rather than applying fixed point thresholds.

### Massive transfusion and blood product requirement prediction

Supervised classifiers combining shock index, base deficit, hemoglobin, FAST exam findings, and injury mechanism have been used to predict need for massive transfusion protocol activation, aiming to reduce both under-triage (delayed blood product availability) and over-triage (unnecessary blood bank strain).

### Resuscitation optimization

Decision-support models integrate dynamic preload-responsiveness indicators (pulse pressure variation, stroke volume variation) with lactate clearance trends to guide fluid versus blood product versus vasopressor sequencing, aiming to avoid both under-resuscitation and dilutional coagulopathy from excess crystalloid.[12-16]

### Imaging-assisted diagnosis

CNN-based image analysis of point-of-care ultrasound and CT angiography has been applied to accelerate identification of active hemorrhage sites, supporting triage to surgical versus endovascular (angioembolization) intervention.

## Model Validation and Performance Characterization

Rigorous evaluation of these models requires reporting beyond raw accuracy, given substantial class imbalance (massive hemorrhage events are comparatively rare): Discrimination: Area under the receiver operating characteristic curve (AUROC) and, given class imbalance, area under the precision-recall curve (AUPRC) are both necessary; AUROC alone can overstate performance on rare-event prediction tasks;

Calibration: Calibration plots and Brier score assess whether predicted probabilities match observed event rates – critical for a model informing transfusion or intervention thresholds;Temporal and external validation: Models should be validated on data from institutions, patient populations, and time periods distinct from training data; performance frequently degrades under external validation, reflecting dataset shift;

Interpretability tooling: SHAP (Shapley Additive Explanations) values, LIME, and attention-weight visualization are increasingly reported to allow clinicians to audit individual predictions rather than treat the model as a black box;

Decision-curve analysis: Net-benefit analysis across a range of clinically plausible decision thresholds is used to determine whether model-guided action would outperform treat-all or treat-none strategies in practice.

## Technical and Implementation Challenges

Class imbalance and rare-event modeling: Massive hemorrhage and mortality outcomes are infrequent, requiring resampling strategies (SMOTE, class-weighted loss functions) or anomaly-detection framings that can introduce their own bias if not carefully validated.Missing and irregularly sampled data: Real-world ICU and trauma-bay data streams have variable sampling frequency and missingness patterns that are often informative (missing-not-at-random), requiring time-aware imputation or masking strategies rather than naive mean imputation. Generalizability and dataset shift: Models trained on a single trauma center's case mix, injury patterns, and resuscitation protocols may not transfer to different populations, resource settings, or blood-product availability contexts – a particularly relevant concern for adaptation in lower-resource health systems.

Interpretability versus performance trade-off: The highest-performing architectures (deep temporal models) are typically the least directly interpretable, creating tension with clinical accountability and regulatory requirements.

Prospective validation gap: The overwhelming majority of published models remain retrospective; few have undergone prospective, randomized evaluation demonstrating that model-guided intervention improves patient-centered outcomes rather than discrimination metrics alone.Regulatory and ethical considerations: Continuous learning systems that update model weights post-deployment raise validation and regulatory challenges (e.g., FDA's evolving framework for adaptive/continuously learning AI/ML-based software as a medical device); patient data privacy, consent for algorithmic decision support, and liability allocation for model-influenced clinical decisions remain unresolved.

Integration into clinical workflow: Alert fatigue from poorly calibrated early-warning systems can erode clinicians' trust; successful deployment requires careful threshold setting, human-in-the-loop design, and integration with existing EHR alerting infrastructure without disrupting workflow.

## Future Directions

Promising directions include federated learning approaches that allow multi-institutional model training without centralizing sensitive patient data, thereby improving generalizability while addressing privacy constraints; multimodal fusion architectures that jointly model structured EHR data, waveform signals, and imaging; and reinforcement-learning-based closed-loop resuscitation systems evaluated first in silico and via retrospective off-policy evaluation before any bedside testing. Prospective, multicenter, ideally randomized trials comparing ML-guided decision support against standard-of-care management – powered on patient-centered outcomes (mortality, transfusion volume, organ dysfunction) rather than discrimination statistics alone – represent the necessary next step before these systems can be considered validated for routine clinical use.

## Conclusion

From a technical standpoint, the machine learning toolkit applicable to hemorrhagic shock has matured substantially: ensemble tree methods and temporal deep learning architectures now consistently outperform static threshold-based scores in retrospective discrimination benchmarks, and interpretability tooling (SHAP, attention visualization) has narrowed – though not eliminated – the black-box concern. The translational gap lies not in model architecture but in external validation, calibration robustness under dataset shift, and, most critically, the near-absence of prospective outcome-based trials. Multicenter prospective validation and careful clinical-workflow integration remain prerequisites for moving these systems from promising retrospective analytics to trusted components of hemorrhagic shock management.

## Declarations

**Conflict of Interest:** The authors declare no competing interests.

**Acknowledgments:** This work was supported by the Shota Rustaveli National Foundation of Georgia (FR-24-389).

## References:

1. Liu NT, Salinas J. Machine Learning in Trauma Care: Advances in Diagnosis, Prognosis, and Treatment. J Trauma Acute Care Surg. 2022;92(3):e57-e65.
2. Sendak MP, D'Arcy J, K ashyap S, et al. A Path for Translation of Machine Learning Products into Healthcare Delivery. EMJ Innov. 2020;4(1):37-43.
3. Topol EJ. High-performance medicine: the convergence of human and artificial intelligence. Nat Med. 2019;25(1):44-56.
4. Esteva A, Robicquet A, Ramsundar B, et al. A guide to deep learning in healthcare. Nat Med. 2019;25(1):24-29.
5. Rajkomar A, Dean J, Kohane IS. Machine Learning in Medicine. N Engl J Med. 2019;380(14):1347-1358.
6. Seymour CW, Kennedy JN, Wang S, et al. Derivation, Validation, and Potential Treatment Implications of Novel Clinical Phenotypes for Sepsis. JAMA. 2019;321(20):2003-2017.
7. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. ACM; 2016:785-794. doi:10.1145/2939672.2939785.
8. Johnson AEW, Pollard TJ, Shen L, et al. MIMIC-III, a freely accessible critical care database. Sci Data. 2016;3:160035.
9. Delano MJ, Rizoli SB, Rhind SG, et al. Advances in the management of hemorrhagic shock. Lancet. 2023;401:1910-1924.
10. Cannon JW. Hemorrhagic Shock. N Engl J Med. 2018;378(4):370-379.
11. Spahn DR, Bouillon B, Cerny V, et al. The European guideline on management of major bleeding and coagulopathy following trauma: sixth edition. Crit Care. 2023;27:80. doi:10.1186/s13054-023-04327-7.
12. Mantskava M, Momtselidze N, Mchedlishvili G. Hemorheological alterations in experimental hemorrhagic shock and perspectives for diagnostic assessment. Series on Biomechanics. 2021;35(2):25-34.
13. Mantskava M, Momtselidze N, Kuchava G, et al. Modern approaches to hemorheological monitoring in experimental medicine. Series on Biomechanics. 2023;37(1):18-29.
14. Ribeiro AH, Ribeiro MH, Paixão GMM, et al. Artificial intelligence and machine learning in cardiovascular medicine. Nat Rev Cardiol. 2020;17(8):507-521.
15. Leone M, Martin-Loeches I, Pirracchio R. Humans versus machine learning models: improving the outcome of trauma patients. Lancet Regional Health – Europe. 2025;55:101356. doi:10.1016/j.lanepe.2025.101356.
16. Sigle M, Boos M, Weiss T, et al. AI-enabled forecasting of prehospital transfusion needs in patients with trauma: a multinational, registry-based, retrospective, machine learning development and validation study. Lancet Digit Health. 2026;8(1):100945. doi:10.1016/j.landig.2025.100945.