Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

First-Day Glycemic Exposure and 28-Day Mortality in the ICU: A Multicohort Study

Joab O. Odera

joab.odera@duke.edu

Duke University

Betsabe G. Blas Achic

Baylor College of Medicine

Julie Cha

Georgia Institute of Technology

Aisha Montgomery

Premier (United States)

Alice A. Ojwang

Diabesity Nutrition Clinic Inc.

Sepiso Masenga

Vanderbilt University

Elizabeth O. Odera

Iris Laboratories Ltd.

Nosayaba Osazuwa-Peters

Duke University

Ananya Yalamanchi

Memorial Hermann–Texas Medical Center

David Han

The University of Texas at San Antonio

Antentor O. Hinton, Jr.

Vanderbilt University

Article

Keywords: glycemic exposure, critical illness, hyperglycemia, survival analysis, artificial intelligence

Posted Date: June 11th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-9710742/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Abstract

Importance: Early glycemic exposure in the ICU is common and clinically modifiable, yet bedside assessment often relies on single glucose values rather than exposure-aware metrics. Interpretable, first-day prediction may support individualized glycemic targets and early intervention.

Objective: To examine the association between first-day time-weighted average glucose (TWAG) and 28-day mortality, and to evaluate GlucoSurvAI, an interpretable ensemble model for first-day risk stratification.

Design, Setting, and Participants: Retrospective cohort study using electornic health records from 13 U.S. hospitals. Among 18,868 adult ICU encounters, 8,048 patients from 7 U.S. hospitals met inclusion criteria (≥1 glucose value and hospital length of stay ≥24 hours).

Exposures: First-day glycemic exposure summarized as TWAG, categorized as <100**,** 100–139 (reference)**,** 140–179, and ≥180 mg/dL. Prespecified covariates included diabetes/prediabetes, first-day insulin and glucose, corticosteroids, vasopressors, shock, cancer, glucose-monitoring intensity, and clinical site.

Main Outcome and Measures: Primary outcome: 28-day all-cause mortality. Associations were estimated with multivariable Cox models (adjusted hazard ratios [aHRs], 95% CIs). GlucoSurvAI performance was assessed using Area Under the Receiver Operating Characteristic (AUROC) and Brier score; SHapley Additive exPlanations (SHAP) provided 28-day interpretability.

Results: Of 8,048 patients, most were euglycemic (70–180 mg/dL) on day 1, although hyperglycemic excursions were frequent. Higher TWAG was associated with higher 28-day mortality: 140–179 mg/dL aHR 1.42 (95% CI, 1.25–1.62)**;** ≥180 mg/dL aHR 1.41 (95% CI, 1.17–1.69). TWAG <100 mg/dL showed a nonsignificant trend toward higher survival. GlucoSurvAI achieved AUROC 0.967 (±0.008) with a low Brier score (~0.026). Adjusted SHAP analyses paralleled Cox results, identifying 100–139 mg/dLas the exposure range associated with decreased mortality, with risk increasing ≥140 mg/dL. First-day vasopressorsand corticosteroids were also associated with higher mortality; insulinexposure marked higher risk after adjustment.

Conclusions and Relevance: During the first ICU day, exposure-aware TWAG assessment identified a practical upper boundary near 140 mg/dL associated with higher 28-day mortality. An interpretable ensemble integrating TWAG, treatments, and physiology provided accurate first-day risk estimates, supporting risk-informed, individualized glycemic targets and earlier intervention in high-risk ICU patients.

Key Points

Question: Among adults admitted to the ICU, is first-day glycemic exposure associated with 28-day mortality, and can an interpretable AI/ML model provide accurate early risk stratification?

Findings: In a multicenter cohort of 8,048 ICU patients, higher first-day time-weighted average glucose (TWAG) showed a strong dose–response relationship with 28-day mortality, with a clear inflection at ≥140 mg/dL. GlucoSurvAI accurately captured this exposure–mortality gradient and outperformed traditional severity scores.

Meaning: Maintaining early glycemic exposure near 100–139 mg/dL may be associated with lower mortality.

Introduction

Hyperglycemia is frequent in critical illness and associated with infection, organ dysfunction, and death.[1,2] Landmark evidence (NICE-SUGAR) demonstrated that intensive glucose targets (81–108 mg/dL) increased mortality and hypoglycemia compared with a conventional target ≤ 180 mg/dL, prompting a shift away from tight control in the ICU.[2,3] Contemporary guidance recommends protocols that minimize hypoglycemia and generally initiate treatment for persistent hyperglycemia (≥ 180 mg/dL), but questions remain about optimal early targets and how best to summarize glycemia during the first ICU day.[3–6]

Most bedside assessments rely on single or sporadic glucose measurements; however, these may misclassify exposure when sampling is intermittent or clinically driven.[7–9] Exposure-aware metrics, such as time-weighted average glucose (TWAG), integrate all measurements and their timing and may better reflect the early glycemic burden that is most relevant to short-term outcomes.[7,9,10] Yet the relationship between first-day TWAG and 28-day mortality, and whether this information can be operationalized into accurate, interpretable first-day risk estimation, remains unclear.[7,9,11–13]

We aimed (1) to examine the association between first-day TWAG and 28-day all-cause mortality in a multicenter ICU cohort, and (2) to evaluate an interpretable ensemble model for early risk stratification that aligns model-based explanations with clinically adjusted time-to-event estimates. We hypothesized that higher first-day TWAG, particularly above a clinically relevant upper boundary, would be associated with increased 28-day mortality and that an interpretable model could capture this exposure–mortality gradient with strong discrimination and calibration.

Methods

Data Source

Our cohort was derived from Bridge2AI CHoRUS AIM-AHEAD60 project, a multicenter resource encompassing diverse ICU populations and clinical contexts from 13 U.S. hospitals. [14,15]

Participants

Eligible patients were aged 18–105 years, had a hospital length of stay ≥ 24 hours, and at least one glucose value recorded during the first ICU day. Among 18,868 unique ICU admissions, 8,048 from 7 of the 13 U.S. hospitals met inclusion criteria after restricting to the index ICU encounter. The participant flow diagram appears in Fig. 1.

Exposures (Glycemic Categories)

Glycemic exposure during the first 24 hours of ICU care was summarized using time-weighted average (TWA) glucose. For clinical interpretability, patients were categorized as hypoglycemic (< 70 mg/dL), euglycemic (70–180 mg/dL), or hyperglycemic (> 180 mg/dL). In analyses examining dose–response, four TWAG bins were evaluated: <100, 100–139 (reference), 140–179, and ≥ 180 mg/dL.

Outcomes

The primary outcome was all-cause mortality by 28 days; mortality by 180 days was a secondary outcome. Survival time was measured from the start of the first 24-hour ICU period. Kaplan–Meier estimates and multivariable time-to-event models were used as described below.

Covariates and Potential Confounders

Prespecified covariates were selected based on clinical relevance and prior literature and included: diabetes status (prediabetes, diabetes), acute severity indicators (e.g., shock, vasopressor score), first-day treatment exposures (insulin, corticosteroids, vasopressors), cancer, and clinical site. To address potential bias from differential glucose monitoring, we adjusted for glucose monitoring intensity and cumulative exposure (e.g., first-day TWAG and number of glucose checks). Variable definitions and coding are detailed in eMethods.

Statistical Analysis

Baseline characteristics are summarized overall and by outcome group (further discussed in eMethods in the Supplement). Survival across glycemic categories was evaluated using Kaplan–Meier curves and log-rank tests (descriptive). The primary inferential analysis used multivariable Cox proportional hazards models to estimate adjusted hazard ratios (aHRs) and 95% CIs for 28-day mortality across glucose categories (reference: 100–139 mg/dL). Models adjusted for the covariates described above, including clinical site, to mitigate confounding due to site-level practice variation; p-values are not emphasized in the narrative, consistent with our focus on aHRs and 95% CIs. Model assumptions and sensitivity analyses are described in eMethods in the Supplement.

To further contextualize glycemic exposure, we calculated restricted mean survival time (RMST) over a fixed 28-day horizon and compared RMST across the prespecified glucose bins (< 100 mg/dL, 100–139

mg/dL, 140–179 mg/dL, and ≥ 180 mg/dL) and across demographic strata (< 50, 50–64, 65–79, and ≥ 80 years; race; and ethnicity). RMST provides an interpretable, time-bounded measure of average survival and is robust when proportional hazards may not fully hold.[16–18]

Prediction Models

For context and benchmarking, we compared traditional severity-score models (SOFA and APACHE II, fit as logistic regressions for 28-day mortality) with a stacked ensemble model (GlucoSurvAI) trained on first-day features. GlucoSurvAI integrates a survival-oriented neural network with gradient-boosted trees and a final meta-learner to yield calibrated short-term mortality predictions. Development procedures (preprocessing, cross-validation, hyperparameter tuning, calibration) and interpretability methods are provided in eMethods in the Supplement.

Performance Assessment

Discrimination for binary 28-day mortality was summarized by AUROC, probability calibration by the Brier score. For time-to-event performance we report Harrell's C-index. Threshold selection and classification metrics are described in eMethods and are not central to clinical interpretation of model comparisons in the main text.

Software

Analyses were performed using Python-based tooling; package versions and code provenance are provided in eMethods.

Reporting and Supplementary Materials

This manuscript adheres to STROBE and TRIPOD + AI recommendations; extended methods, definitions, and robustness checks appear in the eMethods/eTables/eFigures in the Supplement. JAMA Network encourages adherence to EQUATOR reporting guidance and use of the Supplement for technical detail.

Ethics

This multicenter, retrospective study analyzed fully de-identified EHR data contributed by 14 CHoRUS institutions:

1. Ethic approval and consent was waived by the Emory University Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
2. Ethic approval and consent was waived by the University of Florida Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).

3. Ethic approval and consent was waived by the Mass General Brigham Institutional Review Board (covering Massachusetts General Hospital) due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
4. Ethic approval and consent was waived by the University of Pittsburgh Human Research Protection Office Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
5. Ethic approval and consent was waived by the University of Virginia Institutional Review Board for Health Sciences Research due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
6. Ethic approval and consent was waived by the Duke University Health System Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
7. Ethic approval and consent was waived by the Columbia University Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
8. Ethic approval and consent was waived by the University of California, San Francisco Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
9. Ethic approval and consent was waived by the University of California, Los Angeles Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
10. Ethic approval and consent was waived by the Tufts Health Sciences Institutional Review Board (Tufts Clinical and Translational Science Institute) due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
11. Ethic approval and consent was waived by the Committee on Clinical Investigations, the Institutional Review Board for Beth Israel Deaconess Medical Center, due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
12. Ethic approval and consent was waived by the Massachusetts Institute of Technology Committee on the Use of Humans as Experimental Subjects due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46), with oversight conducted in collaboration with Beth Israel Deaconess Medical Center under a reliance agreement where applicable.
13. Ethic approval and consent was waived by the Mayo Clinic Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).
14. Ethic approval and consent was waived by the University of Oregon Institutional Review Board due to the use of fully de-identified data classified as Exempt or Not Human Subjects Research under the U.S. Common Rule (45 CFR 46).

All datasets were de-identified in accordance with the HIPAA Privacy Rule de-identification standard (45 CFR 164.514) using Safe Harbor and/or Expert Determination procedures, and subsequently date-shifted prior to consortium-level curation. No identifiable private information was accessed by the study team.

The research adhered to the ethical principles of the Declaration of Helsinki (2013 Fortaleza amendment) as they pertain to research involving human data, including privacy, confidentiality, and oversight by independent ethics committees.

Ethical review of this data was performed by Eric S. Rosenthal, M.D., Michael J. Young, M.D., MPhil, Prof. Ishan Williams, MS, Ph.D., Ashley Cordes, MA, Ph.D., and Barbara J. Evans, Ph.D., J.D., in collaboration with the Mass General Brigham IRB with local context review by IRBs, HIPAA Privacy Officers, and institutional counsels at contributing medical centers. Individual IRB protocol numbers are maintained by each contributing institution and are available upon reasonable request to the respective IRB offices or the CHoRUS coordinating center.

Results

Cohort selection and first-day glycemic distribution

Among 18,868 ICU admissions, 8,048 unique patients met inclusion criteria (≥ 1 glucose value and hospital length of stay ≥ 24 hours; Fig. 1). By clinical glycemic categories, most patients were euglycemic (70–180 mg/dL) on day 1, with smaller proportions classified as hypoglycemic (< 70 mg/dL) or hyperglycemic (> 180 mg/dL) (distribution in eTable 1). When considering all measurements within the initial 24-hour window, 2,056 patients (25.5%) experienced at least one hyperglycemic episode, whereas 185 (2.3%) had at least one hypoglycemic episode; 7,068 patients (87.8%) remained entirely within the target range (eTable 1 in the Supplement). Of the 8,048 included patients, 158 lacked complete timestamp data for their initial glucose measurements, which prevented calculation of day-one TWAG; these patients were retained in categorical analyses and subsequently included in multivariable models using multiple imputation by chained equations (MICE) for missing time-weighted glucose values, as detailed in the eMethods in the Supplement.

Baseline monitoring and site variation

Baseline monitoring and site differences are summarized in eTable 2 in the Supplement. Mortality differed across sites and was partly explained by case-mix and acuity: centers with higher SOFA/APACHE II burdens also had higher crude mortality. In adjusted models, we included clinical site to mitigate confounding by practice patterns and patient mix; core associations between TWAG bins and mortality were unchanged. Detailed site-level summaries and robustness checks are provided in eResults in the Supplement.

Patient mortality by clinical glycemic category

Kaplan–Meier curves showed significant differences in mortality across clinical glycemic categories (hypoglycemic < 70 mg/dL, euglycemic 70–180 mg/dL, hyperglycemic > 180 mg/dL), with higher mortality in hyperglycemia at both 28 days and 6 months (Fig. 2). In Cox models adjusted for HbA1c category and first-day insulin/glucose administration, hyperglycemia remained associated with increased mortality vs euglycemia (28-day aHR 1.91, 95% CI 1.35–2.70; 6-month aHR 1.63, 95% CI 1.18–2.25), while hypoglycemia was not significant after adjustment. Additional treatment and HbA1c associations are described in eResults in the Supplement.

GlucoSurvAI shows superior predictive performance when compared to AI/ML models

Figure 3B summarizes the performance of seven machine learning and deep learning models predicting 28-day mortality, evaluated across multiple metrics including AUROC, accuracy, sensitivity, specificity, precision, recall, F1 score, and Brier score.

GlucoSurvAI achieved the highest raw AUROC (0.967 ± 0.008) in ROC curve analysis (Fig. 3A), maintained strong accuracy (90%), balanced sensitivity (89.9%) and specificity (90.7%), and a high F1 score (94.3%), and a low Brier score (0.026). A schematic of GlucoSurvAI's architecture is provided in eFigure 1 in the Supplement.

Additionally, site-level performance metrics showed that GlucoSurvAI AUROC values consistently exceeding 0.96 across major sites (Site 13: 0.988; Site 2: 0.980; Site 4: 0.964; Site 3: 0.963; Site 12: 0.975), coupled with high sensitivity (> 0.93) and strong negative predictive values (> 0.89) (eTable 3 in the Supplement).

TWAG exposure bins: SHAP and adjusted hazards

Model explanations from GlucoSurvAI, focused on 28-day survival and adjusted for diabetes/prediabetes, first-day insulin, glucose, corticosteroids, and vasopressors, shock, cancer, monitoring intensity, and clinical site indicated higher 28-day survival for 100–139 mg/dL, somewhat higher survival for < 100 mg/dL, and lower predicted survival for exposures ≥ 140 mg/dL (Fig. 4). These adjusted 28-day SHAP patterns aligned with adjusted 28-day GlucoSurvAI Cox models (Fig. 5, eTable 4A in the Supplement): relative to 100–139 mg/dL, mortality risk was higher for 140–179 mg/dL (aHR 1.42, 95% CI 1.25–1.62) and ≥ 180 mg/dL (aHR 1.41, 95% CI 1.17–1.69).

Beyond glycemic strata, several physiologic and clinical variables were independently associated with mortality after adjustment for confounders and for the reference glucose range (100–139 mg/dL). Among comorbidities, hypertension (aHR 0.76, 95% CI: 0.68–0.85) and COPD (aHR 0.84, 95% CI: 0.71–0.98) were associated with decreased mortality, whereas sepsis (aHR 1.28, 95% CI: 1.13–1.45) and brain injury (aHR 1.53, 95% CI: 1.22–1.93) markedly increased risk. Interestingly, psychological disorders (aHR 0.75, 95% CI: 0.67–0.84) and viral disease (aHR 0.69, 95% CI: 0.50–0.96) were associated with lower hazards, while organ failure involving heart, kidney, or liver showed modest decreased hazards (aHR

0.89, 95% CI: 0.79–0.99); which may indicate complex interactions with critical illness trajectories (eTable 4(A) and eFigure 2 in the Supplement).

When stratified by HbA1c category, risk patterns varied across glucose bins. In the ≥ 180 mg/dL group, patients with diabetes showed lower hazards compared with Normal HbA1c, whereas those with prediabetes exhibited increased risk. Diabetic patients with an end-of-day TWAG between 140–179 mg/dL had significantly decreased risk of mortality, while combinations within < 100 mg/dL and 100–139 mg/dL bins were largely neutral (eFigure 3 in the Supplement). Details on the effect of confounders may be found in eResults and eTable 4B in the Supplement.

GlucoSurvAI Stratified Analysis of Glycemic Exposure by Demographic groups reveals at risk populations

Mortality was lowest in the 100–139 mg/dL range across all strata, with restricted mean survival time (RMST) approximating 27 days. Event rates increased sharply in the 140–179 mg/dL and ≥ 180 mg/dL bins, particularly among patients aged ≥ 80 years, where mortality exceeded 16% and RMST declined to ~ 24–25 days. Younger critically ill patients (< 50 years) also exhibited elevated risk at higher glucose levels, with event rates surpassing 25% in the 140–179 mg/dL and ≥ 180 mg/dL categories.

Across racial groups, White patients consistently demonstrated lower event rates and longer RMST within the optimal glycemic range, whereas Black and Other/Unknown race categories showed higher mortality in hyperglycemic strata. Ethnicity trends were less pronounced, but Hispanic patients in older age groups (≥ 65 years) exhibited increased mortality at glucose levels ≥ 140 mg/dL compared to non-Hispanic counterparts. These patterns were mirrored by higher Cox hazard scores and GlucoSurvAI-predicted mortality probabilities (eTable 5 in the Supplement).

We expanded our analyses to include major ICU admission reasons: injury/trauma, brain injury, neurological conditions, viral and bacterial infections, surgery, organ failure, sepsis, and psychological conditions, stratified by glycemic exposure and age (eTable 6 in the Supplement). Across all strata, sepsis and organ failure consistently exhibited the highest mortality, particularly in hyperglycemic patients aged ≥ 80 years, where event rates exceeded 18–22% and RMST declined to ~ 23–25 days. Elevated risk was also observed in younger patients (< 50 years) admitted for sepsis, with event rates approaching 19–23% in the 140–179 mg/dL and ≥ 180 mg/dL bins.

Conversely, patients admitted for psychological or trauma-related indications had lower mortality in the optimal glycemic range (100–139 mg/dL), with RMST near 27 days. Mortality risk increased substantially under hyperglycemia, with event rates rising to 7–11% in older adults, and up to 33% in younger trauma patients at ≥ 140 mg/dL. Neurologic and brain injury cases showed intermediate risk patterns, with mortality climbing sharply in the highest glucose strata, especially among the oldest patients. (eTable 6).

Discussion

In this multicenter ICU cohort, we developed GlucoSurvAI, an interpretable, stacked survival model that integrates deep learning and gradient boosting to characterize first-day glycemia and predict 28-day mortality using early glycemic and physiologic data. Elevated first-day glucose exposure was associated with higher mortality, underscoring the importance of quantifying per-patient early glycemia rather than focusing solely on hypoglycemia avoidance. These findings align with evolving guidance that favors moderate glycemic targets and emphasizes minimizing both hyperglycemia and hypoglycemia in critical care.[4,12]

Most patients began ICU care within the euglycemic category by conventional clinical cut points, yet hyperglycemic excursions were common over the first 24 hours. We found that time-weighted average glucose (TWAG) offers a more comprehensive summary of cumulative exposure than single-point or short-window metrics, which can misestimate true burden when sampling is intermittent or clustered in time.[3,7,8] By integrating every measurement and its duration, TWAG reduced misclassification evident with shorter windows, a pattern consistent with emerging reports that exposure-based metrics better reflect risk.[7,9]

Using prespecified TWAG exposure bins, both adjusted Cox models and adjusted SHAP explanations from GlucoSurvAI converged on 100–139 mg/dL as the range associated with higher 28-day survival, with higher risk beginning at 140 mg/dL. Relative to the 100–139 mg/dL reference, the 140–179 mg/dL and ≥ 180 mg/dL bins had higher adjusted hazards for 28-day mortality; <100 mg/dL trended toward higher survival but with wider uncertainty due to smaller numbers. The close alignment between adjusted SHAP (model-based explanations) and adjusted Cox (time-to-event estimates) increases confidence that this ~ 140 mg/dL inflection is a clinically meaningful signal in the initial 24 hours of an ICU admission.[7,19,20]

Interpreting glycemia in context is essential. In our adjusted models, first-day vasopressors and corticosteroids were associated with higher mortality, plausibly reflecting illness severity and treatment intensity, while insulin exposure also marked higher risk after adjustment. These associations highlight the interplay between hemodynamic support, pharmacologic therapy, and metabolic control, and reinforce that early glycemic exposure should be considered alongside acuity and therapeutic milieu when formulating risk-informed targets.[21–24]

Traditional ICU mortality prediction scores, such as SOFA and APACHE II, are well validated and have been widely adopted for decades.[25,26] Moreover, both SOFA and APACHE scores provide a robust framework for assessing illness severity and evaluating organ function in critically ill patients.[26,27] However, they were originally developed using older, often single-center cohorts, which may limit their ability to fully reflect contemporary patient complexity and evolving treatment practices.[25,26] In this study, we leveraged a modern, multicenter ICU dataset to benchmark these established scores against advanced AI/ML models and novel glycemic indices derived from time-weighted glucose exposure. Our findings show that while APACHE II and SOFA maintain respectable predictive performance (AUROC = 0.749 and 0.754, respectively), they are substantially outperformed by GlucoSurvAI (AUROC = 0.976),

which integrates dynamic glycemic metrics with physiological and comorbidity features. Additionally, compared with other AI/ML models, GlucoSurvAI avoids extreme trade-offs between precision and recall and offers practical advantages in speed, interpretability, and ease of deployment. These attributes make it a robust and well-rounded choice for survival prediction in this study.[28–30]

Beyond point estimates of accuracy, GlucoSurvAI preserved interpretability via SHAP, allowing clinicians to visualize how exposure bins and treatment covariates contribute to predicted risk at the bedside. Because the model relies on first-day information, it offers a practical path for early risk stratification that complements, and does not replace, established scores and clinical judgment.[30,31]

Stratified summaries suggested that the glycemia–survival gradient is context-dependent: it appeared steeper at higher exposure among older adults and in sepsis/organ-failure admissions, whereas some lower-acuity indications showed flatter gradients until exposure was markedly elevated. These patterns, together with the adjusted bin findings, support risk-informed targets rather than universal thresholds, consistent with literature showing heterogeneous glycemic risk across diagnoses and chronic glycemic status.[8,32]

In our adjusted analyses, patients with established diabetes (higher HbA1c) showed attenuated mortality risk in the upper TWAG bins, whereas those with prediabetes/normal HbA1c had higher risk at the same exposures. This pattern is consistent with prior work that indicate that acute stress hyperglycemia is more prognostic in patients without chronic hyperglycemia, while individuals with long-standing hyperglycemia may exhibit partial physiologic adaptation.[33–36] Taken together, our results support individualized early glucose targets that account for baseline glycemic status (HbA1c) when interpreting first-day exposure.[8,37]

The present results reinforce a growing body of work linking higher early glycemic exposure and lower time-in-range with worse short-term outcomes in ICU populations.[7,9] They also harmonize with the post–NICE-SUGAR shift away from intensive 80–110 mg/dL targets toward moderate ranges that reduce hypoglycemia risk while acknowledging the harm of sustained hyperglycemia.[4,38] Importantly, our analysis quantifies an exposure-oriented signal during the first 24 hours and demonstrates how combining TWAG bins with an interpretable ensemble model can make this signal clinically actionable in real-time.

Strengths of our study include (1) a diverse, multisite cohort; (2) emphasis on exposure-aware metrics that reflect cumulative glycemia; (3) convergence between adjusted Cox and adjusted SHAP results; and (4) rigorous internal validation with discrimination and calibration reporting. Limitations include the retrospective design; variation in measurement frequency and modalities across sites; and the absence of standardized insulin/nutrition protocols, each of which may influence both exposure and outcomes. Although we adjusted for glucose monitoring intensity and included clinical sites to mitigate case-mix and practice differences, residual confounding remains possible. External validation and prospective

evaluation, especially of model-informed glycemic management, are needed before clinical deployment (please see eAppendix).[12,39]

Collectively, our findings support maintaining early exposure in the 100–139 mg/dL range and avoiding sustained exposure ≥ 140 mg/dL during the first ICU day, while recognizing patient-specific factors, competing risks, and the real-world challenges of insulin titration. Pairing exposure-based metrics with interpretable model probabilities may help clinicians identify who is most likely to benefit from early titration and when the risk–benefit balance favors intervention, particularly in older adults and in high-acuity indications such as sepsis and organ failure.[7,9,37]

In conclusion, first-day glycemic exposure is common, measurable, and actionable. TWAG offers a robust summary of early glycemia, and both adjusted Cox and adjusted SHAP indicate 100–139 mg/dL is associated with higher 28-day survival, with higher risk beginning at 140 mg/dL. GlucoSurvAI serves as an interpretable ensemble model that leverages these exposure-aware metrics enables early, individualized risk stratification and provides a foundation for prospective, risk-informed glycemic management in critical care.[12,37]

Declarations

Acknowledgements

We thank Ori J. Lieberman, M.D., Ph.D. (Department of Neurology, University of California San Francisco, San Francisco, CA) for his substantive review of the manuscript, expert clinical insights from critical care practice as a critical care physician, and early contributions to the SQL query development and cohort definition as part of the Bridge2AI Clinical Care Program, Cohort 1 (2025).

Special thanks to the Artificial Intelligence/Machine Learning Consortium to Advance Health Equity and Researcher Diversity (AIM-AHEAD) Coordinating Center and the Bridge2AI Clinical Care team, including Eric S. Rosenthal, M.D., Deandra Maevry, Lesly Grandstaff, M.Ed., Ciera McCary, M.S., Jared Houghtaling, Ph.D., Omar M, Aljawfi, Ph.D. and Toufeeq Ahmed Syed, Ph.D., as well as all visiting lecturers in the Bridge2AI Clinical Care Training Program for their provision of resources, training and advice.

Funding Declaration

This research was, in part, funded by the National Institutes of Health (NIH) Agreement No. 1OT2OD032581.

This work was also supported by the Bridge2AI CHoRUS Clinical Care for AI, funded by the NIH Common Fund (J.O.O, B.B., J.C., A.M., D.H., OT2OD032701), National Institute on Aging (J.O.O., T32AG000029) UNCF/Bristol-Myers Squibb E.E. Just Faculty Fund, Career Award at the Scientific Interface (CASI Award) from the Burroughs Welcome Fund (BWF) ID # 1021868.01, BWF Ad-hoc Award, NIH Small Research Pilot Subaward 5R25HL106365-12 from the National Institutes of Health PRIDE Program, DK020593,

Vanderbilt Diabetes and Research Training Center for DRTC Alzheimer's Disease Pilot & Feasibility Program, CZI Science Diversity Leadership grant number 2022-253529 from the Chan Zuckerberg Initiative DAF, an advised fund of the Silicon Valley Community Foundation (A.H.J).

Author Contributions

J.O.O. conducted all the data processing, model training, formal analysis, table and figure development, and writing of the manuscript. J.O.O., B.B., J.C., A.M. contributed to the study conception and design, data interpretation, and reviewing the manuscript. A.H.J., A.Y., D.H., N.O.P. contributed to the study design and reviewing the manuscript. S.M., E.O.O., and A.A.O. critically reviewed the manuscript.

Competing Interests

The authors declare no competing of interest.

Data availability

This dataset is only available under controlled access after a registration review. One may register for data access at https://chorus4ai.org/dataset/

References

1. Liu, J. et al. Predicting mortality of patients with acute kidney injury in the ICU using XGBoost model. *PLoS One*. **16** (2), e0246306. 10.1371/journal.pone.0246306 (2021).
2. Investigators, N-S-S. et al. Hypoglycemia and risk of death in critically ill patients. *N Engl. J. Med Sep*. **20** (12), 1108–1118. 10.1056/NEJMoa1204942 (2012).
3. American Diabetes Association Professional Practice C. ;47(Suppl 1):S111-S125. Jan 1 (2024). 10.2337/dc24-S006
4. Finfer, S. et al. Clinical review: Consensus recommendations on measurement of blood glucose and reporting glycemic control in critically ill adults. *Crit Care Jun*. **14** (3), 229. 10.1186/cc12537 (2013).
5. Krinsley, J. S. et al. Jul. Glucose Control, Diabetes Status, and Mortality in Critically Ill Patients: The Continuum From Intensive Care Unit Admission to Hospital Discharge. *Mayo Clin Proc*. ;92(7):1019–1029. (2017). 10.1016/j.mayocp.2017.04.015
6. Marik, P. E. & Bellomo, R. Stress hyperglycemia: an essential survival response! *Crit Care Mar.* **6** (2), 305. 10.1186/cc12514 (2013).
7. Feng, M. & Zhou, J. Relationship between time-weighted average glucose and mortality in critically ill patients: a retrospective analysis of the MIMIC-IV database. *Sci Rep Feb*. **27** (1), 4721. 10.1038/s41598-024-55504-9 (2024).
8. Fong, K. M., Au, S. Y. & Ng, G. W. Y. Glycemic control in critically ill patients with or without diabetes. *BMC Anesthesiol Jul*. **16** (1), 227. 10.1186/s12871-022-01769-4 (2022).

9. Naraba, H. et al. Time in blood glucose range 70 to 180 mg/dL and survival rate in critically ill patients: A retrospective cohort study. *PLoS One*. **16** (5), e0252158. 10.1371/journal.pone.0252158 (2021).
10. Wang, Y., Rong, J., Wei, Z., Bai, X. & Deng, Y. Predicting ICU mortality in heart failure patients based on blood tests and vital signs. *Front. Cardiovasc. Med.* **12**, 1590367. 10.3389/fcvm.2025.1590367 (2025).
11. Gunst, J., Umpierrez, G. E. & Van den Berghe, G. Managing blood glucose control in the intensive care unit. *Intensive Care Med Dec.* **50** (12), 2171–2174. 10.1007/s00134-024-07687-y (2024).
12. Honarmand, K. et al. Society of Critical Care Medicine Guidelines on Glycemic Control for Critically Ill Children and Adults 2024. *Crit Care Med Apr*. **1** (4), e161–e181. 10.1097/CCM.0000000000006174 (2024).
13. Fitzgerald, O. et al. Incorporating real-world evidence into the development of patient blood glucose prediction algorithms for the ICU. *J Am. Med. Inf. Assoc Jul*. **30** (8), 1642–1650. 10.1093/jamia/ocab060 (2021).
14. Rosenthal, E. S. B. et al. *A. E. Bridge2AI: Bridge2AI: Patient-Focused Collaborative Hospital Repository Uniting Standards (CHoRUS) for Clinical Care AI* (National Institutes of health (NIH), 2025).
15. Wang, H. et al. Early Sepsis Prediction Using Publicly Available Data: High-Performance AI/ML Models with First-Hour Clinical Information. *Diagnostics (Basel) Oct.* **28** (21). 10.3390/diagnostics15212727 (2025).
16. Royston, P. & Parmar, M. K. Restricted mean survival time: an alternative to the hazard ratio for the design and analysis of randomized trials with a time-to-event outcome. *BMC Med. Res. Methodol Dec.* **7**, 13:152. 10.1186/1471-2288-13-152 (2013).
17. Trinquart, L., Jacot, J., Conner, S. C. & Porcher, R. Comparison of Treatment Effects Measured by the Hazard Ratio and by the Ratio of Restricted Mean Survival Times in Oncology Randomized Controlled Trials. *J Clin. Oncol May*. **20** (15), 1813–1819. 10.1200/JCO.2015.64.2488 (2016).
18. Uno, H. et al. Moving beyond the hazard ratio in quantifying the between-group difference in survival analysis. *J Clin. Oncol Aug*. **1** (22), 2380–2385. 10.1200/JCO.2014.55.2208 (2014).
19. Wang, D. et al. Glycemic variability and its association with short and long term clinical outcomes in critically ill patients with cerebral hemorrhage. *Sci Rep Mar.* **6** (1), 7820. 10.1038/s41598-025-92415-9 (2025).
20. Wang, T-S-L., Wong, S-Y. & Chao, L-T. W-C. Blood Urea Nitrogen to Albumin Ratio Was Associated With Mortality in Critically Ill Septic Patients: A Multicenter Retrospective Propensity–Adjusted Analysis. *International Journal of Clinical Practice*. ;2024 (2024). doi:https://doi.org/10.1155/ijcp/5202122
21. Clore, J. N. & Thurby-Hay, L. Glucocorticoid-induced hyperglycemia. *Endocr Pract Jul-Aug*. **15** (5), 469–474. 10.4158/EP08331.RAR (2009).
22. Deng, Y. et al. Impact of long-acting glucocorticoids on ICU mortality in septic patients with acute respiratory failure: a MIMIC-IV based cohort study. *Front. Pharmacol.* **16**, 1663974.

10.3389/fphar.2025.1663974 (2025).

23. Mitwally, H., Saad, M. O., Mahmoud, S. & Mohamed, A. Hyperglycemia Risk Evaluation of Hydrocortisone Intermittent Boluses vs Continuous Infusion in Septic Shock: A Retrospective Study. *Indian J. Crit. Care Med.* **25** (1), 29–33. 10.5005/jp-journals-10071-23501 (Jan 2021).
24. Motiejunaite, J., Deniau, B., Blet, A., Gayat, E. & Mebazaa, A. Inotropes and vasopressors are associated with increased short-term mortality but not long-term survival in critically ill patients. *Anaesth. Crit. Care Pain Med.* **41** (1), 101012. 10.1016/j.accpm.2021.101012 (Feb 2022).
25. Moreno, R. et al. Rationale and Methodological Approach Underlying the Development of the Sequential Organ Failure Assessment (SOFA)-2 Score: A Consensus Statement. *JAMA Netw. Open Oct.* **1** (10), e2545040. 10.1001/jamanetworkopen.2025.45040 (2025).
26. Mutchmore, A., Lamontagne, F., Chasse, M., Moore, L. & Mayette, M. Automated APACHE II and SOFA score calculation using real-world electronic medical record data in a single center. *J Clin. Monit. Comput Aug*. **37** (4), 1023–1033. 10.1007/s10877-023-01010-8 (2023).
27. Naqvi, I. H., Mahmood, K., Ziaullaha, S., Kashif, S. M. & Sharif, A. Better prognostic marker in ICU - APACHE II, SOFA or SAP II! *Pak J. Med. Sci Sep-Oct*. **32** (5), 1146–1151. 10.12669/pjms.325.10080 (2016).
28. Chan, M. C. et al. Explainable machine learning to predict long-term mortality in critically ill ventilated patients: a retrospective study in central Taiwan. *BMC Med. Inf. Decis. Mak Mar.* **25** (1), 75. 10.1186/s12911-022-01817-6 (2022).
29. Oliullah, K. et al. A stacked ensemble machine learning approach for the prediction of diabetes. *J Diabetes Metab. Disord Jun*. **23** (1), 603–617. 10.1007/s40200-023-01321-2 (2024).
30. Rajkomar, A. et al. Scalable and accurate deep learning with electronic health records. *NPJ Digit. Med.* **1**, 18. 10.1038/s41746-018-0029-1 (2018).
31. Lim, L. et al. Real-time machine learning model to predict short-term mortality in critically ill patients: development and international validation. *Crit Care Mar.* **14** (1), 76. 10.1186/s13054-024-04866-7 (2024).
32. Fitzgerald, O., Perez-Concha, O., Gallego-Luxan, B., Rudd, L. & Jorm, L. The relationship between hyperglycaemia on admission and patient outcome is modified by hyperlactatemia and diabetic status: a retrospective analysis of the eICU collaborative research database. *Sci Rep Sep.* **21** (1), 15692. 10.1038/s41598-023-43044-7 (2023).
33. Tian, J., Zhou, T., Liu, Z., Dong, Y. & Xu, H. Stress hyperglycemia is associated with poor prognosis in critically ill patients with cardiogenic shock. *Front. Endocrinol. (Lausanne)*. **15**, 1446714. 10.3389/fendo.2024.1446714 (2024).
34. You, X. et al. Stress hyperglycemia ratio and 30-day mortality among critically ill patients with acute heart failure: analysis of the MIMIC-IV database. *Acta Diabetol Mar.* **15** 10.1007/s00592-025-02486-3 (2025).
35. Zhang, J., Xia, J., Niu, Z., Zhu, H. & Wang, X. Association of stress hyperglycemia ratio with mortality in sepsis-associated acute kidney injury: a retrospective analysis of the MIMIC-IV database. *Sci Rep*

Aug. **20** (1), 30667. 10.1038/s41598-025-16783-y (2025).

36. Zhang, S. et al. Association between stress hyperglycemia ratio and all-cause mortality in critically ill patients with sepsis: results from the MIMIC-IV database. *Eur J. Med. Res Jan*. **21** (1), 42. 10.1186/s40001-025-02281-4 (2025).
37. Feng, S. et al. Stress hyperglycemia ratio as a mortality predictor in non-diabetic septic patients: a retrospective cohort analysis. *BMC Infect. Dis May*. **25** (1), 752. 10.1186/s12879-025-11151-7 (2025).
38. van den Berghe, G. et al. Intensive insulin therapy in critically ill patients. *N Engl. J. Med Nov*. **8** (19), 1359–1367. 10.1056/NEJMoa011300 (2001).
39. Li, X. et al. Comparison of all-cause mortality with different blood glucose control strategies in patients with diabetes in the ICU: a network meta-analysis of randomized controlled trials. *Ann Intensive Care Apr*. **9** (1), 51. 10.1186/s13613-025-01471-x (2025).

Figures

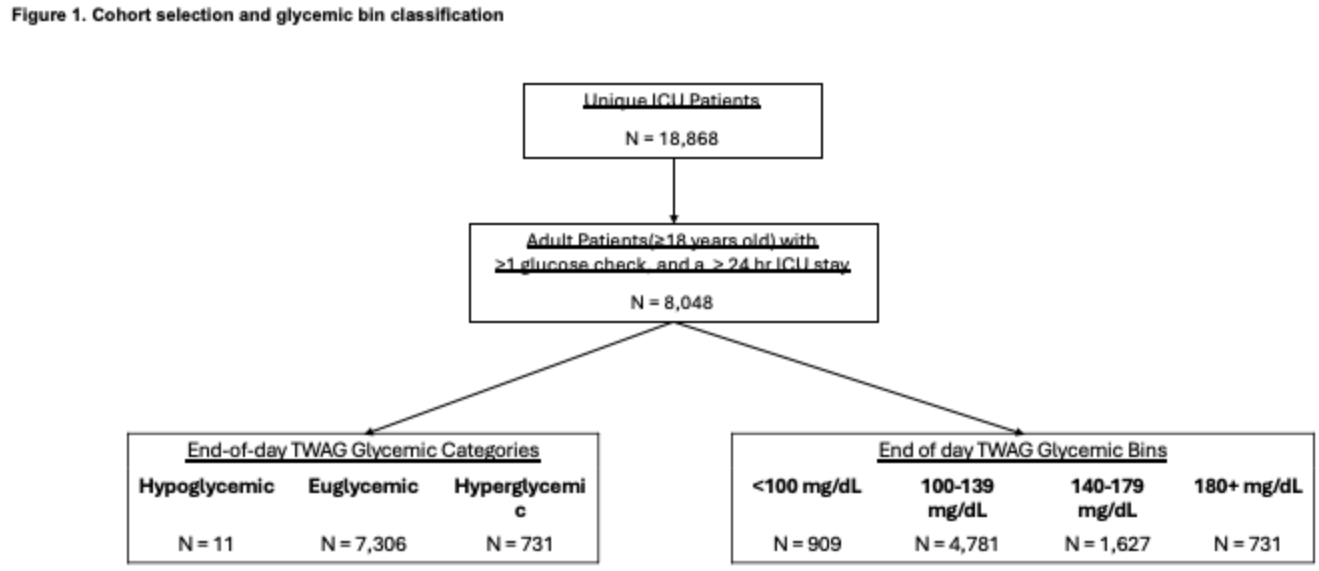
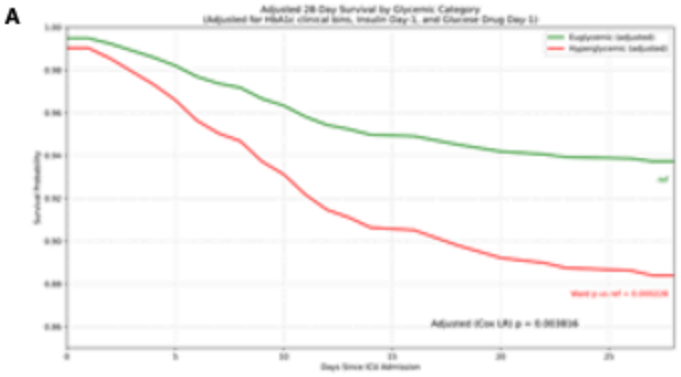


Figure 1

Cohort selection and glycemic bin classification. Flow diagram illustrating the derivation of the study cohort from 18,868 unique ICU patients. After applying inclusion criteria (age ≥18 years, >1 glucose measurement, and ICU stay >24 hours), 8,048 patients were retained for analysis. Patients were categorized into end-of-day time-weighted average glucose (TWAG) bins using two approaches: (1)

clinical glycemic states: hypoglycemic (N = 11), euglycemic (N = 7,306), and hyperglycemic (N = 731); and (2) numeric TWAG ranges: <100 mg/dL (N = 909), 100–139 mg/dL (N = 4,781), 140–179 mg/dL (N = 1,627), and ≥180 mg/dL (N = 731).

Figure 2: Adjusted Kaplan–Meier Survival Curves and Cox Hazard Ratios for Glycemic Categories and Key Covariates



B

Covariable (28-day mortality)	aHR	95%CI	p-value
Glycemic Category			
Euglycemic	Reference		
Hyperglycemic	1.91	1.35-2.70	<0.0001
HbA1c			
Normal HbA1c (HbA1c <5.7)	Reference	1	1
Prediabetic (HbA1c 5.7-6.5)	0.96	0.76-1.22	0.87
Diabetic (HbA1c 6.5 – 9.0)	0.96	0.75-1.24	0.50
Diabetic with poorly controlled glucose (HbA1c >9.0)	0.58	0.41-083	0.003
Insulin Day 1 Administration			
Not administered	Reference	1	1
Administered	2.33	1.62-3.34	<0.0001
Glucose Day 1 Administration			
Not administered	Reference	1	1
Administered	0.67	0.52-0.85	0.001

D

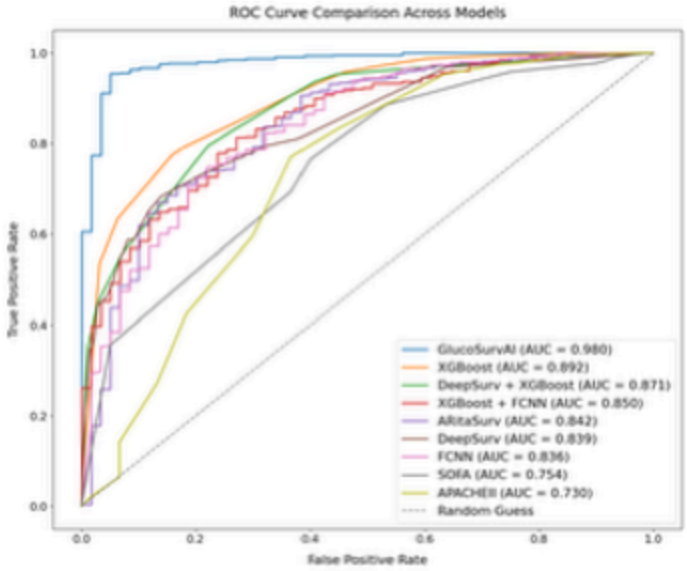
Covariable(6-month mortality)	aHR	95% CI	p-value
Glycemic Category			
Euglycemic	Reference	1	1
Hyperglycemic	1.63	1.18-2.25	<0.001
HbA1c			
Normal HbA1c (HbA1c <5.7)	Reference	1	1
Prediabetic (HbA1c 5.7-6.5)	1.02	0.81-1.28	0.87
Diabetic (HbA1c 6.5 – 9.0)	1.09	0.85-1.39	0.50
Diabetic with poorly controlled glucose (HbA1c >9.0)	0.48	0.33-0.70	<0.0001
Insulin Day 1 Administration			
Not administered	Reference	1	1
Administered	2.49	1.81-3.43	<0.0001
Glucose Day 1 Administration			
Not administered	Reference	1	1
Administered	0.66	0.53-0.83	<0.001

Figure 2

Adjusted Kaplan–Meier Survival Curves and Cox Hazard Ratios for Glycemic Categories and Key Covariates A) Kaplan–Meier curves for 28-day survival stratified by glycemic category (euglycemic, hypoglycemic, hyperglycemic), adjusted for age, comorbidities, and treatment covariates. Hyperglycemia shows the steepest decline in survival, while hypoglycemia also confers elevated risk compared to euglycemia. **B)**Kaplan–Meier curves for 180-day survival by glycemic category, illustrating persistent mortality risk associated with hyperglycemia and hypoglycemia beyond the acute phase. **C)** Survival curves stratified by HbA1c-defined chronic glycemic status (normal, prediabetes, diabetes, poorly controlled diabetes). Patients with poorly controlled diabetes (HbA1c >9%) exhibit distinct long-term survival patterns compared to normoglycemic individuals. **D)**Survival curves comparing patients with and without insulin administration on day 1 of ICU admission. Insulin exposure is associated with higher adjusted mortality risk, reflecting treatment complexity and underlying severity.

Figure 3. Comparative ROC Curves and Performance Metrics Across Candidate Models

A



B

Metric	Logistic SOFA	Logistic APACHEII	XGBoost	DeepSurv	DeepSurv+XGB	FCNN	XGB+FCNN	ARitaSurvAI	GlucoSurvAI
AUROC	0.711 ± 0.029	0.744 ± 0.030	0.887 ± 0.021	0.877 ± 0.024	0.873 ± 0.023	0.847 ± 0.020	0.865 ± 0.034	0.875 ± 0.029	0.967 ± 0.008
Accuracy	0.889 ± 0.008	0.698 ± 0.017	0.939 ± 0.004	0.937 ± 0.004	0.937 ± 0.003	0.795 ± 0.040	0.834 ± 0.046	0.799 ± 0.012	0.900 ± 0.006
Sensitivity	0.934 ± 0.008	0.701 ± 0.017	0.997 ± 0.003	0.805 ± 0.233	0.721 ± 0.232	0.796 ± 0.046	0.838 ± 0.051	0.799 ± 0.012	0.899 ± 0.006
Specificity	0.320 ± 0.033	0.661 ± 0.061	0.207 ± 0.076	0.512 ± 0.369	0.651 ± 0.366	0.786 ± 0.043	0.779 ± 0.065	0.799 ± 0.799	0.907 ± 0.021
Precision	0.945 ± 0.003	0.963 ± 0.006	0.941 ± 0.005	0.747 ± 0.239	0.658 ± 0.239	0.979 ± 0.004	0.980 ± 0.005	0.981 ± 0.981	0.992 ± 0.002
F1 Score	0.940 ± 0.005	0.811 ± 0.012	0.968 ± 0.002	0.771 ± 0.237	0.682 ± 0.235	0.877 ± 0.028	0.903 ± 0.030	0.881 ± 0.881	0.943 ± 0.004
Brier Score	0.063 ± 0.002	0.062 ± 0.002	0.048 ± 0.002	0.050 ± 0.005	0.050 ± 0.005	0.059 ± 0.004	0.051 ± 0.003	0.051 ± 0.051	0.026 ± 0.004

Figure 3

Comparative ROC Curves and Performance Metrics Across Candidate Models. A) Receiver Operating Characteristic (ROC) curves for all candidate models: SOFA, APACHE II, XGBoost, DeepSurv, DeepSurv + XGBoost, FCNN, XGBoost + FCNN, ARitASurv, and GlucoSurvAI, plotted against a "random guess" diagonal baseline. GlucoSurvAI demonstrates the highest discrimination (AUROC ≈ 0.980), followed by XGBoost (≈ 0.894) and XGBoost + DeepSurv (≈ 0.892), while traditional severity scores (SOFA and APACHE II) show substantially lower performance (≈ 0.754 and ≈ 0.730, respectively). **B)**Summary table of cross-validated performance metrics for each model, including AUROC, accuracy, sensitivity, specificity, precision, F1 score, and Brier score (mean ± SD). GlucoSurvAI achieved the best overall calibration and discrimination (AUROC 0.967 ± 0.008; Brier score 0.026 ± 0.004), while maintaining balanced sensitivity (0.899) and specificity (0.907), outperforming both single learners and other ensembles.

Figure 4. SHAP Analysis of Glycemic Bins and Survival Impact in GlucoSurvAI

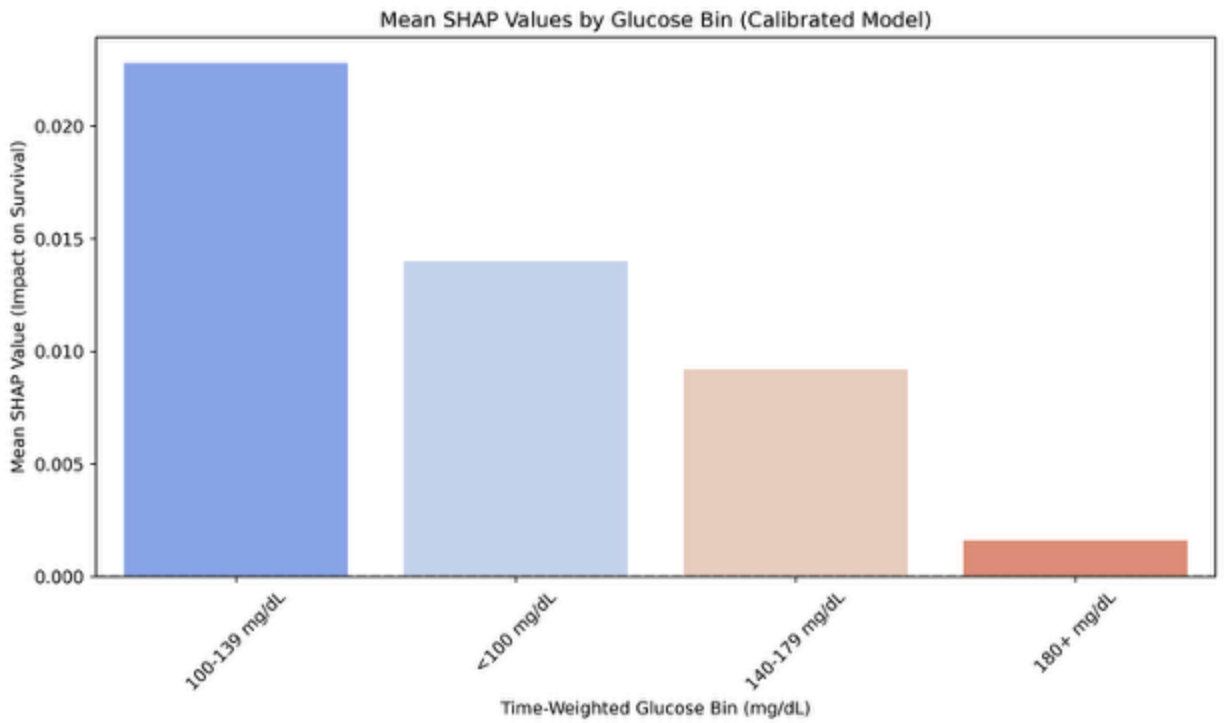


Figure 4

SHAP Analysis of Glycemic Bins and Survival Impact in GlucoSurvAI. SHAP (SHapley Additive exPlanations) summary plot illustrating the contribution of glucose bin features to survival prediction within the GlucoSurvAI ensemble. Glycemic exposures above 140 mg/dL were associated with progressively higher mortality risk, with the strongest negative impact observed for glucose 180+ mg/dL. Conversely, the optimal survival range clustered between 100–139 mg/dL. Feature importance values reflect adjusted models controlling for multiple confounders, including diabetes status, septic or cardiogenic shock, first-day glucose, insulin and corticosteroid administration, vasopressor exposure, time weighted glucose and cancer status. This analysis underscores the additive prognostic value of granular glycemic bins beyond conventional thresholds.

Figure 5. Forest Plot of Adjusted Hazard Ratios for Glycemic Bins

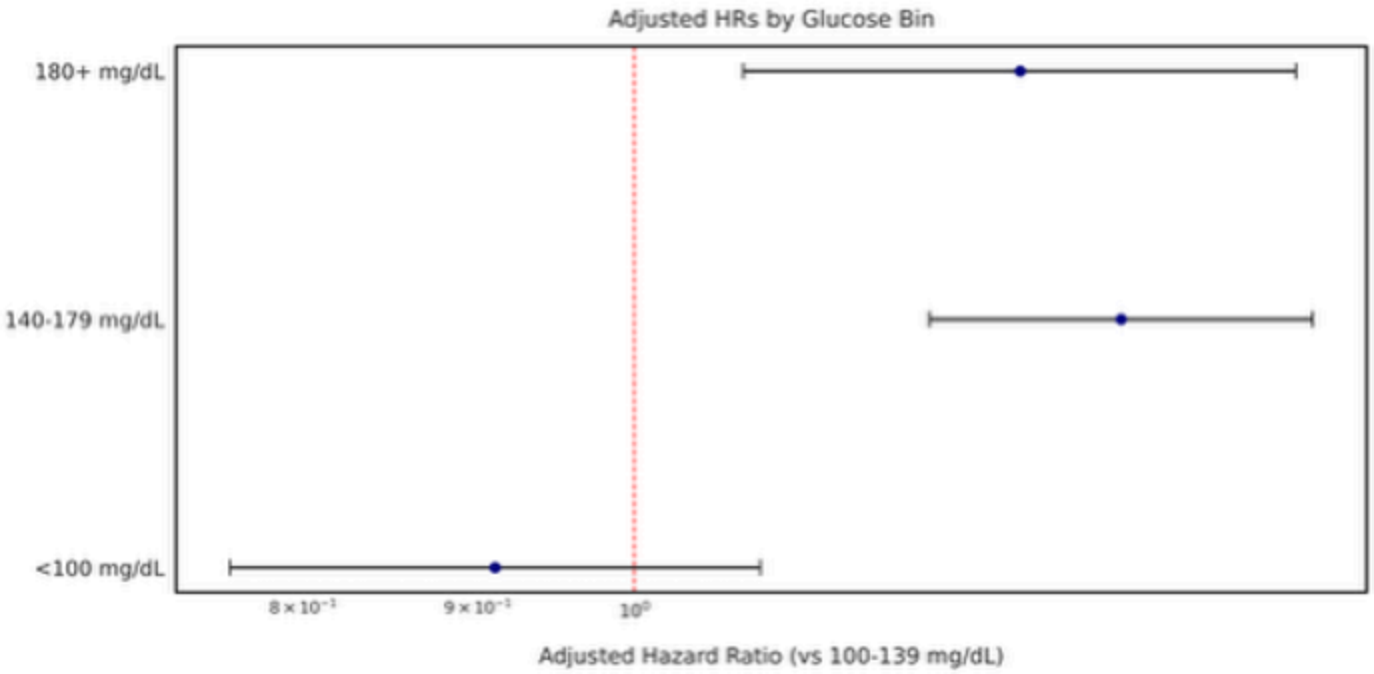


Figure 5

Forest Plot of Adjusted Hazard Ratios for Glycemic Bins. Forest plot displaying adjusted hazard ratios (aHR) and 95% confidence intervals for predefined glucose bins (<100 mg/dL, 100–139 mg/dL, 140–179 mg/dL, 180+ mg/dL) in relation to 28-day and 180-day mortality. Analyses were adjusted for key confounders, including diabetes status, HbA1c category, vasopressor exposure, corticosteroid use, glucose & insulin administration, septic or cardiogenic shock, and other clinical covariates. 100–139 mg/dL range served as a reference.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementalJOO.docx
- eAppendixJOO.docx
- SupplementaleTablesJOO.xlsx