*International Journal of Scientific Research and Modern Technology (IJSRMT)* ijsrmt.com
*Volume 5, Issue 8, 2026*
*DOI: https://doi.org/10.38124/ijsrmt.v5i8.1642*

# INALBIO: An Absorbing Compartmental and Fairness-Constrained Bio-Inspired Modeling for Sepsis Mortality Prediction

Okpara, I. O.[1]; R. T. Abah[2]; M. O. Durojaye[3]

[1,2,3]Department of Mathematics, Faculty of Physical Science, University of Abuja, Abuja, Nigeria

Publication Date: 2026/09/05

**Abstract**

Sepsis mortality prediction requires a clear separation between clinical progression, numerical approximation, predictive error, model complexity, calibration, and subgroup disparity. This study presents INALBIO, an integrated framework combining a seven-state absorbing within-admission model, a ten-variable physiological differential system, numerical integration, bio-inspired featuresubset search, and augmented-Lagrangian fairness constraints. A retrospective descriptive analysis of MIMIC-IV identified 22,507 adult coded-sepsis admissions among 18,041 patients, with 4,395 inhospital deaths (19.5%). The first-admission cohort contained 18,041 admissions and 3,716 deaths (20.6%). Mortality was 28.2% among ICU-associated admissions and 4.8% among admissions without ICU involvement; it increased from 6.5% for sepsis without a severe-sepsis or shock modifier to 15.1% for severe sepsis and 35.2% for septic shock. The compartmental model preserves positivity and total augmented probability mass, has monotone absorbing outcomes, admits no positive endemic equilibrium, and converges to a terminal discharge–mortality equilibrium set under a recurrence threshold below one. Constructed software tests favored fourth-order Runge–Kutta (RK4) over Euler and narrowly over fourth-order Adams–Bashforth–Moulton (ABM4). A particle-swarm verification reduced 20 RK4-derived candidates to four features. In an artificial-group fairness test, the TPR disparity decreased from 0.417 to 0.056 and the maximum TPR/FPR disparity decreased from 0.421 to 0.056; however, log loss increased from 0.610 to 0.668, Brier score from 0.214 to 0.237, and AUROC decreased from 0.690 to 0.639. Thus, the constraint reduced constructed disparity at the cost of predictive performance. INALBIO provides a concise, auditable basis for future leakage-safe fitting and locked clinical evaluation.

***Keywords:*** *Sepsis; Compartmental Model; MIMIC-IV; Numerical Methods; Particle Swarm Optimization; Algorithmic Fairness; Augmented Lagrangian.*

## I. INTRODUCTION

Sepsis is a life-threatening syndrome characterized by organ dysfunction caused by a dysregulated host response to infection [1]. Its heterogeneity and rapid evolution require repeated physiological assessment rather than a purely static description [2]. Electronic health records enable dynamic risk modelling, but published sepsis models vary substantially in cohort definition, observation window, missing-data handling, calibration, and validation [5, 6]. High retrospective discrimination alone therefore does not establish numerical reliability, transportability, clinical usefulness, or equitable performance.

Mechanistic and machine-learning approaches address different parts of this problem. Compartmental differential equations provide interpretable states, transition constraints, and testable properties such as positivity and conservation. Machine-learning models can use complex longitudinal patterns for outcome prediction, but numerical error is often hidden within the modelling pipeline. Fairness is likewise commonly assessed after training, even though unequal sensitivity or false-positive burden may persist across clinically relevant groups [7, 8].

This work addresses these gaps through the Integrated Numerical–Augmented Lagrangian Bio-Inspired Optimization (INALBIO) framework. Its aim is not to claim that one solver or optimizer is universally superior, but to establish an auditable sequence that: (i)

I. O, O., Abah, R. T., & Durojaye, M. O. (2026). INALBIO: An Absorbing Compartmental and Fairness-Constrained Bio-Inspired Modeling for Sepsis Mortality Prediction. *International Journal of Scientific Research and Modern Technology*, *5*(8), 59–70. https://doi.org/10.38124/ijsrmt.v5i8.1642

models within-admission severity; (ii) represents physiological dynamics; (iii) separates trajectory error from mortality-prediction error; (iv) searches for parsimonious trajectory features; and (v) constrains prespecified subgroup disparities before a locked evaluation.

The contribution is both mathematical and methodological. A seven-state probability-flow model is derived and analysed, patient-derived coded-sepsis cohort characteristics are reported, and the numerical, optimization, and fairness components are exercised in constructed verification experiments. Throughout, patient-derived description is distinguished from software verification and from empirical analyses that remain to be completed.

➢ *Research Questions and Contribution*

The dissertation addressed six linked questions: whether the absorbing system preserves non-negativity and probability mass; whether Euler, RK4, and ABM4 differ materially on identical trajectories; whether feature-subset search improves probability loss under a common evaluation budget; whether the selected model is stable and calibrated within development data; whether prespecified TPR and FPR disparities can be constrained without unacceptable loss of utility; and whether the frozen framework improves on conventional baselines in a locked patient-level evaluation.

The present mathematical analysis resolves the first question. The numerical, optimization, and fairness experiments verify software behaviour for selected parts of the remaining questions but do not resolve them clinically. The principal supported contribution is therefore an integrated architecture that keeps the clinical-state, physiological, numerical, predictive, optimization, and fairness layers identifiable and separately auditable.

## II. MATERIALS AND METHODS

➢ *Study Design, Data Source, and Cohort*

The study uses a retrospective longitudinal computational-modelling design based on MIMIC-IV, a deidentified electronic health-record resource containing hospital and intensive-care data [3, 4]. The descriptive phenotype comprises adult admissions with qualifying ICD-coded sepsis. The full sensitivity cohort contains 22,507 admissions among 18,041 unique patients. The recommended primary analytical cohort retains the chronologically first eligible admission for each patient, yielding 18,041 independent patient-admission units.

Hospital admission is the primary time origin. ICU stays are nested care episodes rather than independent observations. The proposed prediction landmark is 12 hours after admission, and only measurements available by that time are eligible. Diagnosis codes identify the retrospective cohort but are not predictors. Death time, discharge information, length of stay, and post-landmark measurements are excluded to prevent information leakage. The primary outcome is in-hospital death after the landmark; 30- and 90-day mortality are secondary outcomes.

The dynamic physiological vector contains heart rate, respiratory rate, mean arterial pressure (MAP), temperature, oxygen saturation, lactate, creatinine, white-cell count, platelets, and total bilirubin. All preprocessing, imputation, scaling, parameter estimation, feature selection, calibration, threshold selection, and fairness settings must be fitted using development patients only.

Table 1 Candidate Physiological Variables and Aggregate Completeness.

| Variable | Unit | Observed *n* | Missing |
|---|---|---|---|
| Heart rate | beats/min | 22,060 | 2.0% |
| Respiratory rate | breaths/min | 21,964 | 2.4% |
| Mean arterial pressure | mmHg | 21,839 | 3.0% |
| Temperature | ◦C | 21,778 | 3.2% |
| Oxygen saturation | % | 22,022 | 2.2% |
| Lactate | mmol/L | 19,774 | 12.1% |
| Creatinine | mg/dL | 21,151 | 6.0% |
| White-cell count | $\times 10^9$9/L | 21,586 | 4.1% |
| Platelets | ×10 /L mg/dL | 21,555 | 4.2% |
| Total bilirubin | | 18,872 | 16.2% |

The aggregate distributions describe cohort coverage but cannot replace timestamped within-patient observations for fitting the physiological system.

➢ *Evidence Classification*

Three evidence classes are used:

- *Patient-derived descriptive evidence*: cohort counts, outcomes, and aggregate physiological summaries calculated from the supplied MIMIC-IV extracts.
- *Constructed software verification*: deterministic or synthetic trajectories, outcomes, and artificial groups used to check equations and code paths.
- *Planned patient-level validation*: timestamped fitting, real-group fairness analysis, and one locked evaluation, which were not available in the supplied project.

This distinction prevents synthetic performance from being interpreted as clinical accuracy or fairness.

➢ *Ordered INALBIO Development Workflow*

The framework follows the stages in table 2. Patient partitioning precedes every fitted transformation, while test data remain locked until all modelling and fairness choices are frozen.

Table 2 Ordered Stages of the INALBIO Methodology.

| Stage | Component | Principal operation |
|---|---|---|
| 1 | Cohort protocol | Freeze unit, landmark, outcome, exclusions, phenotype, predictors, and leakage rules. |
| 2 | Dynamic models | Estimate clinical-state and physiological models using development patients. |
| 3 | Numerical comparison | Apply Euler, RK4, and ABM4 to identical systems, grids, and horizons. |
| 4 | Feature search | Search binary masks under matched unique-evaluation budgets. |
| 5 | Model selection | Apply numerical, calibration, stability, robustness, and complexity gates. |
| 6 | Fairness extension | Fit registered subgroup constraints after freezing the base representation. |
| 7 | Final evaluation | Compare baselines and ablations once on the locked patient set. |

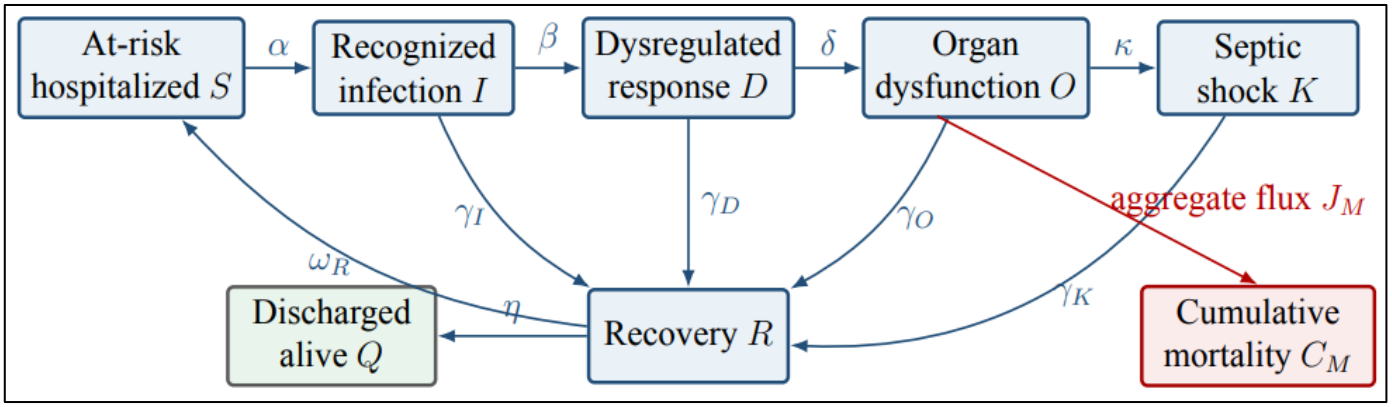


Fig 1 Within-Admission Sepsis Severity Model. Discharge Alive and Mortality are Absorbing Outcomes. The Aggregate Mortality flux JM Contains the State-Specific Outflows from I, D, O, and K.

Constructed data can test the mechanics of Stages 3–6, but only timestamped patient observations and locked evaluation can establish clinical performance.

➢ *Within-Admission Compartmental Model*

This is a within-patient severity model, not a population infection-transmission model. It assumes nonneg-ative state probabilities; mutually exclusive active states for interpretation; nonnegative locally constant or piecewise-constant rates; ordered progression $S \rightarrow I \rightarrow D \rightarrow O \rightarrow K$; recovery from every active severity state; relapse or discharge following recovery; and absorbing discharge and death. Information recorded after time $t$ is ineligible for prediction at $t$, and fitted parameters are not automatically interpreted as causal treatment effects.

For admission $i$, define

$$\mathbf{p}_i(t) = \left[S_i(t), I_i(t), D_i(t), O_i(t), K_i(t), R_i(t), Q_i(t)\right]^{\mathsf{T}},$$

Where $S$ is the at-risk hospitalized state, $I$ recognized infection, $D$ dysregulated systemic response, $O$ organ dysfunction, $K$ septic shock, $R$ recovery during the active admission, and $Q$ cumulative discharge alive. Cumulative in-hospital mortality, $C_{M,i}$, is an absorbing outcome rather than an active compartment. The flow structure is shown in figure 1.

Table 3 Transition Parameters in the Clinical-Severity Model.

| Parameter | Flow | Interpretation |
|---|---|---|
| $\alpha$ | $S \rightarrow I$ | Entry into recognized infection. |
| $\beta, \delta, \kappa$ | $I \rightarrow D \rightarrow O \rightarrow K$ | Sequential progression through dysregulation, organ dysfunction, and shock. |
| $\gamma_I, \gamma_D$ | $I, D \rightarrow R$ | Recovery from infection or dysregulated response. |
| $\gamma_O, \gamma_K$ | $O, K \rightarrow R$ | Recovery from organ dysfunction or shock. |
| $\mu_I, \mu_D$ | $I, D \rightarrow C_M$ | Mortality outflow from earlier severity states. |
| $\mu_O, \mu_K$ | $O, K \rightarrow C_M$ | Mortality outflow from organ dysfunction or shock. |
| $\omega_R$ | $R \rightarrow S$ | Relapse or return to the at-risk state. |
| $\eta$ | $R \rightarrow Q$ | Discharge-alive rate following recovery. |

With time measured in hours and all rates nonnegative, the governing system is

$$\dot{S}_i = -\alpha S_i + \omega_R R_i, \qquad \dot{I}_i = \alpha S_i - (\beta + \gamma_I + \mu_I) I_i, \tag{1}$$

$$\dot{D}_i = \beta I_i - (\delta + \gamma_D + \mu_D) D_i, \qquad \dot{O}_i = \delta D_i - (\kappa + \gamma_O + \mu_O) O_i, \tag{2}$$

$$\dot{K}_i = \kappa O_i - (\gamma_K + \mu_K) K_i, \qquad \dot{R}_i = \gamma_I I_i + \gamma_D D_i + \gamma_O O_i + \gamma_K K_i - (\omega_R + \eta) R_i, \tag{3}$$

$$\dot{Q}_i = \eta R_i, \; \dot{C}_{M,i} = \mu_I I_i + \mu_D D_i + \mu_O O_i + \mu_K K_i. \tag{4}$$

The instantaneous mortality flux is

$$J_{M,i}(t) = \mu_I I_i + \mu_D D_i + \mu_O O_i + \mu_K K_i. \tag{5}$$

For normalized initial data, the feasible region is

$$\Omega = \left\{ \mathbf{X} \in \mathbb{R}^8_{\geq 0} : S + I + D + O + K + R + Q + C_M = 1 \right\}. \tag{6}$$

➢ *Mathematical Properties*

The fixed-rate system is linear, $\dot{X} = AX$, and therefore has a unique global solution $X(t) = e^{A(t-t_0)}X(t_0)$. Because $A$ is a Metzler matrix, nonnegative initial conditions remain nonnegative. Summing equations (1), (3) and (4) gives.

$$\frac{d}{dt}(S + I + D + O + K + R + Q + C_M) = 0, \tag{7}$$

So $\Omega$ is positively invariant and every component remains in [0,1]. Moreover, $\dot{Q} \geq 0$ and $\dot{C}_M \geq 0$, hence the absorbing outcomes cannot decrease. Define total exit rates.

$$b_I = \beta + \gamma_I + \mu_I, \quad b_D = \delta + \gamma_D + \mu_D, \quad b_O = \kappa + \gamma_O + \mu_O,$$

$$b_K = \gamma_K + \mu_K, \qquad b_R = \omega_R + \eta.$$

For $\alpha > 0$ and $\eta > 0$, the terminal equilibrium set is

$$E0 = \{(0,0,0,0,0,0,q,1-q) : 0 \leq q \leq 1\}. \tag{8}$$

No positive endemic equilibrium exists because $\dot{Q} = 0$ implies $R^* = 0$, which successively implies $S^* = I^* = D^* = O^* = K^* = 0$. This is appropriate for an absorbing within-admission model and differs from a population transmission model.

The within-patient recurrence quantity is

$$\mathcal{R}_c = \frac{\omega_R}{b_R}\left[\frac{\gamma_I}{b_I} + \frac{\beta\gamma_D}{b_I b_D} + \frac{\beta\delta\gamma_O}{b_I b_D b_O} + \frac{\beta\delta\kappa\gamma_K}{b_I b_D b_O b_K}\right]. \tag{9}$$

The bracketed term is the probability of reaching recovery through one of four mutually exclusive pathways, while $\omega_R/b_R$ is the probability of relapse before discharge. Thus $0 \leq Rc \leq 1$. When $Rc < 1$, the transient-state matrix is Hurwitz and equation (8) is globally asymptotically stable relative to $\Omega$. This threshold describes recurrence within an admission; it is not an epidemiological basic reproduction number.

➢ *Sensitivity, Identifiability, and Estimation Safeguards*

For a differentiable output $G(\theta)$, the normalized forward sensitivity of $G$ to parameter $\theta_j$ is

$$\Upsilon^G_{\theta_j} = \frac{\theta_j}{G}\frac{\partial G}{\partial \theta_j}. \tag{10}$$

Define the mutually exclusive recovery-path probabilities

$$P_I = \frac{\gamma_I}{b_I}, \quad P_D = \frac{\beta\gamma_D}{b_I b_D}, \quad P_O = \frac{\beta\delta\gamma_O}{b_I b_D b_O}, \quad P_K = \frac{\beta\delta\kappa\gamma_K}{b_I b_D b_O b_K}. \tag{11}$$

Then $P_R = P_I + P_D + P_O + P_K$ and $Rc = (\omega_R/b_R)P_R$. Two immediately interpretable sensitivity indices are

$$\Upsilon^{\mathcal{R}_c}_{\omega_R} = \frac{\eta}{b_R}, \quad \Upsilon^{\mathcal{R}_c}_{\eta} = -\frac{\eta}{b_R}, \quad \Upsilon^{\mathcal{R}_c}_{\alpha} = 0. \tag{12}$$

Thus, increasing the relapse rate raises recurrence, whereas increasing discharge after recovery lowers it.

The entry rate $\alpha$ changes the timing of movement from $S$ to $I$ but not the conditional probability of eventual recovery followed by relapse represented by $Rc$. Mortality outflow rates have negative recurrence sensitivity because greater absorption into $C_M$ reduces the probability of completing a recovery–relapse cycle. These are structural model interpretations, not treatment-effect estimates.

Structural identifiability should be assessed from the observation map and the rank of the sensitivity matrix for the parameters. Practical identifiability requires profile likelihoods, dispersed initializations, parameter-correlation checks, and patient-clustered bootstrap distributions. Parameters that remain on boundaries, vary strongly across restarts, or have broad profiles should not receive clinical interpretation. Transition-rate uncertainty must also be propagated into trajectory features and mortality predictions rather than reporting only point estimates.

Parameter estimation is performed within development folds by combining physiological trajectory error, compartment-state likelihood, and mortality likelihood with regularization. Off-diagonal physiological couplings may receive an $L^1$ penalty, while mortality coefficients receive an $L^2$ penalty. The final analysis must report convergence diagnostics, admissibility checks, uncertainty intervals, and the proportion of trajectories rejected by positivity, conservation, or physiological bound checks.

➢ *Physiological and Mortality-Risk Layers*

Let $x_i(t) \in R^{10}$ contain the standardized physiological variables and $u_i(t)$ contain eligible time-varying inputs. The primary regularized physiological model is

$$\frac{d\mathbf{x}_i}{dt} = \mathbf{b} + A_x\mathbf{x}_i + B\mathbf{u}_i, \qquad \mathbf{x}_i(t_{i0}) = \mathbf{x}_{i0}, \tag{13}$$

Where off-diagonal entries of $A_x$ permit coupled physiological change. Unsupported treatment inputs are omitted rather than imputed as exposure.

The physiological state may be mapped to the seven active clinical states using a scaled softmax:

$$\tilde{p}_{ij}(t) = \frac{\exp(\mathbf{W}_j^{\top}\mathbf{x}_i + d_j)}{\sum_{k=1}^{7}\exp(\mathbf{W}_k^{\top}\mathbf{x}_i + d_k)}, qquad p_{ij}(t) = [1 - C_{M,i}(t)]\tilde{p}_{ij}(t). \tag{14}$$

Consequently, $\sum_{j=1}^{7} p_{ij} = 1 - C_{M,i}$, preserving the remaining non-mortality mass. Leakage-safe trajectory summaries $\mathbf{r}_i$ and eligible baseline covariates $\mathbf{c}_i$ enter a separate mortality model,

$$\hat{p}_i = \Pr(Y_i = 1 \mid \mathcal{H}_i(t_i^*)) = \sigma(\beta_0 + \boldsymbol{\beta}^{\top}\mathbf{r}_i + \boldsymbol{\zeta}^{\top}\mathbf{c}_i) \tag{15}$$

Where $t^*_i$ is the 12-hour landmark and $\sigma(a) = (1 + e^{-a})^{-1}$. Thus the numerical integrator approximates trajectories; it is not itself a classifier.

➢ *Numerical Approximation*

Euler, classical RK4, and ABM4 approximate the same fitted system $x^{\cdot} = f(t,x)$. Their update equations are

$$\mathbf{x}_{n+1}^{E} = \mathbf{x}_n + hf(t_n, \mathbf{x}_n), \tag{16}$$

$$\mathbf{x}^{RK4}_{n+1} = \mathbf{x}_n + \frac{h}{6}(\mathbf{k}_1 + 2\mathbf{k}_2 + 2\mathbf{k}_3 + \mathbf{k}_4), \tag{17}$$

With the usual four RK stages, and

$$\mathbf{x}^{P}_{n+1} = \mathbf{x}_n + \frac{h}{24}(55\mathbf{f}_n - 59\mathbf{f}_{n-1} + 37\mathbf{f}_{n-2} - 9\mathbf{f}_{n-3}), \tag{18}$$

$$\mathbf{x}^{C}_{n+1} = \mathbf{x}_n + \frac{h}{24}(9\mathbf{f}^{P}_{n+1} + 19\mathbf{f}_n - 5\mathbf{f}_{n-1} + \mathbf{f}_{n-2}). \tag{19}$$

Euler has global order $O(h)$, while RK4 and ABM4 have global order $O(h^4)$. Standardized trajectory RMSE, MAE, maximum error, positivity, bounds, step-size convergence, runtime, and failures are evaluated separately from mortality log loss, Brier score, discrimination, and calibration.

For component $j$ and solver $s$, standardized errors are computed after scaling by development-data dispersion:

$$RMSE^{*}_{s,j} = \left[\frac{1}{N_j}\sum_{i,n}\left(z_{i,n,j} - \hat{z}^{(s)}_{i,n,j}\right)^2\right]^{1/2}, \qquad MAE^{*}_{s,j} = \frac{1}{N_j}\sum_{i,n}\left|z_{i,n,j} - \hat{z}^{(s)}_{i,n,j}\right|. \tag{20}$$

The overall progression error is a prespecified weighted sum $E^{(s)}_{\text{prog}} = \sum_{j=1}^{10} w_j RMSE^{*}_{s,j}$, with equal primary weights $w_j = 0.1$. Observed convergence order may be checked by step halving using $p_{\text{obs}} = \log_2(\|xh - xh/2\|/\|xh/2 - xh/4\|)$.

Table 4 Theoretical and Empirical Roles of the Numerical Methods.

| Method | Local error | Global error | New evaluations | Interpretation |
|---|---|---|---|---|
| Euler | $O(h^2)$ | $O(h)$ | 1/step | Simple first-order reference; inexpensive but more sensitive to step size. |
| RK4 | $O(h^5)$ | $O(h^4)$ | 4/step | Self-starting fourth-order method used as the principal high-accuracy comparator. |
| ABM4 | $O(h^5)$ | $O(h^4)$ | 1/step[a] | Predictor–corrector method with an internal correction diagnostic. |

[a]After initialization, which requires starting values, generated here using RK4. The predictor– corrector discrepancy is an internal indicator and is not the exact global error.

➢ *Worked Numerical Calculations*

To make the numerical implementation explicit, consider the severe verification state and its derivative

$$\mathbf{x}_0 = \begin{bmatrix}112 & 28 & 65 & 38.0 & 89 & 4.05 & 1.68 & 16.6 & 140.4 & 2.05\end{bmatrix}^{\top}, \tag{21}$$

$$\mathbf{f}_0 = \begin{bmatrix}2.0 & 0.8 & -1.5 & 0.10 & -0.8 & 0.20 & 0.10 & 0.50 & -2.0 & 0.08\end{bmatrix}^{\top} \text{ h}^{-1}. \tag{22}$$

For a one-hour step, Forward Euler gives

$$\mathbf{x}_1^{E} = \mathbf{x}_0 + h\mathbf{f}_0 = \begin{bmatrix}114 & 28.8 & 63.5 & 38.10 & 88.2 & 4.25 & 1.78 & 17.1 & 138.4 & 2.13\end{bmatrix}^{\top}. \tag{23}$$

For the lactate component, let $L_0 = 4.05$ and let the four RK4 derivative evaluations be $k_1 = 0.20$, $k_2 = 0.23$, $k_3 = 0.22$, and $k_4 = 0.27$. Then

$$L_1^{RK4} = 4.05 + \frac{1}{6}[0.20 + 2(0.23) + 2(0.22) + 0.27] = 4.05 + \frac{1.37}{6} = 4.2783. \tag{24}$$

For the ABM4 lactate calculation, take $L_n = 4.10$, $f_n = 0.20$, $f_n{-}1 = 0.18$, $f_n{-}2 = 0.15$, and $f_n{-}3 = 0.12$. The predictor is

$$L_{n+1}^{P} = 4.10 + \frac{55(0.20) - 59(0.18) + 37(0.15) - 9(0.12)}{24} = 4.3021. \quad (25)$$

If the derivative at the predicted state is $f_{n+1}^{P} = 0.23$, the corrector becomes

$$L_{n+1}^{C} = 4.10 + \frac{9(0.23) + 19(0.20) - 5(0.18) + 0.15}{24} = 4.3133. \quad (26)$$

The displayed predictor–corrector discrepancy is $d_{PC} = L_{n+1}^{C} - L_{n+1}^{P} = 0.0112$; using unrounded intermediate values gives 0.01125. This difference measures the internal correction applied by ABM4 and must not be interpreted as the exact trajectory error.

Mortality predictions are evaluated separately. For $N$ outcomes $y_i \in \{0,1\}$ and probabilities $\widehat{p}_i$, the principal probability losses are

$$L_{\log} = -\frac{1}{N}\sum_{i=1}^{N}[y_i \log(\widehat{p}_i) + (1 - y_i)\log(1 - \widehat{p}_i)], \quad (27)$$

$$BS = \frac{1}{N}\sum_{i=1}^{N}(\widehat{p}_i - y_i)^2. \quad (28)$$

Calibration is examined by fitting $\operatorname{logit}\Pr(Y_i = 1) = a + b\operatorname{logit}(\widehat{p}_i)$, where ideal calibration has intercept $a = 0$ and slope $b = 1$. These predictive quantities cannot be substituted for the numerical trajectory errors in equations (23) to (26).

➢ *Feature-Subset Optimization and Fairness Constraints*

For a fixed RK4 representation, the implemented bio-inspired stage searches a nonempty binary feature mask $m \in \{0,1\}^p$. Particle swarm optimization (PSO), grey wolf optimization (GWO), and artificial bee colony (ABC) are compared as alternative search heuristics [10, 11, 12]. The verification objective combines normalized mortality loss, Brier score, discrimination loss, feature proportion, and numerical error:

$$F_{opt}(m) = [L_{\log}, BS, 1 - \text{AUROC}, \|m\|0/p, E_{num}]^{T}. \quad (29)$$

Valid future comparisons require equal numbers of unique objective evaluations, common developmentderived normalization, repeated seeds, and a frozen selection rule.

For protected group $g$ and reference group $r$, smooth equal-opportunity and false-positive-rate constraints are

$$c_g^{TPR}(\vartheta) = |\widetilde{TPR}_g(\vartheta) - \widetilde{TPR}_r(\vartheta)| - \epsilon_{TPR} \le 0, \quad (30)$$

$$c_g^{FPR}(\vartheta) = |\widetilde{FPR}_g(\vartheta) - \widetilde{FPR}_r(\vartheta)| - \epsilon_{FPR} \le 0. \quad (31)$$

For constraints $c_k(\vartheta) \le 0$, the positive-part augmented Lagrangian is

$$\mathcal{L}_A(\vartheta, \boldsymbol{\lambda}, \rho) = J(\vartheta) + \frac{\rho}{2}\sum_k \left( \left[ c_k(\vartheta) + \frac{\lambda_k}{\rho} \right]_+^2 - \left[ \frac{\lambda_k}{\rho} \right]_+^2 \right), \quad (32)$$

With $[a]_+ = \max(a,0)$ and multiplier update

$$\lambda_k^{(r+1)} = \left[ \lambda_k^{(r)} + \rho^{(r)} c_k(\vartheta^{(r+1)}) \right]_+. \quad (33)$$

Final evaluation must replace smooth surrogates with exact thresholded estimates, uncertainty intervals, calibration, and clinical consequences. No fairness tolerance is clinically universal.

➢ *Statistical and Validation Plan*

Patients, rather than admissions or ICU stays, are the split and bootstrap unit. Inner development folds select preprocessing, regularization, dynamic parameters, feature masks, calibration, thresholds, and fairness settings. Outer development folds estimate selection stability and optimism. After all choices are frozen, a locked patient-level test set is evaluated once. Required comparators include regularized logistic regression, tree-based models, a temporal model where sample size permits, unoptimized trajectory features, the feature-searched model, and the fairness-constrained extension. Reporting follows contemporary clinical prediction guidance [13].

For a test set of $n$ patients, the observed mortality proportion is

$$\widehat{\pi} = \frac{1}{n}\sum_{i=1}^{n} Y_i = \frac{D}{n}, \quad (34)$$

Where $D$ is the number of deaths. At a threshold $\tau$, predicted classes are $\widehat{Y}_i = \mathbb{I}(\widehat{p}_i \ge \tau)$, and the principal threshold statistics are

$$\widehat{\text{Sensitivity}} = \frac{TP}{TP + FN}, \qquad \widehat{\text{Specificity}} = \frac{TN}{TN + FP}, \quad (35)$$

$$\widehat{\text{PPV}} = \frac{TP}{TP + FP}, \qquad \widehat{\text{NPV}} = \frac{TN}{TN + FN}. \quad (36)$$

The alert rate is $(TP + FP)/n$, while the misclassification rate is $(FP + FN)/n$. Calibration intercept and slope are estimated from

$$\operatorname{logit}\Pr(Y_i = 1) = a + b\operatorname{logit}(\widehat{p}_i), \quad (37)$$

With ideal values $a = 0$ and $b = 1$. A slope below one indicates predictions that are too extreme, whereas a nonzero intercept indicates systematic underprediction or overprediction.

Uncertainty is estimated by resampling patients, keeping all admissions from the same individual within one bootstrap cluster. For statistic $T$ and bootstrap replicates $T^{*(1)},\ldots,T^{*(B)}$, the percentile interval is

$$CI_{95\%}^{boot}(T) = \left[Q_{0.025}\{T^{*(b)}\}_{b=1}^{B},\ Q_{0.975}\{T^{*(b)}\}_{b=1}^{B}\right]. \quad (38)$$

For a paired model comparison, $\Delta T = T_A - T_B$ is recalculated within every common bootstrap sample so that the interval reflects patient-level pairing.

Clinical utility at threshold $\tau$ is summarized by decision-curve net benefit,

$$NB(\tau) = \frac{TP(\tau)}{n} - \frac{FP(\tau)}{n}\frac{\tau}{1-\tau}. \quad (39)$$

The factor $\tau/(1-\tau)$ expresses the relative harm assigned to a false-positive alert. Net benefit is reported only over clinically meaningful thresholds and alongside the treat-all and treat-none strategies.

## III. RESULTS

➢ *Patient-Derived Cohort Description*

The full cohort comprised 22,507 admissions among 18,041 patients, including 4,395 in-hospital deaths (19.5%). Thirty-day mortality was 22.5% (5,072 deaths) and 90-day mortality was 31.4% (7,060 deaths).

Table 5 Principal Patient-Derived Descriptive Findings.

| Characteristic | Count | Value |
|---|---|---|
| All coded-sepsis admissions | 22,507 | 100% |
| Unique patients | 18,041 | — |
| In-hospital deaths | 4,395 | 19.5% |
| 30-day deaths | 5,072 | 22.5% |
| 90-day deaths | 7,060 | 31.4% |
| First-admission deaths | 3,716/18,041 | 20.6% |
| ICU-associated admissions | 14,188 | 63.0% |
| Deaths with ICU involvement | 3,994/14,188 | 28.2% |
| Deaths without ICU involvement | 401/8,319 | 4.8% |

Table 6 In-hospital Outcomes by ICD-Coded Severity.

| Severity | Admissions | Deaths | Mortality |
|---|---|---|---|
| No severe/shock modifier | 9,619 | 628 | 6.5% |
| Severe sepsis without shock | 3,816 | 576 | 15.1% |
| Septic shock | 9,072 | 3,191 | 35.2% |
| All coded sepsis | 22,507 | 4,395 | 19.5% |

The first-admission cohort contained 18,041 admissions and 3,716 in-hospital deaths (20.6%). Median age was 68 years (IQR 56–79), 54.7% of admissions were male, and 63.0% involved ICU care.

Mortality increased with coded severity (table 6), supporting distinct organ-dysfunction and shock states while not estimating timed transition rates. Sepsis was recorded as the principal diagnosis in 64.9% of admissions. Mortality was 16.4% when principal and 25.4% when secondary, a descriptive difference potentially influenced by clinical complexity and coding practice.

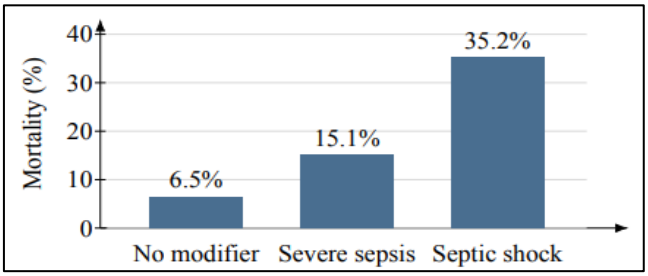


Fig 2 Patient-Derived in-Hospital Mortality by ICD-Coded Sepsis Severity. The Bars show an Unadjusted Descriptive Gradient and do not Represent Causal Effects.

Calculations for Sections 3.1 tables. Each mortality percentage is the number of deaths divided by the relevant admissions and multiplied by 100. For the full and first-admission cohorts,

$$\frac{4,395}{22,507} \times 100 = 19.527\% \approx 19.5\%, \quad \frac{3,716}{18,041} \times 100 = 20.598\% \approx 20.6\%. \quad (40)$$

Similarly, ICU-associated and non-ICU mortality are

$$\frac{3,994}{14,188} \times 100 = 28.151\% \approx 28.2\%, \quad \frac{401}{8,319} \times 100 = 4.820\% \approx 4.8\%. \quad (41)$$

The severity-specific values in table 6 follow from

$$\frac{628}{9,619}(100) = 6.529\%, \quad \frac{576}{3,816}(100) = 15.094\%, \quad \frac{3,191}{9,072}(100) = 35.174\%, \quad (42)$$

Which round to 6.5%, 15.1%, and 35.2%, respectively.The shock-to-no-modifier mortality ratio is 35.174/6.529 = 5.39, a descriptive risk ratio without covariate adjustment.

Vital-sign missingness ranged from 2.0% to 3.2%. Laboratory missingness ranged from 4.1% for whitecell count to 16.2% for bilirubin; lactate was missing in 12.1%.

ICU fields were structurally absent for the 8,319 admissions without ICU involvement and should not be imputed as though ICU care occurred.

➢ *ICU-Associated Admissions*

The 14,188 ICU-associated admissions contained 17,240 ICU stays. Total ICU time per hospital admission had a median of 3.43 days (IQR 1.66–8.64) and a mean of 7.23 days. ICU duration is an outcome of care and is therefore excluded as an admission-time predictor. Mortality varied across the first recorded ICU care unit (table 7), further demonstrating clinical heterogeneity within the coded-sepsis cohort.

Table 7 Patient-Derived Mortality by First ICU Care Unit.

| First care unit | Admissions | Deaths | Mortality |
|---|---|---|---|
| Medical ICU | 5,149 | 1,518 | 29.5% |
| Medical/Surgical ICU | 4,341 | 1,078 | 24.8% |
| Surgical ICU | 1,463 | 380 | 26.0% |
| Trauma Surgical ICU | 1,228 | 330 | 26.9% |
| Coronary Care Unit | 1,036 | 406 | 39.2% |
| Cardiac Vascular ICU | 537 | 154 | 28.7% |
| Neuro Surgical ICU | 155 | 58 | 37.4% |
| Other units | 279 | 70 | 25.1% |
| All ICU coded-sepsis admissions | 14,188 | 3,994 | 28.2% |

The highest descriptive mortality occurred in the Coronary Care Unit (39.2%) and Neuro Surgical ICU (37.4%), whereas the Medical/Surgical ICU had the lowest unit-specific estimate (24.8%). These differences are unadjusted and may reflect case mix, referral patterns, comorbidity, interventions, and small subgroup sizes; they do not measure unit quality or causal treatment effects.

Table 8 Held-Out Constructed Numerical-Solver Errors ($n$ = 60).

| Method | RMSE | MAE | Maximum error | Rank |
|---|---|---|---|---|
| Euler | $9.9999 \times 10^{-3}$ | $7.6030 \times 10^{-3}$ | $3.4105 \times 10^{-2}$ | 3 1 |
| RK4 | $7.2310 \times 10^{-5}$ | $5.0573 \times 10^{-5}$ | $2.4860 \times 10^{-4}$ | 2 |
| ABM4 | $7.2372 \times 10^{-5}$ | $5.0603 \times 10^{-5}$ | $2.4909 \times 10^{-4}$ | |

For example, Medical ICU mortality is

$$\frac{1,518}{5,149} \times 100 = 29.481\% \approx 29.5\%, \quad (43)$$

While Medical/Surgical ICU mortality is 1,078/4,341 × 100 = 24.833% ≈ 24.8%. The two highest displayed rates are

$$\frac{406}{1,036} \times 100 = 39.189\% \approx 39.2\%, \quad \frac{58}{155} \times 100 = 37.419\% \approx 37.4\%. \quad (44)$$

Summing the unit rows gives 14,188 admissions and 3,994 deaths, reproducing the overall ICU rate 3,994/14,188 × 100 = 28.151%.

➢ *Mathematical Findings*

The analysis established existence and uniqueness, positivity, boundedness, augmented mass conservation, and monotonicity of discharge and mortality probabilities. The system admits the terminal equilibrium set in equation (8), but no biologically admissible positive endemic equilibrium. The active states converge to zero when R$c$ < 1, leaving final discharge-alive and mortality probabilities that sum to one.

Adding the eight compartment equations cancels every internal flow:

$$N^{\cdot} = (-\alpha S + \omega_R R) + (\alpha S - b_I I) + (\beta I - b_D D) + (\delta D - b_O O) + (\kappa O - b_K K) + (\gamma_I I + \gamma_D D + \gamma_O O + \gamma_K K - b_R R) + \eta R + J_M = 0, \quad (45)$$

Because $b_I = \beta + \gamma_I + \mu_I$, $b_D = \delta + \gamma_D + \mu_D$, $b_O = \kappa + \gamma_O + \mu_O$, $b_K = \gamma_K + \mu_K$, $b_R = \omega_R + \eta$, and $J_M = \mu_I I + \mu_D D + \mu_O O + \mu_K K$. Hence $N(t) = N(0) = 1$. At equilibrium, $0 = Q^{\cdot} = \eta R^*$ with $\eta > 0$ gives $R^* = 0$; then $0 = S^{\cdot} = -\alpha S^*$ gives $S^* = 0$, and the cascade $0 = -b_I I^*$, $0 = -b_D D^*$, $0 = -b_O O^*$, and $0 = -b_K K^*$ yields $I^* = D^* = O^* = K^* = 0$. Thus only $Q^*$ and $C_M^*$ remain, with $Q^* + C_M^* = 1$.

➢ *Constructed Numerical Verification*

The ten-variable physiological code was checked using constructed trajectories. All fitted derivative models achieved $R^2 > 0.99999$. On 60 held-out constructed cases at a one-hour step, RK4 produced the lowest standardized errors, narrowly outperforming ABM4 and substantially outperforming Euler (table 8). These values demonstrate numerical implementation, not patient-level trajectory accuracy.

The Euler-to-RK4 RMSE ratio is

$$\frac{9.9999241 \times 10^{-3}}{7.23098 \times 10^{-5}} = 138.293, \quad (46)$$

So the RK4 RMSE is [1−(7.23098×10$^{-5}$/9.9999241×10$^{-3}$)] 100 = 99.277% lower in this constructed experiment. RK4 and ABM4 are much closer:

$$(7.23715 - 7.23098) \times 10^{-5} = 6.17 \times 10^{-8}, \quad (47)$$

And RK4 is only 0.085% lower relative to ABM4. These calculations explain why both fourth-order methods greatly outperform Euler while RK4 ranks narrowly first.

➢ *Constructed Feature-Search Verification*

The verification cohort contained 300 constructed cases split into 180 training, 60 validation, and 60 heldout cases. Twenty candidates represented the 12-hour value and 0–12-hour slope of each physiological variable. PSO selected four 12-hour variables—MAP, lactate, creatinine, and platelets—and achieved a lower overall decision score than GWO and ABC. Relative to the full RK4 baseline, held-out log loss decreased from 0.653 to 0.610, Brier score from 0.229 to 0.214, and AUROC increased from 0.673 to 0.690. The feature count decreased by 80%. However, fixed-threshold misclassification increased from 0.333 to 0.400. The result supports probabilistic software verification and parsimony, not a clinically stable four-variable signature.

Table 9 Constructed Comparison of Feature-Subset Searches for the RK4 Representation.

| **Method** | **Val. log loss** | **Val. Brier** | **Val. AUROC** | **Features** |
|---|---|---|---|---|
| Full RK4 baseline | 0.539 | 0.179 | 0.798 | 20 |
| PSO | 0.506 | 0.167 | 0.852 | 4 |
| GWO | 0.503 | 0.166 | 0.860 | 5 |
| ABC | 0.506 | 0.167 | 0.852 | 4 |
| | | | | |
| **Method** | **Test log loss** | **Test Brier** | **Test AUROC** | **Time (s)** |
| Full RK4 baseline | 0.653 | 0.229 | 0.673 | — |
| PSO | 0.610 | 0.214 | 0.690 | 1.234 |
| GWO | 0.612 | 0.215 | 0.686 | 0.739 |
| ABC | 0.610 | 0.214 | 0.690 | 4.979 |

GWO obtained the lowest validation log loss and Brier score, but PSO produced slightly lower held-out probability losses, retained one fewer feature, and shared the best held-out values with ABC at substantially lower runtime. The dissertation's scalarized decision rule therefore carried PSO forward in the verification pipeline. Because this was a constructed feature-mask experiment rather than a strictly dominance-driven clinical multi-objective study, no universal optimizer ranking is inferred.

Using the unrounded values, the PSO reduction in held-out log loss is

$$\frac{0.6525310 - 0.6096418}{0.6525310} \times 100 = 6.573\%, \quad (48)$$

And the Brier-score reduction is

$$\frac{0.2288143 - 0.2142293}{0.2288143} \times 100 = 6.374\%. \quad (49)$$

The AUROC gain is 0.689866−0.672772 = 0.017094, while the feature reduction is (20−4)/20×100 = 80%. PSO and ABC produced the same displayed held-out values, but ABC required 4.979/1.234 = 4.03 times the PSO runtime in this verification run.

➢ *Artificial-Group Fairness Verification*

An artificial binary group was used solely to test equations (31) to (33). The maximum smooth constraint violation decreased from 0.0607 to 0.0003. The final held-out comparison used 60 constructed cases, with 30 cases in each artificial group. The fairness and prediction values are reported in table 10.

Table 10 Constructed Fairness-Procedure Trade-off on the Held-Out Software-Verification Set.

| **Measure** | **Baseline** | **Constrained** | **Change** |
|---|---|---|---|
| TPR disparity point estimate | 0.417 | 0.056 | −0.361 |
| Maximum TPR/FPR disparity | 0.421 | 0.056 | −0.365 |
| Brier-score group gap | 0.082 | 0.010 | −0.072 |
| Overall log loss | 0.610 | 0.668 | +0.059 |
| Overall Brier score | 0.214 | 0.237 | +0.023 |
| Overall AUROC | 0.690 | 0.639 | −0.051 |

*Note*. Change is constrained minus baseline and was calculated from the unrounded estimates; consequently, a displayed difference may vary by 0.001 from subtraction of the three-decimal entries. Lower disparity, log loss, Brier score, and Brier-score gap are preferable, whereas higher AUROC is preferable. These are constructed software-verification results, not patient-derived fairness estimates.

For group $g$, the thresholded true-positive and false-positive rates were calculated as

$$TPR_g = \frac{TP_g}{TP_g + FN_g}, \qquad FPR_g = \frac{FP_g}{FP_g + TN_g}. \quad (50)$$

With two artificial groups, the TPR disparity was $|TPR_1 - TPR_0|$, while the maximum TPR/FPR disparity was

$$d_{\max} = \max\{|TPR_1 - TPR_0|, |FPR_1 - FPR_0|\}. \quad (51)$$

The baseline TPR gap was 0.417 and its FPR gap was 0.421, giving $d_{\max} = 0.421$. After constraint, the TPR and FPR gaps were 0.056 and 0.048, respectively, giving $d_{\max} = 0.056$. Hence, the TPR disparity fell by 0.361 and the largest rate disparity fell by 0.365.

The exact thresholded rate gaps were obtained from the two constructed groups as

$$d_{TPR}^{base} = |0.750000 - 0.333333| = 0.416667 \approx 0.417, \quad (52)$$

$$d_{FPR}^{base} = |0.611111 - 0.190476| = 0.420635 \approx 0.421, \quad (53)$$

$$d_{TPR}^{con} = |0.500000 - 0.444444| = 0.055556 \approx 0.056, \quad (54)$$

$$d_{FPR}^{con} = |0.333333 - 0.285714| = 0.047619 \approx 0.048. \quad (55)$$

The Brier-score group gaps are $|0.255352 - 0.173107| = 0.082245 \approx 0.082$ before constraint and $|0.242249 - 0.231841| = 0.010408 \approx 0.010$ afterward. The change column uses $\Delta$ = constrained − baseline; for example, $0.055556 - 0.416667 = -0.361111 \approx -0.361$, whereas the AUROC change is $0.639 - 0.690 = -0.051$.

The Brier score was $BS = n^{-1}\sum_i(\hat{p}_i - y_i)^2$, and the Brier-score group gap was the absolute difference between the two group-specific Brier scores. Its reduction from 0.082 to 0.010 indicates that the groups had more similar probability-error levels after constraint. Overall log loss measured the penalty assigned to incorrect probabilistic predictions, while AUROC measured ranking discrimination across all thresholds.

The table therefore shows a clear trade-off. The constrained model produced smaller between-group rate and Brier-score gaps, but overall probability accuracy worsened: log loss increased by 0.059 and Brier score increased by 0.023. Discrimination also declined, with AUROC decreasing by 0.051. The augmented-Lagrangian procedure achieved its constructed disparity objective by changing the risk coefficients and decision behaviour, but those changes reduced overall predictive performance. Because the groups and outcomes were constructed and subgroup counts were small, these values demonstrate implementation behaviour only and do not establish demographic fairness or clinical acceptability.

## IV. DISCUSSIONS

INALBIO connects an interpretable clinical-state model to physiological trajectories and a separately evaluated mortality predictor. This separation is important: a solver can approximate an ODE accurately while the underlying physiological model is misspecified, and a well-fitted trajectory can still produce a poorly calibrated mortality estimate. Likewise, an optimizer may improve log loss while worsening thresholded classification, as observed in the constructed PSO verification.

The patient-derived findings show substantial heterogeneity. ICU-associated mortality was nearly six times the non-ICU value, while coded septic shock was associated with markedly higher mortality than less severe coded disease. These descriptive gradients support the ordered severity structure but must not be interpreted as causal effects or timed transition estimates. The low vital-sign missingness and higher lactate and bilirubin missingness also indicate that leakage-safe handling of measurement timing and missingness will be central to patient-level fitting.

The numerical experiment behaved consistently with theory: Euler accumulated first-order error, whereas RK4 and ABM4 were closely matched. RK4 was retained for subsequent verification because it achieved the lowest held-out constructed errors. This is not sufficient to establish superiority on irregular clinical observations, where interpolation, stiffness, measurement noise, and parameteruncertainty may dominate truncation error.

The fairness experiment illustrates a central clinical-AI tension. Reducing group-rate disparities can worsen overall loss and discrimination. A defensible constraint therefore requires adequately supported real groups, a clinically justified harm model, confidence intervals, subgroup calibration, alert burden, and decision-curve analysis. Fairness should be treated as a constrained clinical design choice rather than a universal scalar score [9, 7].

➢ *Supported Contributions and Evidence Status*

The study contributes a separation between within-patient severity modelling and population transmission modelling; a seven-state probability-flow formulation with external cumulative mortality; a physiological-to-compartment mapping that preserves remaining non-mortality mass; a leakage-aware development protocol; and an evidence taxonomy distinguishing description, constructed verification, and locked validation. The status of each research objective is summarized in table 11.

Table 11 Evidence Status of the Six Research Objectives.

| Objective | Evidence class | Defensible conclusion |
|---|---|---|
| I | Theory and patient description | The absorbing model and its mathematical properties are established; timed patient-level rate fitting is pending. |
| II | Constructed numerical verification | Euler, RK4, and ABM4 code paths are checked; patient-level solver superiority is not established. |
| III | Constructed feature search | Fixed RK4 feature masks were searched; joint dynamicparameter optimization is not demonstrated. |
| IV | Not established on patient data | A verification candidate exists, but no patient-level optimal model has been selected and locked. |
| V | Artificial-group verification | The constraint mechanism reduces constructed gaps; realdemographic fairness is not established. |
| VI | Registered validation design | Nested development, baselines, calibration, utility analysis, and locked testing are specified but not executed. |

➢ *Limitations*

- The diagnosis-code phenotype and aggregate physiological summaries do not provide the complete timestamped trajectories required for patient-level transition-rate estimation.
- The retrospective single-centre design and absence of external or prospective validation limit the generalizability of the findings.

➢ *Clinical-Use Boundary and Future Work*

The next phase should extract timestamped chart and laboratory events with identifiers, item mappings, units, provenance, duplicate-resolution rules, and range checks. The 12-hour landmark cohort should then be partitioned by patient before preprocessing. Dynamic fitting, numerical comparison, feature search, calibration, and subgroup constraints must occur entirely within development data, followed by one locked patient-level evaluation with bootstrap uncertainty and conventional baselines.

## V. CONCLUSION

This study condenses a mathematical, numerical, predictive, optimization, and fairness workflow for sepsis mortality research into one auditable framework. The patient-derived evidence establishes a large heterogeneous coded-sepsis cohort and its descriptive mortality profile. The mathematical evidence establishes a positive, mass-conserving absorbing progression model with a stable terminal equilibrium set. Constructed experiments verify numerical integration, feature-search, and augmented-Lagrangian code paths while revealing important error and fairness trade-offs. The framework is ready for timestamped patient-level fitting and locked validation, but current evidence does not establish clinical predictive accuracy, demographic fairness, or deployment readiness.

## DECLARATIONS

➢ *Ethics and Data Governance*

This study used deidentified MIMIC-IV data. PhysioNet authorized the researchers' access before the data were obtained, subject to its applicable credentialing and data-use requirements. This authorization concerns access to the PhysioNet resource and should not be described as a local institutional review-board approval unless a separate institutional approval, waiver, or exemption was obtained.

➢ *Data Availability*

MIMIC-IV is available to credentialed users through PhysioNet [4]. Derived cohort specifications, item mappings, split identifiers, and analysis code should accompany the final empirical study subject to the data-use agreement.

➢ *Funding*

No funding was obtained for this research.

➢ *Competing Interests*

The author should confirm the competing-interest statement before submission.

➢ *Author Contributions*

Okpara, I. O.: conceptualization, mathematical formulation, methodology, software, analysis, and original manuscript preparation. R. T. Abah and M. O. Durojaye: supervision, methodological guidance, critical review, and manuscript revision. All authors reviewed and approved the manuscript.

## REFERENCES

[1]. 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. doi:10.1001/jama.2016.0287.

[2]. 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. doi:10.1007/s00134-021-06506-y.

[3]. Johnson AEW, Bulgarelli L, Shen L, et al. MIMIC-IV, a freely accessible electronic health record dataset. *Sci Data*. 2023;10:1. doi:10.1038/s41597-022-01899-x.

[4]. Johnson A, Bulgarelli L, Pollard T, et al. MIMIC-IV (Version 3.1). PhysioNet; 2024. doi:10.13026/kpb9-mt58.

[5]. Moor M, Rieck B, Horn M, Jutzeler CR, Borgwardt K. Early prediction of sepsis in the ICU using machine learning: a systematic review. *Front Med*. 2021;8:607952. doi:10.3389/fmed.2021.607952.

[6]. Deng H-F, Sun M-W, Wang Y, et al. Evaluating machine learning models for sepsis prediction: a systematic review of methodologies. *iScience*. 2022;25:103651. doi:10.1016/j.isci.2021.103651.

[7]. ChenRJ,WangJJ,WilliamsonDFK,etal.Algorithmic fairnessinartificialintelligenceformedicine and healthcare. *Nat Biomed Eng*. 2023;7:719–742. doi:10.1038/s41551-023-01056-8.

[8]. Ueda D, Kakinuma T, Fujita S, et al. Fairness of artificial intelligence in healthcare: review and recommendations. *Jpn J Radiol*. 2024;42:3–15. doi:10.1007/s11604-023-01474-3.

[9]. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. *Ann Intern Med*. 2021;174:866–872. doi:10.7326/M20-6261.

[10]. Kennedy J, Eberhart R. Particle swarm optimization. In: *Proceedings of the IEEE International Conference on Neural Networks*. 1995;4:1942–1948. doi:10.1109/ICNN.1995.488968.

[11]. Mirjalili S, Mirjalili SM, Lewis A. Grey wolf optimizer. *Adv Eng Softw*. 2014;69:46–61. doi:10.1016/j.advengsoft.2013.12.007.

[12]. Karaboga D, Basturk B. A powerful and efficient algorithm for numerical function optimization: artificial bee colony algorithm. *J Glob Optim*. 2007;39:459–471. doi:10.1007/s10898-007-9149-x.

[13]. 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. doi:10.1136/bmj-2023-078378.