This MICCAI paper is the Open Access version, provided by the MICCAI Society. It is identical to the accepted version, except for the format and this watermark; the final published version is available on SpringerLink.

ArdsTracker: Time-series physiological neural networks for predicting respiratory failure based on chest X-ray

Weiru Wang[1†], Jianming Wei[2†], Rosa Cornelia Jacoba Kraaijveld[2], Hong-Yu Zhou[3], Stephan Ursprung[4], Yue Sun[5], Tao Tan[5,6], Yuan Gao[2,6*], and Xin Wang[6*]

[1] Graduate School of Life Sciences, Utrecht University, Utrecht, The Netherlands
[2] University Medical Center Utrecht, Utrecht University, Utrecht, The Netherlands
[3] Tsinghua University, Beijing, China
[4] Department of Diagnostic and Interventional Radiology, Tübingen University Hospital, Tübingen, Germany
[5] Faculty of Applied Sciences, Macao Polytechnic University, Macao SAR, China
[6] Netherlands Cancer Institute, Amsterdam, The Netherlands
[†] These authors contributed equally to this work.
[*] Corresponding authors: yuangaowx@gmail.com; xinwanggy@gmail.com

Abstract. Early prediction is essential for timely intervention in acute respiratory distress syndrome (ARDS), especially within the first 12 hours, when optimizing ventilation strategy and fluid management can mitigate progression to severe hypoxemia. Chest X-ray (CXR) is non-invasive and routinely acquired in the ICU, and it has potential value for the detection, evaluation, and monitoring of ARDS. However, most prior ARDS-assessment AI approaches focused on current static status modeling as a discrete classification, without explicitly incorporating temporal dynamics or underlying physiologic constraints. In this context, we present ArdsTracker, a physiology-informed (PI) framework designed to characterize the dynamic evolution of key gas-exchange (GE) markers, i.e., the clinical indices of ARDS severity. Specifically, it forecasts future trajectories of PaO_2/FiO_2 (PF ratio) and $PaCO_2$ from a single baseline CXR. To address irregular temporal sampling across patients and the challenges of long-horizon prediction, we develop a Trajectory-Aware Dual Attention (TADA) module to highlight trajectory-relevant features. Furthermore, to enhance physiologic plausibility, we impose a gas-exchange –guided loss informed by clinical prior knowledge. The ArdsTracker was evaluated in 4,744 MIMIC-CXR and MIMIC-IV studies using 5-fold cross-validation. ArdsTracker non-invasively improves GE markers' trajectory regression, reducing MSE and MAE by 6.5% and 4.0%, respectively. It also achieves superior ARDS-related risk stratification across multiple time series, with AUCs at 3, 6, 9, and 12 hours of 0.756, 0.780, 0.774, and 0.802, respectively. Moreover, for more immediate clinical use (within 1–3 hours), point-wise predictions of the PF ratio and $PaCO_2$ exhibited significant separation between subsequent ARDS and non-ARDS events, further supporting early ICU decision-making. Code will be released publicly.

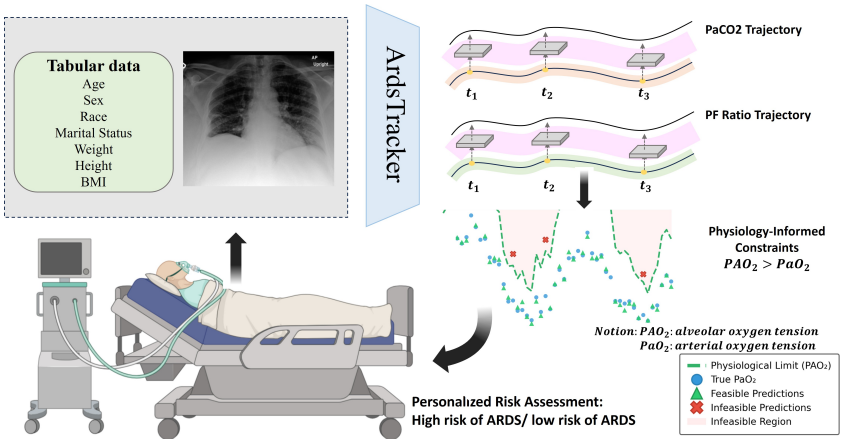

Fig. 1: Overview of ArdsTracker. Baseline chest X-ray and tabular clinical data are integrated to forecast continuous PF ratio and $PaCO_2$ trajectories under physiology-informed constraints for personalized ARDS risk assessment.

Keywords: Respiratory trajectory prediction · Time-encoding attention · ARDS · Chest X-ray

1 Introduction

Acute respiratory distress syndrome (ARDS) is a life-threatening condition marked by diffuse alveolar damage and severe hypoxemia [1]. Early assessment is critical, particularly within the first 12 hours, when optimization of ventilation strategies and fluid management can mitigate progression to severe hypoxemia [2,3]. Clinical assessment relies on gas-exchange (GE) indicators such as PaO_2/FiO_2 (PF ratio) and $PaCO_2$. However, these measurements often depend on arterial blood gas sampling or continuous monitoring and may be intermittently available in practice [4,5]. Chest X-ray (CXR) is non-invasive and routinely acquired in the ICU, providing a low-cost and effective assessment of evolving lung pathology, yet its potential to map radiographic patterns to continuous trajectories of GE remains largely unexplored.

Most existing AI studies focus on discrete ARDS classification [6,7,8], leaving the challenge of forecasting irregular, nonlinear physiological dynamics from imaging unresolved. Temporal modeling approaches for disease progression typically rely on discrete-time recurrent architectures (e.g., RNNs, LSTMs, and GRUs), but they treat time as a discrete index and often degrade under irregular sampling [9,10]. Neural ordinary differential equations (NODEs) address this by modeling hidden-state dynamics in continuous time [10,11,12]. Nevertheless, NODEs remain sensitive to noise and missing observations and suffer from an information bottleneck, where future dynamics are predominantly determined by the initial condition, limiting adaptability to downstream observations [10,13].

Moreover, purely data-driven temporal models may violate physiologically implausible predictions. Physiology-informed (PI) learning addresses this limitation by integrating domain knowledge into neural network training to encourage biologically consistent dynamics. For example, Sel et al. incorporated physiological priors as an auxiliary loss for cuffless blood pressure estimation [14].

To bridge these gaps, we propose ArdsTracker, a framework for predicting continuous ARDS-related respiratory trajectories (PF ratio and $PaCO_2$) from a single baseline chest X-ray, as illustrated in Fig. 1. ArdsTracker integrates image and textual clinical information via BiomedCLIP [15], models physiological time series with a designed Trajectory-Aware Dual Attention (TADA) module, and introduces PI constraints. Fig. 2 shows the scheme of ArdsTracker and principle of TADA. We evaluate ArdsTracker on MIMIC-IV [16] and MIMIC-CXR [17], demonstrating consistent improvements over baselines in both physiological trajectory prediction and ARDS risk stratification. Our main contributions are as follows: (1). To our knowledge, ArdsTracker is the first ARDS-oriented time-series modeling framework that non-invasively predicts continuous gas-exchange trajectories from a single baseline chest X-ray, achieving robust predictive performance over a 12-hour horizon. [18]. (2). We develop a robust baseline image-guided forecasting paradigm that models future physiological trajectories while integrating PI constraints to enhance long-horizon reliability and interpretability. (3). We demonstrate ultra-early (1–3h) clinically actionable physiological prediction, enabling effective discrimination of subsequent ARDS events for timely ICU decision-making.

2 Method

2.1 Problem Formulation

A baseline chest X-ray $\mathbf{I} \in \mathbb{R}^{H \times W}$ and clinical text $\mathbf{P}$ are encoded by BiomedCLIP [15] into visual and textual features, which are fused via MLPs and LayerNorm to produce a patient-specific context $\mathbf{c}_0$ at t_0. Given query times $\{t_i\}_{i=1}^{n}$, we predict respiratory trajectories $\hat{\mathbf{Y}} \in \mathbb{R}^{n \times 2}$ (PF ratio, $PaCO_2$). The TADA module encodes t_i and Δt_i into temporal features, concatenated with $\mathbf{c}_0$ and a time-conditioned latent $\mathbf{z}_i$ to form $\mathbf{x}_i$. The resulting sequence $\mathbf{X}$ is refined by Trajectory Dynamics Attention (TDA) and Physiological Saliency Attention (PSA), and final predictions are obtained by averaging three decoders (LSTM, time-aware LSTM (TA-LSTM), and controlled differential equation (CDE)).

2.2 Trajectory-Aware Dual Attention (TADA)

Temporal Encoding. To represent continuous time and handle irregular sampling, we encode both absolute timestamps t_i and relative intervals Δt_i using sinusoidal embeddings:

$$\mathrm{TE}(t) = [\sin(\omega_0 t), \cos(\omega_0 t), \ldots, \sin(\omega_{d/2-1} t), \cos(\omega_{d/2-1} t)]^\top, \tag{1}$$

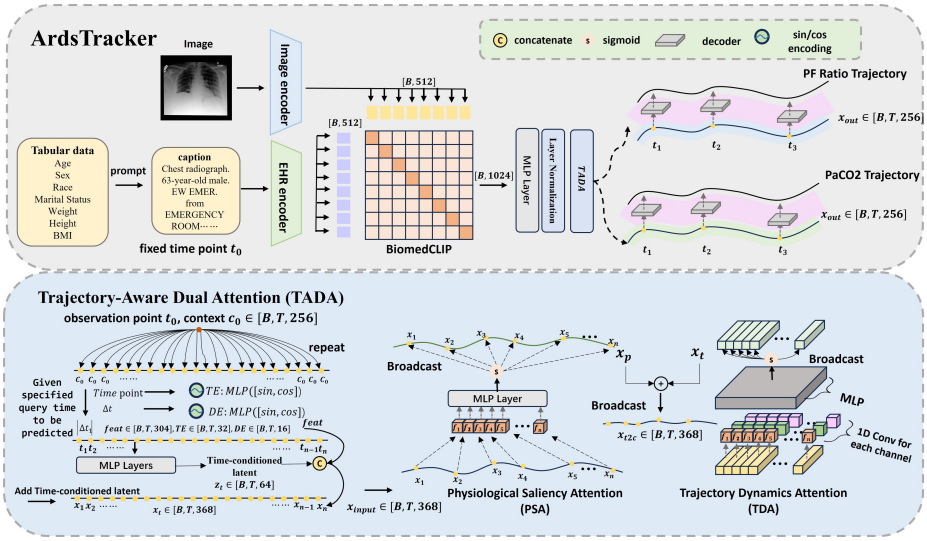

Fig. 2: Architecture of ArdsTracker. Multimodal inputs are encoded by BiomedCLIP into a context feature (i.e., $\mathbf{c}_0$), refined by the TADA module (Trajectory Dynamics Attention and Physiological Saliency Attention, and decoded into continuous gas-exchange (GE) trajectories (PF ratio and $PaCO_2$).

where frequencies $\{\omega_k\}$ are logarithmically spaced between 10^{-2} and 10^2, enabling the representation of multi-frequency temporal patterns. The resulting embeddings are projected through MLPs to obtain learned temporal representations.

Time-Conditional Latent Evolution. To capture patient-specific temporal dynamics beyond a single initial condition, we introduce a time-conditional latent refinement mechanism. For each query time t_i, the latent state is initialized from the temporal embedding and iteratively refined over L layers conditioned on the baseline context and temporal features.

$$\mathbf{z}_i^{(0)} = \mathbf{te}_i, \quad \mathbf{z}_i^{(\ell)} = f_\ell\Big([\mathbf{z}_i^{(\ell-1)}; \mathbf{c}_0; \mathbf{te}_i; \mathbf{de}_i]\Big), \quad \ell = 1, \ldots, L, \tag{2}$$

where $f_\ell(\cdot)$ denotes a nonlinear transformation. The refined latent $\mathbf{z}_i^{(L)}$ encodes a time-conditioned physiological state, and the final representation at t_i is $\mathbf{x}_i = [\mathbf{c}_0; \mathbf{te}_i; \mathbf{de}_i; \mathbf{z}_i^{(L)}] \in \mathbb{R}^{d_x}$.

Physiology-Trajectory Attention Mechanism. Respiratory trajectories exhibit coupled dynamics across both temporal evolution and physiological variables. To capture these interdependencies, we introduce concurrent attention mechanisms on the sequence $\mathbf{X} = [\mathbf{x}_1, \ldots, \mathbf{x}_n] \in \mathbb{R}^{n \times d_x}$.

Table 1: Trajectory regression performance on MIMIC-IV.

Method	PF Ratio		$PaCO_2$	
	MSE↓	MAE↓	MSE↓	MAE↓
RNN [21]	$0.885_{\pm 0.030}$	$0.756_{\pm 0.017}$	$1.008_{\pm 0.045}$	$0.698_{\pm 0.031}$
LSTM [22]	$0.855_{\pm 0.015}$	$0.741_{\pm 0.010}$	$1.028_{\pm 0.060}$	$0.709_{\pm 0.042}$
TA-LSTM [25]	$0.849_{\pm 0.027}$	$0.736_{\pm 0.015}$	$0.978_{\pm 0.029}$	$0.695_{\pm 0.016}$
CDE [24]	$0.899_{\pm 0.047}$	$0.760_{\pm 0.033}$	$1.018_{\pm 0.023}$	$0.703_{\pm 0.013}$
TCN [23]	$0.913_{\pm 0.027}$	$0.774_{\pm 0.018}$	$1.047_{\pm 0.045}$	$0.711_{\pm 0.020}$
ArdsTracker	$\mathbf{0.794}_{\pm 0.004}$	$\mathbf{0.707}_{\pm 0.005}$	$\mathbf{0.903}_{\pm 0.007}$	$\mathbf{0.661}_{\pm 0.003}$

Table 2: Risk stratification performance on MIMIC-IV.

Method	C-index↑	AUC@3h↑	AUC@6h↑	AUC@9h↑	AUC@12h↑
RNN [21]	$0.759_{\pm 0.015}$	$0.747_{\pm 0.013}$	$0.770_{\pm 0.013}$	$0.751_{\pm 0.020}$	$0.774_{\pm 0.016}$
LSTM [22]	$0.750_{\pm 0.007}$	$0.744_{\pm 0.010}$	$0.762_{\pm 0.010}$	$0.744_{\pm 0.007}$	$0.761_{\pm 0.002}$
TA-LSTM [25]	$0.762_{\pm 0.009}$	$0.753_{\pm 0.006}$	$0.772_{\pm 0.013}$	$0.757_{\pm 0.010}$	$0.771_{\pm 0.009}$
CDE [24]	$0.757_{\pm 0.005}$	$0.746_{\pm 0.005}$	$0.767_{\pm 0.005}$	$0.755_{\pm 0.010}$	$0.760_{\pm 0.004}$
TCN [23]	$0.746_{\pm 0.005}$	$0.735_{\pm 0.009}$	$0.759_{\pm 0.005}$	$0.753_{\pm 0.006}$	$0.757_{\pm 0.008}$
ArdsTracker	$\mathbf{0.774}_{\pm 0.003}$	$\mathbf{0.756}_{\pm 0.008}$	$\mathbf{0.780}_{\pm 0.005}$	$\mathbf{0.774}_{\pm 0.002}$	$\mathbf{0.802}_{\pm 0.002}$

Physiological Saliency Attention recalibrates feature importance by performing temporal squeeze through masked global average pooling:

$$\mathbf{s}_p = \frac{\sum_{i=1}^{n} m_i \, \mathbf{x}_i}{\sum_{i=1}^{n} m_i}, \tag{3}$$

where $m_i \in \{0, 1\}$ denotes the validity mask. This vector is transformed through a two-layer bottleneck MLP with reduction ratio r to generate channel recalibration weights $\mathbf{g}_p = \sigma(\mathbf{W}_2(\mathbf{W}_1\mathbf{s}_p)) \in \mathbb{R}^{d_x}$, where $\mathbf{W}_1 \in \mathbb{R}^{(d_x/r)\times d_x}$ and $\mathbf{W}_2 \in \mathbb{R}^{d_x \times (d_x/r)}$. The weight vector is broadcast across all time steps to recalibrate $\mathbf{X}$ channel-wise as $\mathbf{X}_p = \mathbf{g}_p \odot \mathbf{X}$.

Trajectory Dynamics Attention highlights critical time windows via channel squeeze, implemented by feature averaging: $\mathbf{q} = \mathbf{X}\mathbf{1}_{d_x}/d_x$, where $\mathbf{q} \in \mathbb{R}^n$. The vector $\mathbf{q}$ is passed through a 1D convolution (kernel size 3) followed by a sigmoid activation to generate temporal gating weights $\mathbf{g} = \sigma(\text{Conv1D}_3(\mathbf{q})) \in \mathbb{R}^n$, which recalibrate the sequence as $\mathbf{X}_t = [g_1\mathbf{x}_1, \ldots, g_n\mathbf{x}_n]$.

The final TADA output combines both recalibration pathways: $\hat{\mathbf{X}} = \mathbf{X}_t + \mathbf{X}_p$. This concurrent recalibration enables the model to adaptively emphasize trajectory-relevant features and temporal phases during disease progression.

2.3 Physiology-Informed Training

Alveolar Gas Exchange Constraint. We impose a physiological loss derived from the alveolar gas equation to constrain predicted GE values during training:

$$P_{\mathrm{A}}O_2 = F_{\mathrm{i}}O_2 \cdot (P_{\mathrm{atm}} - P_{\mathrm{H_2O}}) - \frac{P_{\mathrm{a}}CO_2}{R}, \tag{4}$$

with R=0.8. Arterial oxygen tension P_aO_2, estimated from the predicted PF ratio via $P_aO_2 = \mathrm{PF} \cdot F_iO_2$, is physiologically bounded by the alveolar oxygen tension P_AO_2. Predictions violating $P_aO_2 \leq P_AO_2$ are considered infeasible [19].

Physiology-Guided Loss. We translate this constraint into a differentiable training objective via adaptive reweighting. For each prediction, the constraint violation is defined as $\mathrm{viol} = \mathrm{ReLU}(P_aO_2 - P_AO_2)$ to compute a penalty

$$w = 1 + \alpha \cdot \frac{\mathrm{viol}}{\mathrm{viol} + c}, \tag{5}$$

where α controls penalty strength and c ensures numerical stability. The final loss combines masked regression with physiology-guided weighting:

$$\mathcal{L}_{\mathrm{total}} = \frac{\sum_{i,j} w_{ij}\, m_{ij}\, (\hat{y}_{ij} - y_{ij})^2}{\sum_{i,j} m_{ij}}. \tag{6}$$

where $m_{ij} \in \{0, 1\}$ is a validity mask restricting the loss to observed entries. The differentiable formulation penalizes physiologically infeasible predictions during training without imposing hard constraints at inference.

3 Experimental Settings

Dataset We evaluated ArdsTracker on trajectory prediction of $PaCO_2$, and PF ratio using 4,744 studies of 1460 patients from MIMIC-IV [16] and MIMIC-CXR [17]. Only studies with paired CXR images and physiological measurements were retained. Demographic and clinical metadata (e.g., age, sex, race, height) were collected for each subject. Using the CXR acquisition time as baseline, all subsequent PaO_2 and $PaCO_2$ measurements with timestamps were extracted. PF ratio was estimated by pairing PaO_2 with the nearest FiO_2 measurement. All CXRs were padded and resized to 512×512. $PaCO_2$ was z-score normalized, while PF ratio was log-transformed followed by z-score normalization.

Implementation All models were implemented in PyTorch. Frozen Biomed-CLIP [15] and PubMedBERT encoders [20] were used to construct a 256- dimensional multimodal context, followed by the TADA module and an average ensemble of three temporal decoders (CDE, LSTM, and TA-LSTM). Models were trained using AdamW with a learning rate of 10^{-4} and weight decay of 10^{-4}. Monte Carlo dropout (p=0.1) was applied for uncertainty estimation. Training was performed for 20 epochs with a batch size of 8. The physiology-informed loss used fixed coefficients (α=1.0, c=10.0). Evaluation used 5-fold subject-level cross-validation with bootstrap-based model selection. All experiments were run with a fixed random seed on a single NVIDIA RTX 5090 GPU.

Baselines We compared our approach against five temporal modeling baselines: LSTM [22], RNN [21],CDE [24], Temporal-Convolution Network (TCN) [23] and TA-LSTM[25].

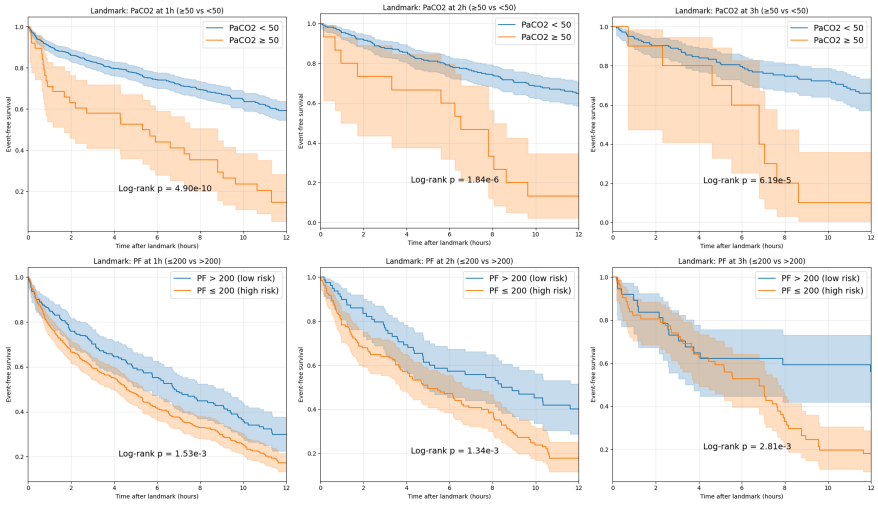

Fig. 3: Kaplan–Meier survival curves stratified by ArdsTracker's 12-hour forecasts of $PaCO_2$ and PF ratio at landmark time points of 1h, 2h, and 3h after ICU admission.

Evaluations Trajectory prediction was evaluated using MAE and MSE in normalized space. Risk stratification was assessed with the concordance index (C-index) and time-dependent AUC@t at horizons $t \in \{3, 6, 9, 12\}$ hours. All metrics are reported as mean ± standard deviation (mean ± std) across cross-validation folds.

Ablation Studies Ablation studies on CDE [24], LSTM [22], and TA-LSTM [25] backbones evaluated the impact of the TADA module and physiology-guided loss using regression (MAE, MSE) and risk stratification metrics (C-index, AUC@t).

4 Results and Discussion

Trajectory regression reults Table 1 reports trajectory regression results on MIMIC-IV. ArdsTracker achieves the lowest error across all physiological variables. For PF ratio prediction, it improves upon the best baseline (TA-LSTM), reducing MSE from 0.849 to 0.794 and MAE from 0.736 to 0.707 (relative reductions: 6.5% and 4.0%). Consistent improvements are observed for $PaCO_2$ across both metrics.

Risk stratification Table 2 summarizes risk stratification performance. ArdsTracker achieves the best C-index (0.774 vs. 0.762 for TA-LSTM) and the highest AUC@12h (0.802 vs. 0.771–0.774 for baselines). Improvements at shorter horizons (3h/6h) are more modest, reflecting the intrinsic uncertainty of early prediction from a single baseline image.

Table 3: Trajectory regression ablation results comparing backbone models with ArdsTracker (ensemble).

Backbone	TADA	PI constraints	PF Ratio		$PaCO_2$	
			MSE↓	MAE↓	MSE↓	MAE↓
LSTM [22]	✗	✗	$0.855_{\pm 0.015}$	$0.741_{\pm 0.010}$	$1.028_{\pm 0.060}$	$0.709_{\pm 0.042}$
TA-LSTM [25]	✗	✗	$0.849_{\pm 0.027}$	$0.736_{\pm 0.015}$	$0.978_{\pm 0.029}$	$0.695_{\pm 0.016}$
CDE [24]	✗	✗	$0.899_{\pm 0.047}$	$0.760_{\pm 0.033}$	$1.018_{\pm 0.023}$	$0.703_{\pm 0.013}$
LSTM [22]	✓	✗	$0.805_{\pm 0.014}$	$0.709_{\pm 0.006}$	$0.923_{\pm 0.017}$	$\mathbf{0.667}_{\pm 0.006}$
TA-LSTM [25]	✓	✗	$0.809_{\pm 0.015}$	$0.715_{\pm 0.012}$	$0.919_{\pm 0.042}$	$0.680_{\pm 0.037}$
CDE [24]	✓	✗	$0.851_{\pm 0.016}$	$0.732_{\pm 0.004}$	$0.955_{\pm 0.015}$	$0.688_{\pm 0.011}$
LSTM [22]	✓	✓	$0.801_{\pm 0.008}$	$0.710_{\pm 0.004}$	$0.930_{\pm 0.012}$	$0.670_{\pm 0.002}$
TA-LSTM [25]	✓	✓	$\mathbf{0.798}_{\pm 0.015}$	$\mathbf{0.709}_{\pm 0.014}$	$\mathbf{0.910}_{\pm 0.024}$	$0.672_{\pm 0.027}$
CDE [24]	✓	✓	$0.853_{\pm 0.021}$	$0.731_{\pm 0.008}$	$0.960_{\pm 0.021}$	$0.691_{\pm 0.011}$
ArdsTracker	✓	✓	$\mathbf{0.794}_{\pm 0.004}$	$\mathbf{0.707}_{\pm 0.005}$	$\mathbf{0.903}_{\pm 0.007}$	$\mathbf{0.661}_{\pm 0.003}$

Table 4: Risk stratification ablation results comparing backbone models with ArdsTracker (ensemble).

Backbone	TADA	PI constraints	C-index↑	AUC@3h↑	AUC@6h↑	AUC@9h↑	AUC@12h↑
LSTM [22]	✗	✗	$0.750_{\pm 0.007}$	$0.744_{\pm 0.010}$	$0.762_{\pm 0.010}$	$0.744_{\pm 0.007}$	$0.761_{\pm 0.002}$
TA-LSTM [25]	✗	✗	$0.762_{\pm 0.009}$	$0.753_{\pm 0.006}$	$0.772_{\pm 0.013}$	$0.757_{\pm 0.010}$	$0.771_{\pm 0.009}$
CDE [24]	✗	✗	$0.757_{\pm 0.005}$	$0.746_{\pm 0.005}$	$0.767_{\pm 0.005}$	$0.755_{\pm 0.010}$	$0.760_{\pm 0.004}$
LSTM [22]	✓	✗	$0.766_{\pm 0.017}$	$0.755_{\pm 0.016}$	$\mathbf{0.780}_{\pm 0.014}$	$0.763_{\pm 0.013}$	$0.787_{\pm 0.009}$
TA-LSTM [25]	✓	✗	$0.762_{\pm 0.003}$	$0.741_{\pm 0.008}$	$0.772_{\pm 0.007}$	$0.766_{\pm 0.006}$	$0.788_{\pm 0.008}$
CDE [24]	✓	✗	$0.763_{\pm 0.010}$	$0.741_{\pm 0.008}$	$0.764_{\pm 0.011}$	$0.766_{\pm 0.010}$	$0.789_{\pm 0.015}$
LSTM [22]	✓	✓	$\mathbf{0.768}_{\pm 0.013}$	$\mathbf{0.756}_{\pm 0.014}$	$0.779_{\pm 0.009}$	$\mathbf{0.769}_{\pm 0.009}$	$0.792_{\pm 0.004}$
TA-LSTM [25]	✓	✓	$0.766_{\pm 0.005}$	$0.749_{\pm 0.012}$	$0.772_{\pm 0.010}$	$0.765_{\pm 0.005}$	$0.792_{\pm 0.008}$
CDE [24]	✓	✓	$0.760_{\pm 0.008}$	$0.755_{\pm 0.012}$	$0.773_{\pm 0.009}$	$0.768_{\pm 0.004}$	$\mathbf{0.793}_{\pm 0.012}$
ArdsTracker	✓	✓	$\mathbf{0.774}_{\pm 0.003}$	$\mathbf{0.756}_{\pm 0.008}$	$\mathbf{0.780}_{\pm 0.005}$	$\mathbf{0.774}_{\pm 0.002}$	$\mathbf{0.802}_{\pm 0.002}$

Fig. 3 further shows that ArdsTracker yields clear early event discrimination (1–3 h). Using 12-hour forecasts, stratification by predicted $PaCO_2$ (≥ 50 vs. < 50 mmHg) and PF ratio (≤ 200 vs. > 200) produces significant separation of time-to-ARDS curves ($PaCO_2$: $p < 0.001$; PF ratio: $p < 0.005$), underscoring the clinical value of the predicted trajectories for early risk assessment.

Ablation Results in Tables 3 and 4 demonstrate complementary contributions of both components across all backbones. LSTM benefits most from the TADA module, achieving the largest PF ratio MSE reduction (0.855→0.805), while PI constraints yield more consistent gains in risk stratification than in regression metrics alone. Removing either component leads to degraded performance, confirming their complementary roles in trajectory modeling and risk prediction.

5 Conclusion

In this study, we present ArdsTracker, a non-invasive, physiology-informed framework for forecasting continuous respiratory trajectories (PF ratio and $PaCO_2$) from a single baseline chest X-ray. By integrating the TADA module with a

physiology-informed loss, ArdsTracker consistently improves trajectory regression and ARDS event discrimination across all baselines, with ablation studies confirming their complementary contributions. This study is limited by its retrospective single-center design. Future work will focus on multi-center validation, inference-time physiological constraints, and broader cardiopulmonary modeling.

Prospects of Application: ArdsTracker may support early ICU risk stratification by forecasting respiratory trajectories from routine baseline chest X-rays, enabling non-invasive monitoring of PF ratio and $PaCO_2$ trends and ultra-early identification of patients at risk of subsequent ARDS events.

References

1. Ranieri, V.M., Rubenfeld, G.D., Thompson, B.T., Ferguson, N.D., Caldwell, E., Fan, E., Camporota, L., Slutsky, A.S.: Acute respiratory distress syndrome: The Berlin Definition. JAMA **307**, 2526–2533 (2012)
2. Mercat, A., Richard, J.-C.M., Vielle, B., et al.: Positive end-expiratory pressure setting in adults with acute lung injury and acute respiratory distress syndrome: a randomized controlled trial. JAMA **299**, 646–655 (2008)
3. Duggal, A., Camporota, L.: Sequencing interventions in ARDS: the critical role of timing and order in standardized management. Intensive Care Med. **50**(7), 1133–1136 (2024)
4. Swenson, K.E., Swenson, E.R.: Pathophysiology of acute respiratory distress syndrome and COVID-19 lung injury. Critical Care Clinics **37**, 749–776 (2021)
5. Munshi, L., Mancebo, J., Brochard, L.: Noninvasive respiratory support for adults with acute respiratory failure. New England Journal of Medicine **387**, 1688–1698 (2022)
6. Rajpurkar, P., Irvin, J., Zhu, K., Yang, B., Mehta, H., Duan, T., Ding, D., Bagul, A., Ball, R.L., Langlotz, C., Lungren, M.P., Ng, A.Y.: CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning. arXiv preprint arXiv:1711.05225 (2017)
7. Cohen, J.P., Dao, L., Roth, K., Morrison, P., Bengio, Y., Abbasi, A.F., Shen, B., Mahsa, H.K., Ghassemi, M., Li, H., Duong, T.Q.: Predicting COVID-19 pneumonia severity on chest X-ray with deep learning. Cureus **12**, e9448 (2020)
8. Sjoding, M.W., Hofer, T.P., Co, I., Courey, A., Cooke, C.R., Iwashyna, T.J., et al.: Deep learning to detect acute respiratory distress syndrome on chest radiographs: A retrospective study with external validation. The Lancet Digital Health **3**(6), e340–e348 (2021)
9. Jia, N., Jia, T., Zhang, Z.: A residual GRU method with deep cross fusion for Alzheimer's disease progression prediction using missing variable-length time series data. Biomedical Signal Processing and Control **102**, 107253 (2025). https://doi.org/10.1016/j.bspc.2024.107253
10. Rubanova, Y., Chen, T.Q., Duvenaud, D.K.: Latent ordinary differential equations for irregularly-sampled time series. In: Wallach, H., Larochelle, H., Beygelzimer, A., d'Alché-Buc, F., Fox, E., Garnett, R. (eds.) Advances in Neural Information Processing Systems 32, pp. 5320–5330 (2019)
11. Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations. In: Advances in Neural Information Processing Systems, pp. 6571–6583 (2018)

12. Park, W., Ahn, S., Hernandez, M., Maniega, S., Park, J.: BrainODE: Neural shape dynamics for age- and disease-aware brain trajectories. In: Advances in Neural Information Processing Systems (NeurIPS) (2023)
13. De Brouwer, E., Simm, J., Arany, A., Moreau, Y.: GRU-ODE-Bayes: Continuous modeling of sporadically-observed time series. arXiv preprint arXiv:1905.12374 (2019)
14. Sel, K., Mohammadi, A., Pettigrew, R.I., et al.: Physics-informed neural networks for modeling physiological time series for cuffless blood pressure estimation. npj Digital Medicine **6**, 110 (2023). https://doi.org/10.1038/s41746-023-00853-4
15. Zhang, S., Xu, Y., Usuyama, N., Xu, H., Bagga, J., Tinn, R., Preston, S., Rao, R., Wei, M., Valluri, N., Wong, C., Tupini, A., Wang, Y., Mazzola, M., Shukla, S., Crabtree, B., Piening, B., Bifulco, C., Lungren, M.P., Naumann, T., Wang, S., Poon, H.: BiomedCLIP: A multimodal biomedical foundation model pretrained from fifteen million scientific image-text pairs. arXiv preprint arXiv:2303.00915 (2023). https://doi.org/10.48550/arXiv.2303.00915
16. Johnson, A.E.W., Bulgarelli, L., Pollard, T.J., et al.: MIMIC-IV, a freely accessible electronic health record dataset. Scientific Data **10**, 1 (2023). https://doi.org/10.1038/s41597-022-01899-x
17. Johnson, A.E.W., Pollard, T.J., Berkowitz, S.J., et al.: MIMIC-CXR, a large publicly available database of labeled chest radiographs. Scientific Data **6**, 317 (2019). https://doi.org/10.1038/s41597-019-0322-0
18. Tasaka, S., Ohshimo, S., Takeuchi, M., et al.: ARDS Clinical Practice Guideline 2021. Journal of Intensive Care **10**, 32 (2022). https://doi.org/10.1186/s40560-022-00615-6
19. West, J.B.: Respiratory physiology: The essentials. 9th edn. Lippincott Williams & Wilkins, Philadelphia (2012)
20. Gu, Y., Tinn, R., Cheng, H., Lucas, M., Usuyama, N., Liu, X., Naumann, T., Gao, J., Poon, H.: Domain-specific language model pretraining for biomedical natural language processing. In: Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics (ACL), pp. 5893–5907 (2021)
21. Elman, J.L.: Finding structure in time. Cognitive Science **14**(2), 179–211 (1990)
22. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation **9**(8), 1735–1780 (1997)
23. Bai, S., Kolter, J.Z., Koltun, V.: An empirical evaluation of generic convolutional and recurrent networks for sequence modeling. arXiv preprint arXiv:1803.01271 (2018)
24. Kidger, P., Morrill, J., Foster, J., Lyons, T.: Neural controlled differential equations for irregular time series. In: Advances in Neural Information Processing Systems, pp. 6694–6705 (2020)
25. Baytas, I.M., Xiao, C., Zhang, X., Wang, F., Jain, A.K., Zhou, J.: Patient subtyping via time-aware LSTM networks. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 65–74 (2017)