**Table S-7.** Calibration metrics for RNN predictions at all timesteps of all test set episodes.

|  |  |
| --- | --- |
| **Metric** | **Value** |
| Brier Score | 0.054 |
| Hosmer-Lemeshow |  |
| C\* Statistic | 24,681 |
| p-value | 0.00 |
| Cox Line Regression |  |
| Intercept | -5.378 [-5.395, -5.361] |
| Slope | 0.733 [0.730, 0.736] |
| Chi-squared | 15.66 |
| p-value | 0.00 |

The Hosmer-Lemeshow metrics were computed using 50 degrees of freedom (DOF), while the Cox regression metrics with logit transformation were computed using 2 DOF.