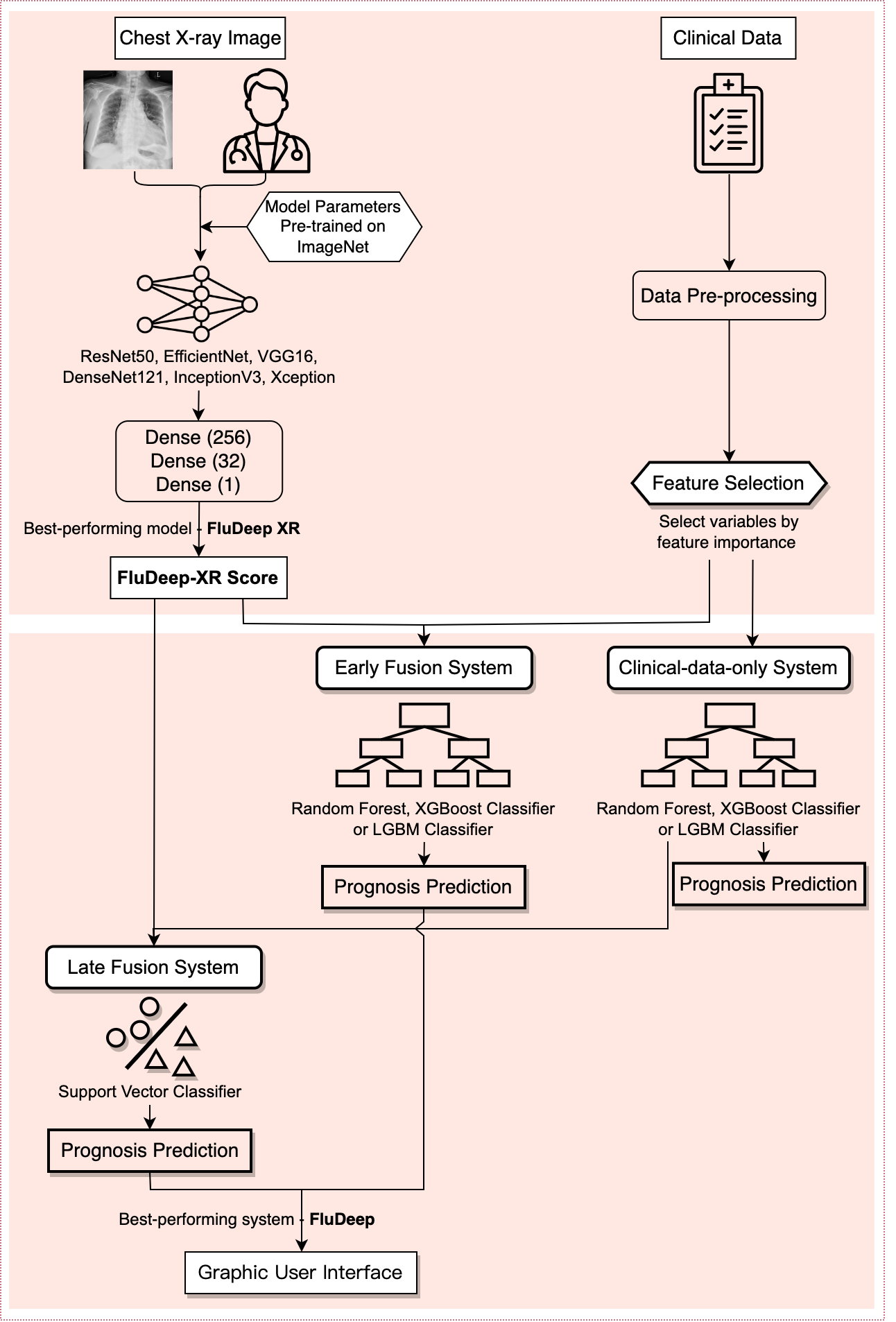


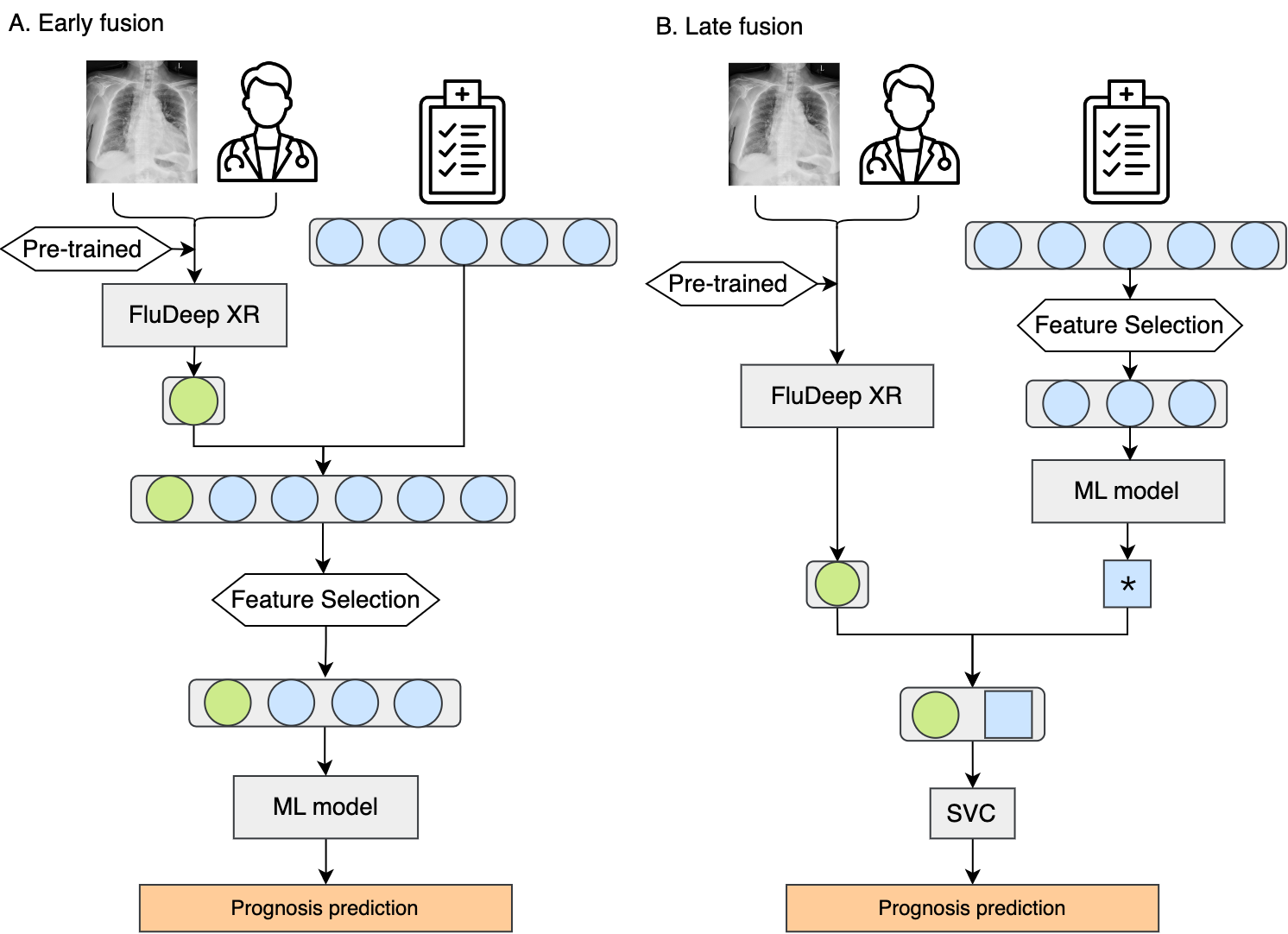
* Figure 1. The flowchart of patient selection criteria and study design

| ***1*** | ***2*** | ***3*** | ***4*** | ***5*** |
| --- | --- | --- | --- | --- |
|  |  |  |  |  |

* Table 1. Criteria of the severity scoring system of the chest X-rays. 1 for normal lung; 2 for hyperinflation and/or bronchial wall thickening; 3 for focal alveolar consolidation limited in a single lobe; 4 for multifocal consolidation; 5 for diffuse alveolar consolidation and/or pleural effusion involving the upper lobe on at least one lung.



* Figure 2. The framework of developing prognosis prediction systems for influenza patients.



* Figure 3. The illustration of early fusion and late fusion model in this study. \* indicates the prediction of patient severity without using the chest X-ray images. SVC stands for support vector classifier, ML stands for machine learning.

|  | Total  (N=765) | Survivors  (n=723) | Non-Survivors  (n=42) | P-value  (T-test) |
| --- | --- | --- | --- | --- |
| Demographics | | | | |
| Gender |  |  |  | 0.5041 |
| Male (%) | 408 (53%) | 25 (60%) | 383 (53%) |  |
| Age (years) | 65.9 ± 18.2 | 65.2 ± 18.3 | 78.3 ± 12.8 | **<0.0001** |
| BMI (kg/m2) | 23.4 ± 4.6 | 23.4 ± 4.6 | 22.6 ± 4.9 | 0.2870 |
| Comorbidities | | | | |
| Cardiovascular disease (%) | 192 (25%) | 184 (25%) | 8 (19%) | 0.4253 |
| Hypertension (%) | 338 (44%) | 312 (43%) | 26 (62%) | **0.0265** |
| COPD (%) | 107 (14%) | 97 (13%) | 10 (24%) | 0.0971 |
| Diabetes (%) | 190 (25%) | 180 (25%) | 10 (24%) | 0.9858 |
| Chronic liver disease (%) | 134 (18%) | 127 (18%) | 7 (17%) | 0.7531 |
| Chronic kidney disease (%) | 152 (10%) | 143 (20%) | 9 (21%) | 0.9505 |
| Malignant tumor (%) | 187 (24%) | 175 (24%) | 12 (29%) | 0.1316 |
| HIV (%) | 18 (2.4%) | 16 (2.2%) | 2 (4.8%) | 0.5920 |
| Vital signs | | | | |
| Respiratory rate (BPM) | 22.3 ± 11.6 | 22.2 ± 11.6 | 24.0 ± 11.7 | 0.3036 |
| DBP (mmHg) | 72.8 ± 13.6 | 73.0 ± 13.5 | 69.2 ± 14.6 | 0.0752 |
| Laboratory markers | | | | |
| Hemoglobin (g/dL) | 11.7 ± 2.4 | 11.7 ± 2.4 | 10.6 ± 2.2 | **0.0040** |
| Hematocrit (%) | 33.3 ± 6.5 | 33.5 ± 6.5 | 29.8 ± 5.7 | **0.0002** |
| Lymphocyte percentage (%) | 14.9 ± 11.6 | 15.2 ± 11.7 | 9.0 ± 7.8 | **0.0007** |
| Platelet count (10³/uL) | 200.9 ± 105.9 | 201.0 ± 105.5 | 200.5 ± 114.2 | 0.9766 |
| Lactic acid (mmol/L) | 2.4 ± 3.2 | 2.4 ± 3.3 | 2.3 ± 1.4 | 0.8510 |
| Glucose (mg/dL) | 152.4 ± 72.8 | 151.7 ± 71.1 | 164.7 ± 97.5 | 0.2612 |
| pH of arterial blood | 7.4 ± 0.1 | 7.4 ± 0.1 | 7.4 ± 0.1 | 0.2257 |
| PO2 (mmHg) | 57.4 ± 35.0 | 57.0 ± 33.6 | 62.8 ± 54.1 | 0.3009 |
| Bicarbonate (mmol/L) | 23.2 ± 4.3 | 23.2 ± 4.2 | 22.5 ± 6.1 | 0.3457 |
| BUN (mg/dL) | 31.6 ± 24.2 | 31.2 ± 23.6 | 37.8 ± 32.5 | 0.0866 |
| eGFR (mL/min/1.73m²) | 82.9 ± 49.8 | 83.1 ± 48.8 | 80.3 ± 65.8 | 0.7264 |
| INR | 1.1 ± 0.2 | 1.1 ± 0.2 | 1.1 ± 0.2 | 0.6224 |
| CRP (mg/dL) | 8.6 ± 6.3 | 8.4 ± 6.2 | 12.0 ± 5.5 | **0.0003** |
| Troponin T (ng/L) | 67.3 ± 186.2 | 67.4 ± 189.7 | 66.5 ± 111.9 | 0.9770 |
| Chest X-Ray Severity Score | | | | |
| DeepFlu-XR Score | 3.03 ± 1.2 | 3.0 ± 1.2 | 3.9 ± 1.1 | **<0.0001** |
| 1 |  | 202 (27.9%) | 7 (16.7%) |  |
| 2 |  | 114 (15.8%) | 1 (2.4%) |  |
| 3 |  | 141 (19.5%) | 3 (7.1%) |  |
| 4 |  | 126 (17.4%) | 10 (23.8%) |  |
| 5 |  | 140 (19.4%) | 21 (50.0%) |  |

|  |  | **Total**  **(N = 765)** | **Survivors**  **(n = 723)** | **Non-Survivors**  **(n = 42)** | **P-value**  **(T-test)** |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Demographics** | | | | |
| Gender | | | | 0.5041 |
| Male | 408 (53%) | 25 (60%) | 383 (53%) |  |
| Female | 357 (47%) | 17 (40%) | 340 (47%) |  |
| Age, years | 65.9 ± 18.2 | 65.2 ± 18.3 | 78.3 ± 12.8 | <0.0001 |
| BMI | 23.38 ± 4.62 | 23.42 ± 4.60 | 22.64 ± 4.90 | 0.2870 |
| **Vital signs** | | | | |
| Respiratory rate (BPM) | 22.3 ± 11.6 | 22.2 ± 11.6 | 24.0 ± 11.7 | 0.3036 |
| Diastolic blood pressure  (mmHg) | 72.8 ± 13.6 | 73.0 ± 13.5 | 69.2 ± 14.6 | 0.0752 |

|  | **Laboratory markers** | | | | |  |
| --- | --- | --- | --- | --- | --- | --- |
| Hemoglobin (g/dL) | 11.67 ± 2.42 | 11.73 ± 2.42 | 10.62 ± 2.23 | 0.0040 |
| Hematocrit (%) | 33.31 ± 6.47 | 33.52 ± 6.46 | 29.75 ± 5.65 | 0.0002 |
| Lymphocyte percentage  (%) | 14.89 ± 11.63 | 15.23 ± 11.72 | 9.03 ± 7.83 | 0.0007 |
| Platelet count (10³/uL) | 200.92 ± 105.93 | 200.95 ± 105.51 | 200.45 ± 114.16 | 0.9766 |
| Lactic acid (mmol/L) | 2.35 ± 3.18 | 2.35 ± 3.25 | 2.26 ± 1.36 | 0.8510 |
| Glucose (mg/dL) | 152.44 ± 72.77 | 151.73 ± 71.10 | 164.71 ± 97.46 | 0.2612 |
| pH of arterial blood | 7.40 ± 0.06 | 7.40 ± 0.06 | 7.39 ± 0.08 | 0.2257 |
| PO2 (mmHg) | 57.36 ± 35.00 | 57.04 ± 33.59 | 62.793 ± 54.05 | 0.3009 |
| Bicarbonate (mmol/L) | 23.15 ± 4.29 | 23.19 ± 4.17 | 22.54 ± 6.13 | 0.3457 |
| Blood urea nitrogen  (mg/dL) | 31.59 ± 24.15 | 31.23 ± 23.55 | 37.80 ± 32.52 | 0.0866 |
| eGFR (mL/min/1.73m²) | 82.93 ± 49.82 | 83.08 ± 48.78 | 80.31 ± 65.83 | 0.7264 |
| INR | 1.10 ± 0.21 | 1.09 ± 0.21 | 1.11 ± 0.18 | 0.6224 |
| CRP (mg/dL) | 8.59 ± 6.25 | 8.39 ± 6.24 | 11.98 ± 5.48 | 0.0003 |
| Troponin T (ng/L) | 67.31 ± 186.22 | 67.36 ± 189.70 | 66.51 ± 111.93 | 0.9770 |
|  | **Total**  **(N = 765)** | **Survivors**  **(n = 723)** | **Non-survivors**  **(n = 42)** | **P-value**  **(T-test)** |
| **Comorbidities** | | | | |
| Cardiovascular disease | 192 (25%) | 184 (25%) | 8 (19%) | 0.4253 |
| Hypertension | 338 (44%) | 312 (43%) | 26 (62%) | 0.0265 |
| COPD | 107 (14%) | 97 (13%) | 10 (24%) | 0.0971 |
| Diabetes | 190 (25%) | 180 (25%) | 10 (24%) | 0.9858 |
| Chronic liver disease | 134 (18%) | 127 (18%) | 7 (17%) | 0.7531 |
| Chronic kidney disease | 152 (10%) | 143 (20%) | 9 (21%) | 0.9505 |
| Malignant tumor | 187 (24%) | 175 (24%) | 12 (29%) | 0.1316 |
| HIV | 18 (2.4%) | 16 (2.2%) | 2 (4.8%) | 0.5920 |

|  | **CXR image severity score, each patient’s first chest x-ray image (within 24 hours of presentation)** | | | | |  |
| --- | --- | --- | --- | --- | --- | --- |
| DeepFlu-XR Score | 3.03 ± 1.18 | 2.98 ± 1.17 | 3.91 ± 1.05 | <0.0001 |
| 1 |  | 202 (27.9%) | 7 (16.7%) |  |
| 2 |  | 114 (15.8%) | 1 (2.4%) |  |
| 3 |  | 141 (19.5%) | 3 (7.1%) |  |
| 4 |  | 126 (17.4%) | 10 (23.8%) |  |
| 5 |  | 140 (19.4%) | 21 (50.0%) |  |

* Table 2. The characteristics of flu patients in this study. The patients are stratified by 30-day survival.

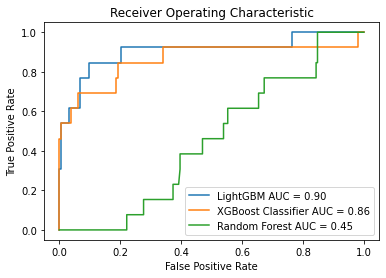
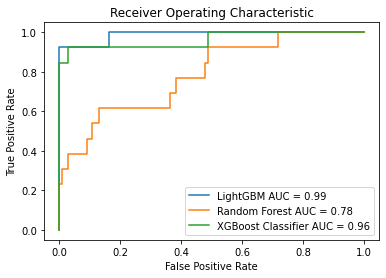
|  | **Models\MSE** | **Non-pretrained model** | **ImageNet pretrained model** |  |
| --- | --- | --- | --- | --- |
|  | **ResNet50** | 2.3708 | 2.3005 |  |
|  | **EfficientNet** | 2.2516 | 2.3904 |  |
|  | **VGG16** | 2.3118 | 2.3114 |  |
|  | **DenseNet121** | 2.3114 | 2.3146 |  |
|  | **InceptionV3** | 2.4594 | 1.6313 |  |
|  | **Xception** | 2.7874 | 1.1390 |  |
|  | | | | |

* Table 3. The performance of different deep learning-based chest-X ray severity scoring systems with and without pre-training on the validation dataset. MSE stands for Mean Squared Error.

|  |  | **AUC** | **Accuracy** | **Sensitivity** | **Specificity** | **Precision** | **F1** | **P-value** |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Random Forest** | | | | | | |  |
| Clinical-data-only system | 0.729 | 0.787 | 0.539 | 0.802 | 0.140 | 0.222 | - |
| Early fusion system | 0.816 | 0.791 | 0.769 | 0.793 | 0.182 | 0.294 | 0.7937 |
| Late fusion system | 0.785 | 0.944 | 0.000 | 1.000 | 0.000 | 0.000 | 0.0009 |
| **XGBoost Classifier** | | | | | | |  |
| Clinical-data-only system | 0.818 | 0.944 | 0.308 | 0.982 | 0.500 | 0.381 | - |
| Early fusion system | 0.861 | 0.961 | 0.539 | 0.986 | 0.700 | 0.609 | 0.0932 |
| Late fusion system | 0.960 | 0.991 | 0.846 | 1.000 | 1.000 | 0.917 | 0.0278 |
| **LGBM Classifier** | | | | | | |  |
| Clinical-data-only system | 0.859 | 0.948 | 0.308 | 0.986 | 0.571 | 0.400 | - |
| Early fusion system | 0.904 | 0.970 | 0.539 | 0.995 | 0.875 | 0.667 | 0.0742 |
| Late fusion system | 0.988 | 0.991 | 0.846 | 1.000 | 1.000 | 0.917 | 0.0037 |
|  |  |  |  |  |  |  |  |  |  |

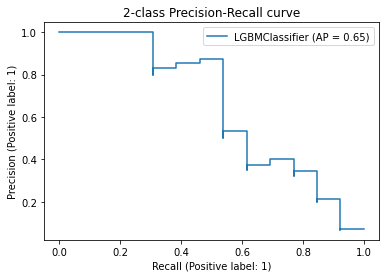
* Table 4. The results of the clinical-data-only model, early fusion model and late fusion flu prognosis prediction model. The p-values of early fusion models or late fusion models are calculated by comparing them to the clinical-data-only model built with the same ML model.

A B

****

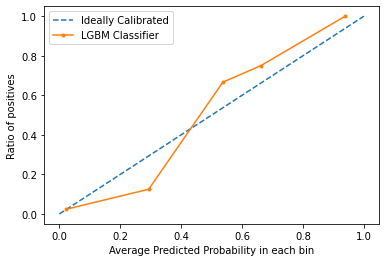
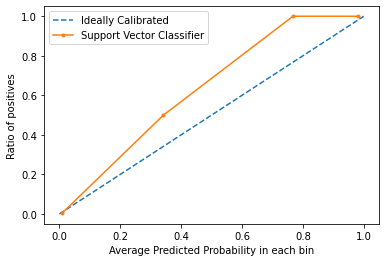
* Figure 4. The Receiver Operating Characteristic (ROC) curve of the fusion models incorporates both chest X-ray severity and clinical features. 4A, the early fusion model; 4B, the late fusion model.

A B

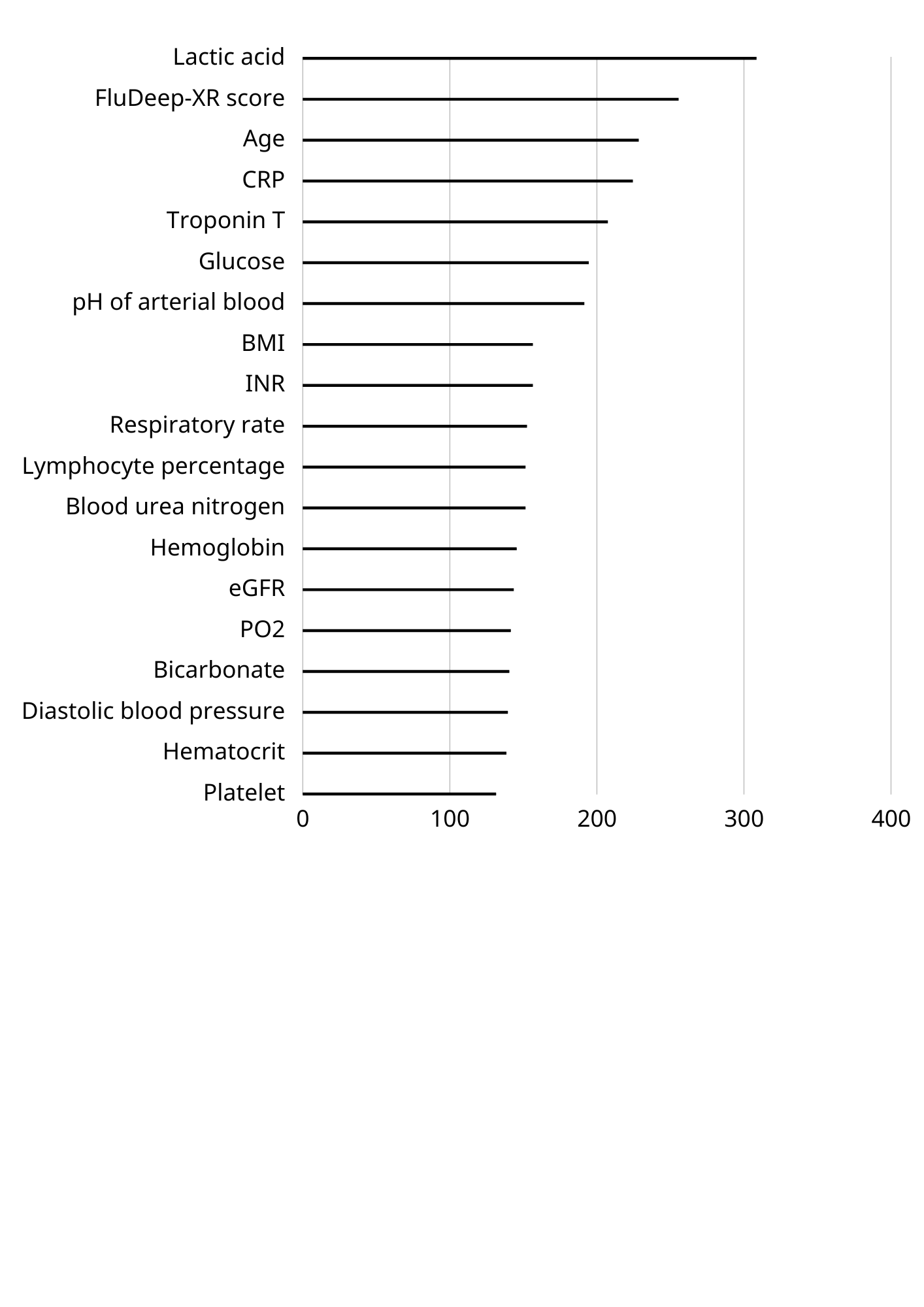
****

* Figure 5. The precision-recall curve of LightGBM-based early and late fusion prediction models. 5A, the early fusion model; 5B, the late fusion model. AP stands for average precision.

A B

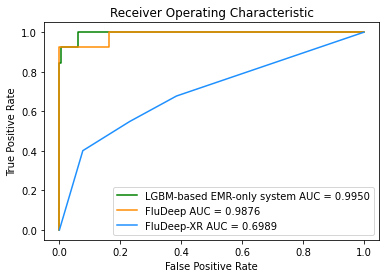
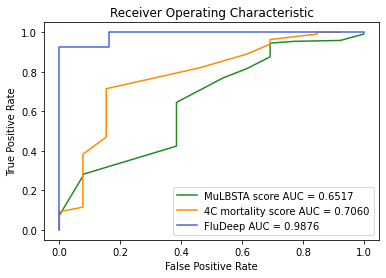
****

* Figure 6. The calibration plot for LightGBM-based early and late fusion prediction systems. 6A, the early fusion model; 6B, the late fusion model.



* Figure 7. The importance of features ranked by LightGBM. CRP: C-reactive protein, BMI: Body Mass Index, INR: international normalized ratio, eGFR: estimated glomerular filtration rate, PO2: arterial partial pressure of oxygen.

A B

* Figure 8. Comparison of Receiver Operating Characteristic (ROC) curve of Deepflu and other prediction systems.   
  8A Comparison of the ROC curve of LGBM-based clinical-data-only systems, FluDeep-XR, FluDeep. The FluDeep had higher AUC than LGBM-based clinical-data-only systems and FluDeep-XR (AUROC 0.9876 versus 0.8589 and 0.6989, p-value 0.0037 and < 0.0001).   
  8B Comparison of the ROC curve of MuLBSTA score, 4c-mortality score, and FluDeep on our patients. The FluDeep had higher AUC than MuLBSTA score and 4c-mortality score (AUROC 0.9876 versus 0.6517 and 0.7060, p-value 0.0001 and 0.0074).