medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Deep learning-based prognosis prediction among preeclamptic pregnancies using electronic health record data

Xiaotong Yang[1], Hailey K Ballard[2], Aditya D Mahadevan[4], Xu Ke[2], David G Garmire[5], Elizabeth S Langen[6], Dominick J Lemas[2,3], Lana X Garmire[1*]

1 Department of Computational Medicine and Bioinformatics, University of Michigan, Ann Arbor, Michigan, United States of America

2 Department of Health Outcomes and Biomedical Informatics, University of Florida, Gainesville, Florida, United States of America

3 Department of Obstetrics & Gynecology, University of Florida College of Medicine, Gainesville, Florida

4 Department of Physiology and Functional Genomics, University of Florida, Gainesville, FL, United States

5 Department of Electronical Engineering and Computer Science, University of Michigan, Ann Arbor, Michigan, United States of America

6 Department of Obstetrics and Gynecology, University of Michigan, Ann Arbor, Michigan, United States of America

*: Corresponding author email: lgarmire@umich.edu

NOTE: This preprint reports new research that has not been certified by peer review and should not be used to guide clinical practice.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

ABSTRACT

Background

Preeclampsia (PE) is one of the leading factors in maternal and perinatal mortality and morbidity worldwide. The only cure for PE to date is to deliver the placenta and stop gestation. However, the timing of delivery among PE patients is essential to minimize the risk of severe maternal morbidities, and at the same time ensure the survival of the baby.

Methods

In this study, we constructed a series of deep learning-based models to predict the prognosis, or the time to delivery, since the initial diagnosis of PE using electronic health record (EHR) data. We extracted and processed 1578 pregnancies in Michigan Medicine at the University of Michigan in 2015-2021 as the discovery cohort. Using the Cox-nnet v2 algorithm, we built the baseline model with EHR information prior to diagnosis, as well as the full model including baseline information and lab testing results and vital signs at the time of diagnosis. We evaluated the models using the C-index and log-rank p-values in KM survival curves, using both 20% testing data of the Michigan cohort, as well as 1177 PE pregnancy EHR data from the Medical Center of the University of Florida.

Results

The baseline prognosis model for time to delivery since PE diagnosis achieved C-index values of 0.75 and 0.72 on the training and testing set respectively. While the full model reached C-indices of 0.77 and 0.74 in the same training and testing sets. Both models performed better than their Cox-PH model counterparts. The seven most important features in the baseline model in descending order were diagnosis gestational age, severe PE, past PE, age, parity, gravidity, and

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

uncomplicated diabetes. Meanwhile, 14 most important features were selected and interpreted in the full model, including diagnosis gestational age, parity, severe PE, past PE, features in lab tests (white blood cell, platelet, and red blood cell counts, AST value), min respiratory rate, and features measuring blood pressure (minimum, mean and standard deviation of systolic blood pressure, and maximum and standard deviation of diastolic blood pressure).

Conclusion

The time to delivery predicting models provide clinicians valuable tools and options to quantify the delivery risks and make better decisions on the optimal delivery time of PE patients at the time of diagnosis. Implementation of these actionable models into PE clinical care practice is expected to significantly improve the management of PE patients.

INTRODUCTION

Preeclampsia (PE), characterized by hypertension with proteinuria during gestation, is a severe pregnancy complication that affects about 2% to 8% of all pregnancies worldwide. It's one of the worldwide leading causes of maternal and perinatal mortality and morbidity[1,2]. PE is a heterogenous complication and consists of various subtypes[3]. Based on the severity, PE patients can be classified as severe PE and mild PE; based on onset gestational age, PE occurring before 34 weeks of gestation is classified as early-onset PE (EOPE) and called onset PE (LOPE) if symptoms show up after 34 weeks[4,5]. Current clinical management of PE includes detecting and controlling blood pressure, preventing maternal seizures, and limiting the damage to fetuses. However, to date the only effective treatment for PE is the delivery of the babies, many of them being pre-terms[6].

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Clinicians often face a dilemma when treating PE patients: earlier delivery can potentially prevent severe maternal morbidities including seizure, stroke, organ dysfunction, but this leads to premature birth and various neonatal complications[7]. Current guidelines on deciding the delivery time in PE pregnancies depend mainly on the diagnosis time and severity of PE: in general, patients beyond 37 weeks of gestation or beyond 34 weeks but diagnosed with severe PE should be delivered immediately; patients with severe PE but less than 34 weeks of gestation should only continue pregnancies if intensive care resources are available; other related medical history and complications should also be considered when deciding delivery time[8,9]. These guidelines identified some important clinical determinants of PE delivery, however, neither the individual impact nor combined impacts were quantified based on each patient's unique profile. In particular, the determination of delivery time for patients with complicated conditions still largely relies on clinicians' judgment. Additionally, risk factors of developing PE, such as the patients' demographics, social status, lifestyle, and other comorbidities may also have influences on the timing of delivery, but they are rarely discussed8. Therefore, a comprehensive quantitative model to assess the contributions of these risk factors to PE patients' prognosis for the time of delivery is very informative to help clinicians to make decisions, particularly among the challenging early onset PE cases.

Here for the first time, we report the time to delivery of predicting models since the initial diagnosis of PE, utilizing electronic health record (EHR) data. The model is built upon a previously reported state-of-the-art neural network method called Cox-nnet, which showed consistently better performances over the conventional Cox-PH models under various conditions, including EHR data[10-12]. We extracted and processed 1576 PE complicated pregnancy records from the University of Michigan Medicine Healthcare System from year 2015 to 2021. The EHR

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

features cover patients' demographics, social status, lifestyle, comorbidities, medical history, lab results and vital signs at the earliest PE diagnosis were collected. Our objectives are: (1) to predict the time to delivery from the earliest diagnosis, by constructing and validating prognosis models utilizing EHR data; (2) to assess the contributions of critical EHR features informative of delivery time among PE patients.

METHODS

Data source and description

The data were obtained through Precision Health DataDirect, a web-based interface to access de-identified EHR data of more than 4 million Michigan Medicine patients. Data usage was approved by the Institutional Review Board (IRB) of the University of Michigan Medical School (HUM#00168171). The need for consent was waived. We conducted a query on all antepartum PE patients (mild and severe) with ICD-10 code and excluded HELLP syndromes that require delivery immediately. We extracted all pregnancy records associated with these selected patients and removed the pregnancies with no PE diagnosis or with only postpartum PE diagnosis. PE pregnancies without known delivery gestational age data were also removed from the analysis. In total, we extracted a total of 1576 antepartum PE pregnancy records from 1512 patients, between year 2015 and 2021 from Michigan Medicine, the academic health care system of the University of Michigan, Ann Arbor.

EHR Feature Processing

The EHR records provide a variety of heterogeneous features from demographics, lifestyle, related medical history, pregnancy information, comorbidities, lab test results and vital signs. All

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

features except lab test results and vital signs are considered baseline features, meaning that they exist before the diagnosis of PE.

Among the baseline features, demographic data include age, race, ethnicity, smoking status, and illegal drug use status at the earliest diagnosis of PE. Cumulated medical histories related to each PE were extracted: past PE, history of gestational diabetes, history of renal disease, history of placental abruption, history of IUGR, history of preterm labor and history of C-section. All medical history features are coded as binary variables with 1 denoting the present and 0 denoting the absence of a history. Pregnancy information features include parity, gravidity, number of fetuses in current pregnancy, PE onset gestational age, and the severity of PE at the initial diagnosis (mild vs severe). Elixhauser comorbidity index was used to categorize patients' comorbidities at the earliest PE diagnosis[13]. Elixhauser comorbidity index summaries a large number of common patient comorbidities ICD diagnosis codes into 29 categories, with each category being a binary variable, indicating whether this comorbidity presents or not. It is commonly used to reduce feature dimensionality and improve interpretability.

Lab test results were collected around the earliest diagnosis of each PE pregnancy, for the purpose of delivery prediction at the time of PE diagnosis. The most frequent lab test results ordered within 5 days of the initial diagnosis of PE were used for data processing. For tests ordered multiple times in the time window, we used the mean result value in the analysis. Results with more than 20% missing values among patients were removed. As a result, 13 lab test features were used in the final dataset, including 10 blood test features, 2 liver function test features, and 1 urine test feature. Similar to the lab test results, we used vital signs collected within 5 days before the initial diagnosis of PE. We included the statistics (maximum, minimum, mean and standard deviation) of systolic and diastolic blood pressures, oxygen saturation level

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

and respiratory rate, similar to others[14]. We removed other vital signs due to a large number of missing values.

All EHR features with more than 20% of missing values were removed from the analysis. As a result, 71 features were kept (**Supplementary Table 1).** The remaining missing values were imputed using the PMM algorithm from R package "mice"[15]. All categorical features were converted into dummy variables. Scaling was done to all numeric variables, dividing their values by their root mean square[16].

Model Construction

We developed one baseline model and one full model using a subset or all EHR features. The baseline model consists of baseline features including demographic, medical history, comorbidities and PE diagnosis gestational age. The full model incorporated all features from the baseline model, as well as additional lab test results and vital signs measured at the earliest diagnosis of PE.

We constructed both baseline and full models using the Cox-nnet v2 algorithm. The Cox-nnet model is an artificial neural network (ANN) prognosis predicting model based on Cox Proportional Hazard (Cox-PH). The model consists of one input layer, one hidden layer and one Cox-PH output layer10,11. The updated version of Cox-nnet, Cox-nnet v2, significantly reduced the computational cost to make it suitable for large-scale input like EHR data. To fit each model, we randomly divided the corresponding dataset into a training set (80%) and a testing set (20%). Hyperparameters in the Cox-nnet v2 model were auto tunned using Adam Optimizer[17]. For

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

comparison, we used the Cox-PH model with ridge regulation. We compared Cox-nnet v2 and Cox-PH models using the concordance index (C-index), as done previously10,11.

To derive a subset of clinically significant and easily interpretable features, we reduced Cox-nnet features based on both their importance scores and significance levels. To do so, we first selected the top 15 most important features based on their permutation importance score generated by the Cox-nnet model. Then we created single-variable Cox-ph models using each of the 15 features and obtained their log-rank p-values. Only features with significant log-rank test results ($p<0.05$) were kept in the analysis.

Model Validation

The validation dataset containing 3407 pregnancy records diagnosed with antepartum PE was obtained from the University of Florida (**Table 1**). Data processing and feature engineering on the validation set follow strictly the same procedures as those of the training and testing set. We validated the Cox-nnet model by plotting K-M survival curves dichotomized by the median of each significant feature from the reduced baseline model. If the plots generated using validation data present similar trends and p-values from the log-rank tests are of similar magnitude, we conclude that the time-to-delivery model is adaptable.

Software

R 4.0.2 and Python 2.7 were used for all analysis[18,19]. R package “dplyr”, “mice” were used in data preparation[20,21]. Python 2.7 was used to construct the cox-nnet v2 model. R package

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

"survival", "pec", "glmnet" were used to build and evaluate the cox-ph model with ridge regulation[22-24]. "ggplot2" was used for visualization[25,26].

RESULT

Study overview

To obtain the discovery cohort to predict the prognosis, or the time to delivery, since diagnosis of PE during pregnancy, we queried all antepartum PE patients (mild and severe) with the ICD-10 code in Michigan Medicine's Electronic Health Record (EHR) system. We excluded the cases for postpartum PE where the time to delivery is negative values and those with HELLP syndromes that require delivery immediately. To avoid confounding from prior PE history, we only used the EHR records of PE patients during the pregnancies when this disease is diagnosed for the first time. As a result, 1576 antepartum PE pregnancies remain in the final discovery dataset (**Figure 1A**).

The EHR records provide a variety of heterogeneous features from demographics, social status, related medical history, pregnancy information, comorbidities, lab test results and vital signs (**Supplementary Table 1**). After preprocessing, such as removing variables with more than 20% missing values, imputation and normalization, a total of 71 variables were used in this study (**Figure 1A**). The overall patient characteristics are presented in **Table 1**. The majority of the population are Caucasians (74%), followed by African Americans (18.3%) and Asians (6.6%). The average age at PE diagnosis is 30 (+/- 5.73). Each patient on average has 0.72 (+/- 1.15) past parity and 2.35 (+/-1.82) past gravidity. The time between first PE diagnosis and delivery is heavily left-skewed, with a mean of 6.5 days and a median of 2 days (**Figure 1C**). About 13% of patients were delivered on the same day as the presence of PE.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

The baseline model of time to delivery since PE diagnosis

We first aim to construct and validate a baseline model which only basic information, including demographics, social status, related medical history, pregnancy information and comorbidities. To predict the time to delivery, or prognosis of PE, we employed the state-of-the-art survival prediction method using Cox-nnet version 2 algorithm11. The Cox-nnet model is an artificial neural network (ANN) prognosis predicting model, based on the Cox-PH model[7]. Cox-nnet consists of one hidden layer and one Cox-PH fitting output layer (**Figure 1B**). The updated version of Cox-nnet, Cox-nnet v2, significantly reduced the computational cost to make it suitable for large-scale EHR data. For modeling, we randomly divided the dataset into a training set (80%) and a testing set (20%). The resulting Cox-nnet v2 model shows high performing prediction on time-to-delivery, with median C-index scores of 0.75 and 0.72 in training and hold-out testing data, as compared to 0.74 and 0.71 in the conventional Cox-PH models (**Figure 2A**).

To identify the most important variables among the EHR data that are predictive of PE prognosis time to delivery, we next calculated the importance scores of the input features in the Cox-nnet model and ranked the top 15 features (**Figure 2B, 2C**). We also calculated the p-value of each of these 15 features in single variate Cox-PH fitting and kept the 7 features with log-rank p-values < 0.05. The 7 features in the descending order of their importance scores are diagnosis gestational age, severe PE, past PE, age, parity, gravidity, and uncomplicated diabetes, and their details are listed in **Table 2.** Diagnosis gestational age is shown the most predictive of time to delivery, with an importance score of 464. It is subsequently followed by severe PE, past PE and maternal age. We denote this Cox-nnet model of seven features as the “reduced baseline model”. It still

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

reaches a C-index of 0.71 on the testing set, higher than the C-index of 0.70 from the Cox-PH model based on the same features.

The predicted prognosis index (PI) scores, or the fitted prognosis scores from the Cox-nnet model, can be used to illustrate the clinically discriminative values of the computational model. Patients with a higher prognosis score are expected to have shorter times to delivery. In the hold-out testing data, the PI scores of the reduced baseline model range from -2.06 to 2.25, with an average of 0.93. We stratify patients by the median PI score into the high-risk and low-risk groups and plot their Kaplan Meier (KM) curves (**Figure 2D**). The two risk groups have significant differences in the survival status (log-rank p-value < 0.0001). The associated log-rank test has a p-value close to 0, proving the strong distinction power of the prognosis scores. In associating the two risk groups with other risk factors, we identified that the high-risk group has a significantly larger diagnosis gestational age, more severe PE, and fewer patients with PE history (**Figure 2E**), confirming the observations based on the feature importance scores earlier **(Figure 2B, C).** To validate the significance of the features derived from the University of Michigan cohort, we extracted and processed a PE EHR cohort from the University of Florida, following the same inclusion and exclusion criteria and data processing pipeline. or illustration. In figure 2F we showed the dichotomized KM curves of the most important and significant features: diagnosis gestational age, severe PE and past PE using the University of Florida validation set. have a very similar trend, directionality, and p-values, confirming the validity of our baseline model.

Results of the full model of time to delivery since PE diagnosis

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

We next investigated the contribution to time of delivery, from all 71 variables, including the baseline variables as well as lab testing results and vital signs that were obtained around the time of diagnosis. We call this the full model. We conducted model building, optimization and testing the same as the baseline model, using the Cox-nnet v2 algorithm. As shown in **Figure 3A**, the full model shows higher predictive accuracy on time-to-delivery compared to the baseline model, with median C-index scores of 0.77 and 0.74 in training and hold-out testing data, respectively. It is also more accurate than the conventional Cox-PH model using the same 71 variables, with C-index scores of 0.76 and 0.73 in the same training and testing datasets.

Similar to the baseline model, we identified the top 15 most important variables among the EHR data that are predictive of time to delivery (**Figure 3B)**. We also calculated the p-value of each of these 15 features in single-variate Cox-PH fitting. Except for maternal age, all other 14 features are also significant in the single variate Cox-PH models (**Figure 3C**). Diagnosis gestational age is still shown as the most predictive of time to delivery, with an importance score of 407. It is subsequently followed by the minimum systolic blood pressure, aspartate aminotransferase (AST) value and standard deviation of diastolic blood pressure **(Table 3).** The PI scores predicted by the full model range from -2.62 to 2.76 with a mean of 0.99. Compared to the baseline model, the full model prediction results are less skewed, suggesting a better separation. The KM curves applied to testing samples, stratified by median PI scores in the training dataset, are drastically different with the log-rank p-value close to 0, confirming the high distinction of the full model (**Figure 3D**).

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

To demonstrate the significance of each of these 14 features, we again stratified the patients into the high-risk vs low-risk groups based on the median value of each feature using the testing set, and then compared the KM curves of two groups respectively (Figure 3E). The KM curves in each plot are significantly different, showing strong distinction power of all features. Larger gestational age and severe PE are associated with shorter times to delivery, consistent with the observations from the baseline model. Different from the baseline model, in the reduced full model patients with higher white blood cell count, higher AST, lower platelet count and red blood cell count have significantly short times to delivery. Among vital signs collected at PE diagnosis, larger variation in systolic and diastolic blood pressure, higher mean systolic blood pressure, and lower minimum respiratory rate are all associated with shorter times to delivery. Nulliparous women also have significantly shorter times to delivery. Additionally, we found that patients with a past PE have longer time to delivery than those who were diagnosed with PE for the first time.

DISCUSSION

Implementing time to delivery predicting models may provide clinicians with valuable information in determining patient-specific optimal delivery time. In this study, we successfully constructed and evaluated a time to delivery model for PE patients at the time of diagnosis utilizing EHR data. In addition, the model reveals and quantifies important factors informative to predict the delivery time.

In total, we identified 7 features from the baseline model and 14 features from the full model that present strong distinctions among the delivery time of PE patients. In both models, diagnosis gestational age and PE severity presented some of the strongest influences on delivery time.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Patients with larger gestational age and severe PE tend to be delivered faster than patients with smaller gestational age and mild PE, which agree with the clinical guidelines. Additionally, older patient age and prior past PE have protective effects in both models, associated with longer time to delivery. Surprisingly, with the addition of more quantitative EHR variables at the time of diagnosis, such as blood testing results and blood pressure measurements, the resulting full model (C-index 0.74) is only slightly more accurate than the baseline model (C-index 0.72) based on the same testing data. A possible explanation is that the test results and vital signs are related to the severity of PE.

Nevertheless, the inclusion of lab tests and vital signs at the time of PE diagnosis provides additional predictive, and much more interpretable information for the time to delivery. Patients with lower platelet counts and higher AST values were also delivered significantly faster compared to others. Low platelet count and high AST are signs of thrombocytopenia and impaired lever function respectively, which are important severe features of PE3. Patients with those two characteristics are likely to have comorbid severe PE and require more urgent delivery. Patients with low red blood cell count and high white blood cell counts are associated with shorter times to delivery. Low red blood cell count is a symptom of anemia, a risk factor for preterm birth and comorbidity of PE[27,28]. High white blood cell count is a common indicator of maternal inflammation or infection. We also found that larger fluctuation in blood pressure contributes to the risk of faster delivery. Specifically, patients with larger standard deviation in blood pressure measures and patients with lower minimum systolic blood pressure have a shorter time to delivery. A lower respiratory rate is associated with a shorter time to delivery. This effect

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

agrees with the guidelines that for patients with severe PE and unstable conditions, delivery is recommended soon after the maternal stabilization.

Interestingly, nulliparous patients presented shorter times between diagnosis and delivery in the full model but not the baseline model. Rather, in the baseline model, no prior PE history suggests a shorter time to delivery. While the selection of these two variables is largely dependent on other associating variables in their respective models, both variables suggest that prior history of pregnancy or PE helps patients to adapt and increase the duration between PE diagnosis to delivery in a particular pregnancy. The latter observation of a longer time to delivery given the prior past PE is particularly interesting since the past PE is a risk factor for developing PE. Further mechanistic studies are warranted to understand the complex relationships among them.

A few caveats are worth pointing out. This is a retrospective study, and it was not clear if a patient was delivered at the most optimal time with the minimum damage. It will be interesting to check how well this model will perform in perspective studies from other hospitals. Also, like all other EHR studies, prediction is dependent on the quality of the EHR data, which are influenced by clinicians' previous judgment, past hospital protocols, intensive care resources availability and patients' own intentions. These factors cannot be quantified and included in the model but can potentially affect the prognosis for delivery. Despite these potential issues, to our best knowledge, this is the first report to build a quantitative model to predict the delivery time among PE patients using EHR data. The ICD diagnosis codes we constructed the features from are widely used among health systems across the world and can be converted to older coding schemes such as Read Codes or OPCS4 used in Britain. Thus, the modeling methodology

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

reported here can be applied to the abundant EHR data to address prognosis or survival issues in PE as well as many other diseases.

In conclusion, we successfully constructed and evaluated high-performing time to delivery models for PE patients utilizing EHR data. We identified that factors such as diagnosis gestational age, parity, severe PE, past PE, AST, white blood cell count, AST, platelet count, red blood cell count and blood pressure are important in predicting delivery time among PE patients. Using this model in PE care practice could provide clinicians with a strong tool to quantify the delivery risk level and decide the optimal delivery time for PE patients at the time of diagnosis.

REFERENCE

1 Ives CW, Sinkey R, Rajapreyar I, Tita ATN, Oparil S. Preeclampsia-Pathophysiology and Clinical Presentations: JACC State-of-the-Art Review. *J Am Coll Cardiol*. 2020;76(14):1690-1702. doi:10.1016/j.jacc.2020.08.014

2 US Preventive Services Task Force. Screening for PE: US Preventive Services Task Force Recommendation Statement. *JAMA*. 2017;317(16):1661–1667. doi:10.1001/jama.2017.3439

3 Roberts JM, Rich-Edwards JW, McElrath TF, Garmire L, Myatt L; Global Pregnancy Collaboration. Subtypes of Preeclampsia: Recognition and Determining Clinical Usefulness. Hypertension. 2021;77(5):1430-1441. doi:10.1161/HYPERTENSIONAHA.120.14781

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

4 Publications Committee, Society for Maternal-Fetal Medicine, Sibai BM. Evaluation and management of severe preeclampsia before 34 weeks' gestation. *Am J Obstet Gynecol*. 2011;205(3):191-198. doi:10.1016/j.ajog.2011.07.017

5 von Dadelszen P, Magee LA, Roberts JM. Subclassification of preeclampsia. *Hypertens Pregnancy*. 2003;22(2):143-148. doi:10.1081/PRG-120021060

6 Amaral LM, Wallace K, Owens M, LaMarca B. Pathophysiology and Current Clinical Management of PE. *Curr Hypertens Rep*. 2017;19(8):61. doi:10.1007/s11906-017-0757-7

7 Hollegaard B, Lykke JA, Boomsma JJ. Time from pre-eclampsia diagnosis to delivery affects future health prospects of children. Evol Med Public Health. 2017;2017(1):53-66. Published 2017 Feb 8. doi:10.1093/emph/eox004

8 Hypertension in Pregnancy, Obstetrics & Gynecology: November 2013 - Volume 122 - Issue 5 - p 1122-1131 doi: 10.1097/01.AOG.0000437382.03963.88

9 ACOG Practice Bulletin No. 202: Gestational Hypertension and Preeclampsia. *Obstet Gynecol*. 2019;133(1):1. doi:10.1097/AOG.0000000000003018

10 Wang D, Jing Z, He K, Garmire LX. Cox-nnet v2.0: improved neural-network based survival prediction extended to large-scale EMR data [published online ahead of print, 2021 Jan 30]. *Bioinformatics*. 2021;37(17):2772-2774. doi:10.1093/bioinformatics/btab046

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

11 Ching T, Zhu X, Garmire LX. Cox-nnet: An artificial neural network method for prognosis prediction of high-throughput omics data. *PLoS Comput Biol*. 2018;14(4):e1006076. Published 2018 Apr 10. doi:10.1371/journal.pcbi.1006076

12 Zhan Z, Jing Z, He B, et al. Two-stage Cox-nnet: biologically interpretable neural-network model for prognosis prediction and its application in liver cancer survival using histopathology and transcriptomic data. NAR Genom Bioinform. 2021;3(1):lqab015. Published 2021 Mar 22. doi:10.1093/nargab/lqab015

13 Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care* 1998;36:8-27

14 Zhao J, Feng Q, Wu P, et al. Learning from Longitudinal Data in Electronic Health Record and Genetic Data to Improve Cardiovascular Event Prediction. *Sci Rep*. 2019;9(1):717. Published 2019 Jan 24. doi:10.1038/s41598-018-36745-x

15 Becker, R. A., Chambers, J. M. and Wilks, A. R. (1988) The New S Language. Wadsworth & Brooks/Cole.

16 van Buuren S, Groothuis-Oudshoorn K (2011). “mice: Multivariate Imputation by Chained Equations in R.” *Journal of Statistical Software*, **45**(3), 1-67. doi: 10.18637/jss.v045.i03.

17 Kingma D.P., Ba J. (2015) Adam: a method for stochastic optimization. In: Proceedings of the 3rd International Conference on Learning Representations (ICLR 2015)

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

18 R Core Team (2021). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

19 Van Rossum, G., & Drake Jr, F. L. (1995). *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam.

20 Hadley Wickham, Romain François, Lionel Henry and Kirill Müller (2020). dplyr: A Grammar of Data Manipulation. R package version 1.0.2. https://CRAN.R-project.org/package=dplyr

21 Therneau T (2020). _A Package for Survival Analysis in R_. R package version 3.2-7, <URL: https://CRAN.R-project.org/package=survival>.

22 Terry M. Therneau, Patricia M. Grambsch (2000). _Modeling Survival Data: Extending the Cox Model_. Springer, New York. ISBN 0-387-98784-3.

23 Ulla B. Mogensen, Hemant Ishwaran, Thomas A. Gerds (2012). Evaluating Random Forests for Survival Analysis Using Prediction Error Curves. Journal of Statistical Software, 50(11), 1-23. URL https://www.jstatsoft.org/v50/i11.

24 Jerome Friedman, Trevor Hastie, Robert Tibshirani (2010). Regularization Paths for Generalized Linear Models via Coordinate Descent. Journal of Statistical Software, 33(1), 1-22. URL http://www.jstatsoft.org/v33/i01/.

25 H. Wickham. ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York, 2016.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

26 Gregory R. Warnes, Ben Bolker, Lodewijk Bonebakker, Robert Gentleman, Wolfgang Huber, Andy Liaw, Thomas Lumley, Martin Maechler, Arni Magnusson, Steffen Moeller, Marc Schwartz and Bill Venables (2020). gplots: Various R Programming Tools for Plotting Data. R package version 3.1.1. https://CRAN.R-project.org/package=gplots

25 Haider BA, Olofin I, Wang M, Spiegelman D, Ezzati M, Fawzi WW; Nutrition Impact Model Study Group (anaemia)Anaemia, prenatal iron use, and risk of adverse pregnancy outcomes: systematic review and meta-analysis.BMJ2013346f3443

27 Smith C, Teng F, Branch E, Chu S, Joseph KS. Maternal and Perinatal Morbidity and Mortality Associated With Anemia in Pregnancy. *Obstet Gynecol*. 2019;134(6):1234-1244. doi:10.1097/AOG.0000000000003557

Tables

Table 1: The Characteristics of most important features in the PE discovery cohort from University of Michigan and validation cohort from University of Florida

Variable Name	***University of Michigan Cohort (n = 1576) Mean (sd)***	***University of Florida Cohort (n = 3407) Mean (sd)***
Maternal Age	*30.09(5.73)*	*28.05(6.49)*
Gravidity Count	*2.35(1.82)*	*2.25(2.06)*
Parity Count	*0.72(1.15)*	*0.62(1.52)*
Number of fetuses	*1.07(0.27)*	*NA*
Diagnosis Gestational Age(day)	*249.74(27.13)*	*260.20(6.49)*
Time to delivery (day)	*6.50(13.57)*	*7.12(16.81)*
Race (%)		
African American	*18.30%*	*36.31%*
American Indian or Alaska Native	*0.40%*	*0.12%*
Asian	*6.60%*	*1.41%*
Caucasian	*74.30%*	*49.72%*
Native Hawaiian and Other Pacific Islander	*0.40%*	*0.06%*
Unknown or Other	*0.00%*	*12.38%*

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

Ethnicity (%)		
Hispanic	*5.50%*	*10.60%*
Non-Hispanic	*94.50%*	*88.11%*
Unknown	*0.00%*	*1.29%*
Illegal Drug Use Status (%)		
Yes	*9.52%*	*0.09%*
No	*90.48%*	*99.91%*
Past PE (%)		
Yes	*7.99%*	*9.33%*
No	*92.01%*	*90.67%*

Table 2: Feature details of reduced baseline model

Feature Name	***Permutation Importance Score***	***Single variate Cox-PH P-value***
Diagnosis GA	*464.226835*	*6.52E-112*
Severe PE	*115.886885*	*1.75E-12*
Past PE	*9.907836*	*6.77E-08*
Age	*7.231669*	*4.69E-02*
Parity	*4.091492*	*9.09E-07*
Gravidity	*3.045364*	*1.09E-06*
Uncomplicated Diabetes	*2.654092*	*6.72E-03*

Table 3: Feature details of reduced full model

Variable Name	***Permutation Importance Score***	***Single variate Cox-PH P-value***
Diagnosis GA	*407.389062*	*2.66E-95*
Min Systolic BP	*33.087822*	*8.68E-35*
AST	*32.956872*	*1.27E-04*
Diastolic BP Standard Deviation	*21.886153*	*6.36E-39*
Max Diastolic BP	*21.173801*	*4.19E-32*
Severe PE	*19.780133*	*2.97E-11*
Mean Systolic BP	*12.821062*	*2.23E-03*
Platelet Count	*10.860128*	*6.39E-09*
Systolic BP Standard Deviation	*9.986978*	*1.93E-31*
Min Respiratory Rate	*9.74149*	*2.82E-13*
Parity	*9.525487*	*4.40E-05*
White Blood Cell Count	*6.461414*	*5.93E-09*
Past PE	*6.266599*	*4.48E-06*
Red Blood Cell Count	*5.861864*	*2.06E-02*

Figure and Figure Legends

Figure 1: Overview of study design, model construction and validation. A: Flowchart of model development B: Cox-nnet neural network architecture. C. The KM curves of time to delivery since PE diagnosis, in the University of Michigan discovery cohort.

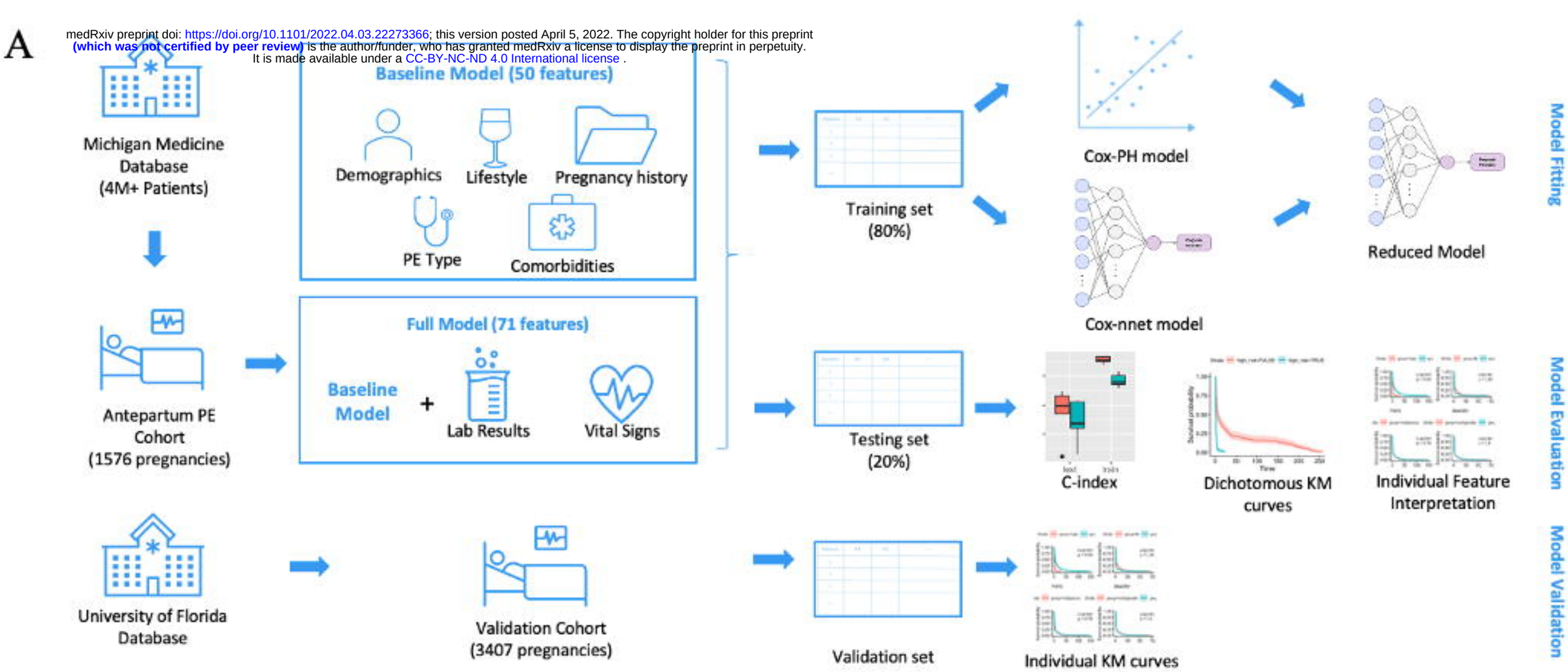
Figure 2: Baseline model results, interpretation, and evaluation. A: The boxplots of C-index values from Cox-PH (green) and Cox-nnet (red) models on the 80% training and 20% testing set, using University of

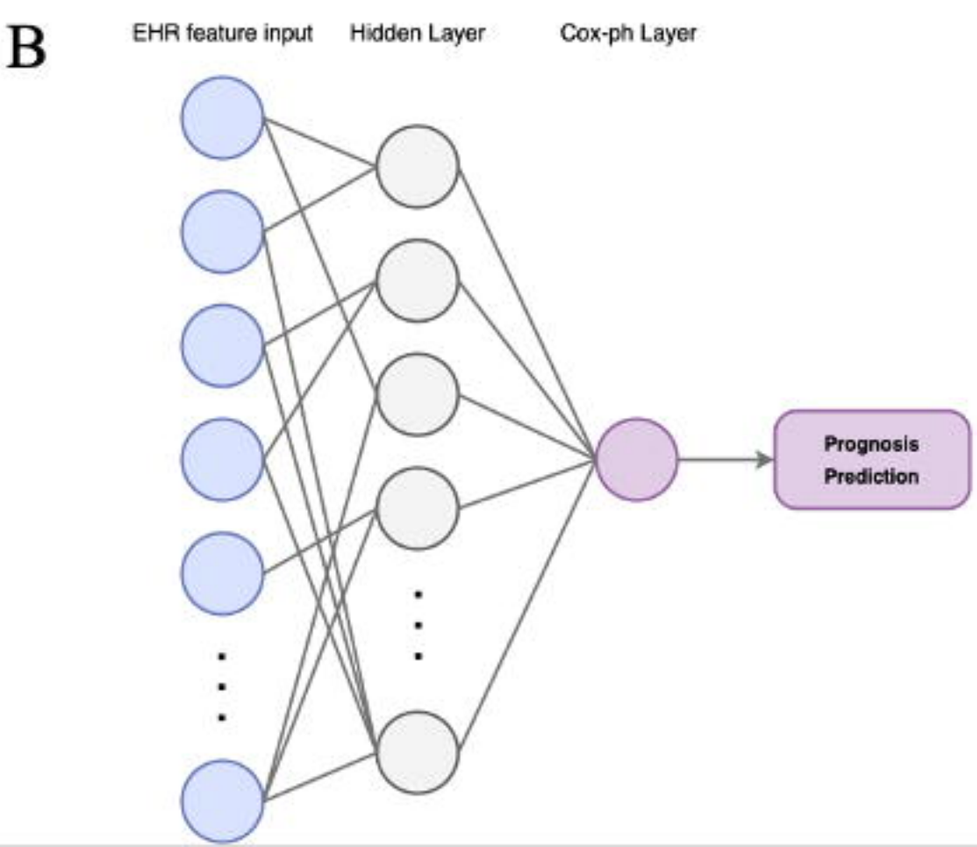
medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .

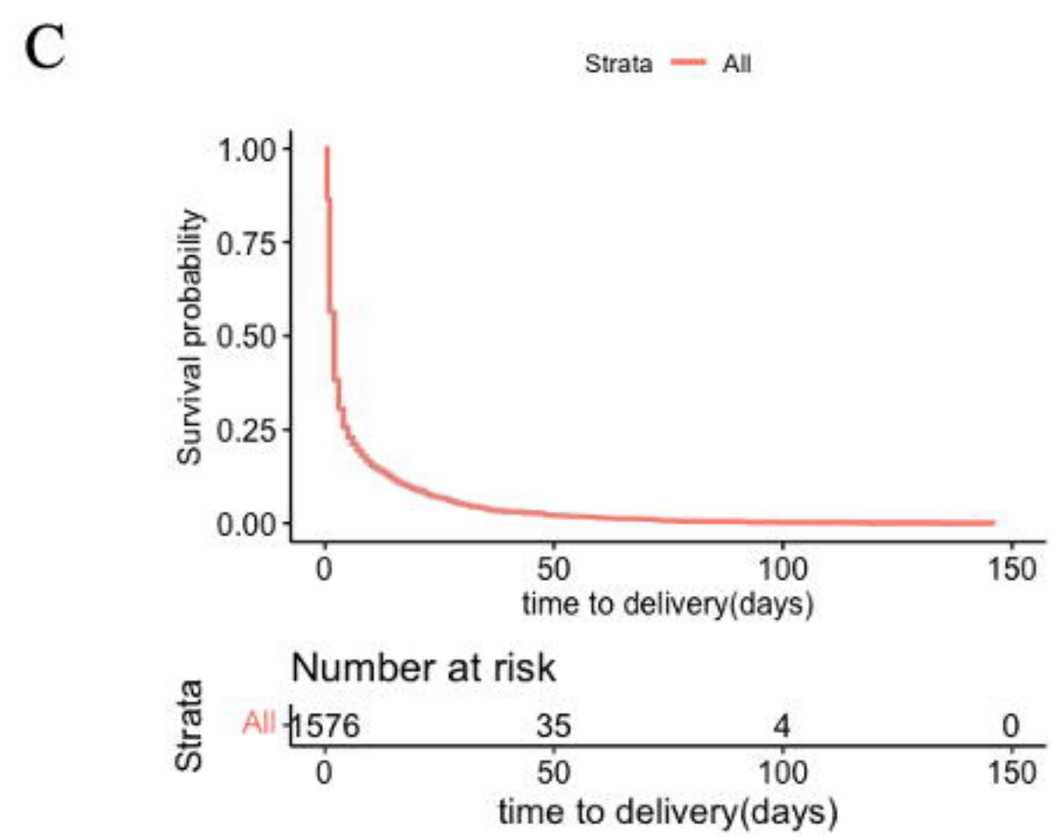
Michigan cohort. B: The ln-transformed permutation importance score of the top 15 most important features. Positive and negative signs indicate that higher value in the feature is associated with increased and reduced risk (hazard), respectively. C. The heatmap of the top 15 most important features sorted by the Cox-nnet permutation importance score and their ranks by p-value from individual Cox-PH fitting. Features with star in their name were selected in the reduced model. D: The KM curves of high-risk and low-risk groups on the testing set, dichotomized by the median value of predicted prognosis score. E: The KM curves on the testing set, based on each of the seven features from the reduced model dichotomized by their median values. F: The KM plots dichotomized by median value of diagnosis GA, severe PE and past PE, generated using validation set.

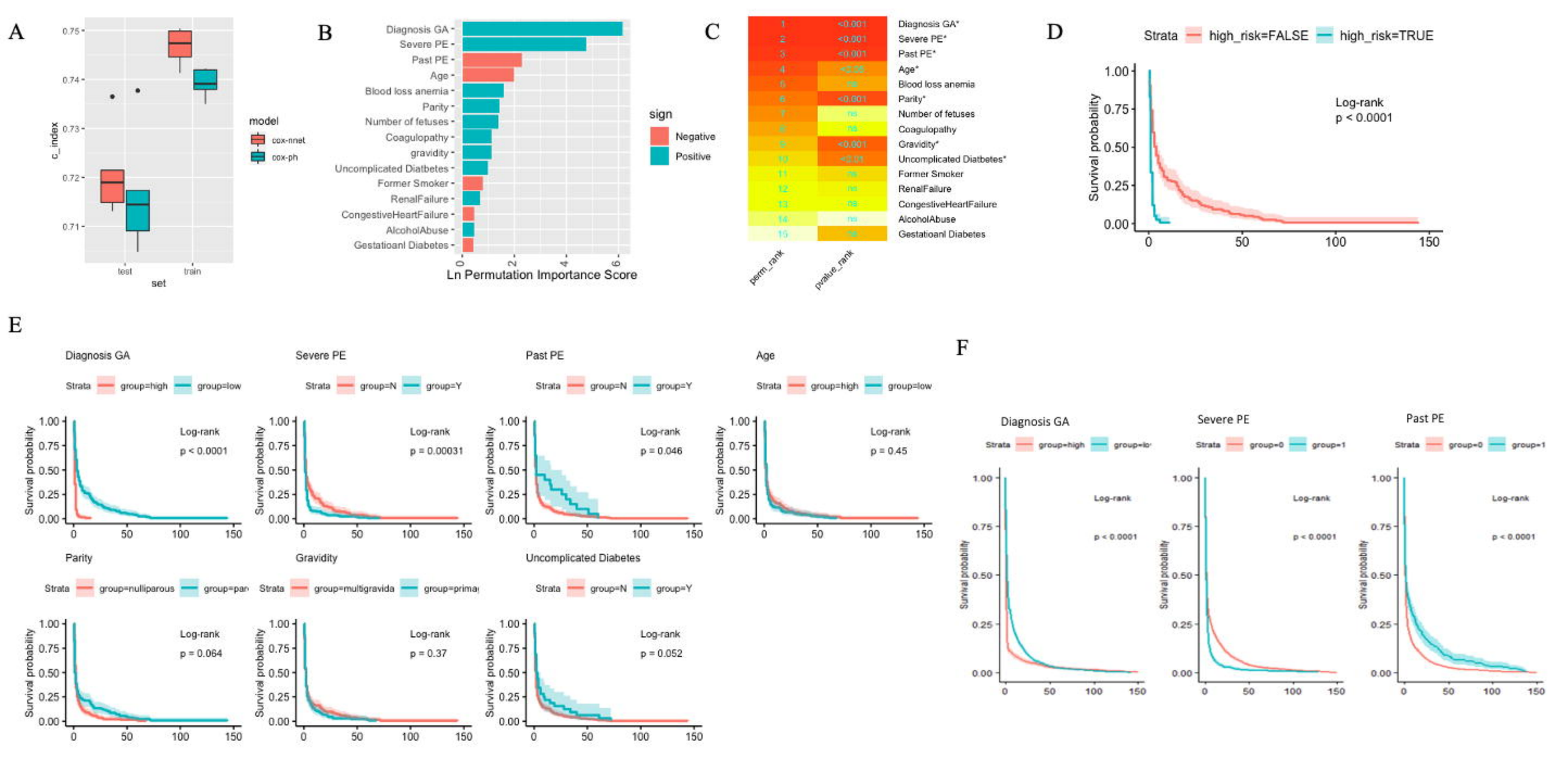
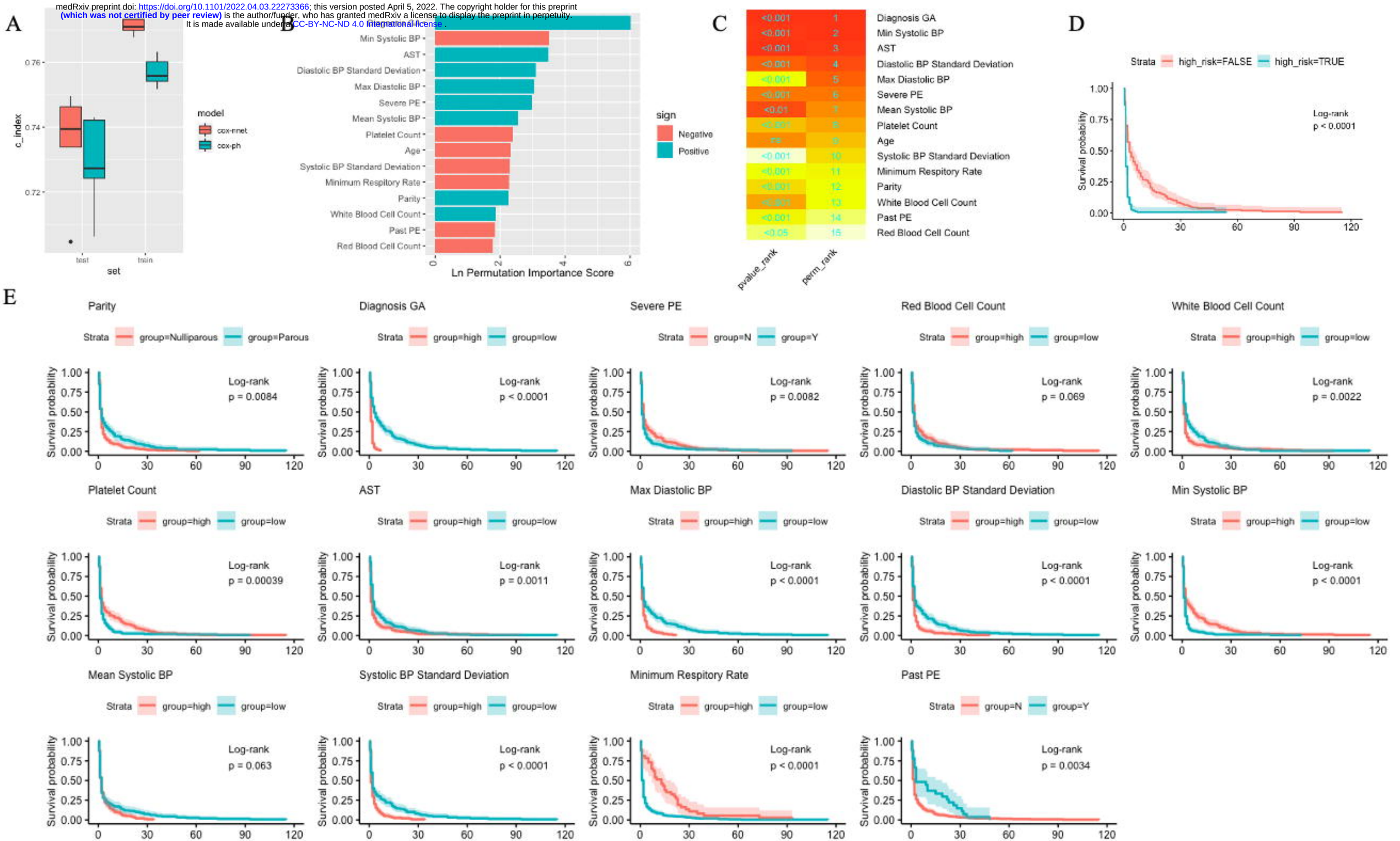
Figure 3: Full model results, interpretation and evaluation. A: The boxplots of C-index values from Cox-PH (green) and Cox-nnet (red) models on the 80% training and 20% testing set, using University of Michigan cohort. B: The ln-transformed permutation importance score of the top 15 most important features. Positive and negative signs indicate that higher value in the feature is associated with increased and reduced risk (hazard), respectively. C. The heatmap of the top 15 most important features sorted by the Cox-nnet permutation importance score and their ranks by p-value from individual Cox-PH fitting. Features with star in their name were selected in the reduced model. D: The KM curves of high-risk and low-risk groups on the testing set, dichotomized by the median value of prognosis score predicted by reduced full model. E: The KM curves on the testing data, based on each of the 14 features from the reduced model dichotomized by their median values.

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .








A
c_index
set
test
train
model
cox-nnet
cox-ph
B
Diagnosis GA
Severe PE
Past PE
Age
Blood loss anemia
Parity
Number of fetuses
Coagulopathy
gravidity
Uncomplicated Diatbetes
Former Smoker
RenalFailure
CongestiveHeartFailure
AlcoholAbuse
Gestatioanl Diabetes
Ln Permutation Importance Score
sign
Negative
Positive
C
perm_rank
pvalue_rank
D
Strata
high_risk=FALSE
high_risk=TRUE
Survival probability
Log-rank
p < 0.0001
E
Diagnosis GA
Severe PE
Past PE
Age
Parity
Gravidity
Uncomplicated Diabetes
Log-rank
p = 0.00031
p = 0.046
p = 0.45
p = 0.064
p = 0.37
p = 0.052
F
Diagnosis GA
Severe PE
Past PE

medRxiv preprint doi: https://doi.org/10.1101/2022.04.03.22273366; this version posted April 5, 2022. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted medRxiv a license to display the preprint in perpetuity. It is made available under a CC-BY-NC-ND 4.0 International license .