Oxygenation Fluctuations Associated with Severe Retinopathy of Prematurity

Insights from a Multimodal Deep Learning Approach

Wei-Chun Lin, MD, PhD,[1] Brian K. Jordan, MD, PhD,[2] Brian Scottoline, MD, PhD,[2] Susan R. Ostmo, MS,[1] Aaron S. Coyner, PhD,[1] Praveer Singh, PhD,[3] Jayashree Kalpathy-Cramer, PhD,[3] Deniz Erdogmus, PhD,[4] R.V. Paul Chan, MD, MSc,[5] Michael F. Chiang, MD,[6,7] J. Peter Campbell, MD, MPH[1]

Purpose: Retinopathy of prematurity (ROP) is one of the leading causes of blindness in children. Although the role of oxygen in the pathophysiology of ROP is well established, a precise understanding of the dynamic relationship between oxygen exposure ROP incidence and severity is lacking. The purpose of this study was to evaluate the correlation between time-dependent oxygen variables and the onset of ROP.

Design: Retrospective cohort study.

Participants: Two hundred thirty infants who were born at a single academic center and met the inclusion criteria were included. Infants are mainly born between January 2011 and October 2022.

Methods: Patient data were extracted from electronic health records (EHRs), with sufficient time-dependent oxygen data. Clinical outcomes for ROP were recorded as none/mild or moderate/severe (defined as type II or worse). Mixed-effects linear models were used to compare the 2 groups in terms of dynamic oxygen variables, such as daily average and the coefficient of variation (COV) fraction of inspired oxygen (FiO_2). Support vector machine (SVM) and long-short-term memory (LSTM)-based multimodal models were trained with fivefold cross-validation to predict which infants would develop moderate/severe ROP. Gestational age (GA), birth weight, and time-dependent oxygen variables were used to develop predictive models.

Main Outcome Measures: Model cross-validation performance was evaluated by computing the mean area under the receiver operating characteristic (AUROC) curve, precision, recall, and F1 score.

Results: We found that both daily average and COV of FiO_2 were associated with more severe ROP (adjusted $P < 0.001$). With fivefold cross-validation, the multimodal LSTM models had higher performance than the best static models (SVM using GA and 3 average FiO_2 features) and SVM models trained on GA alone (mean AUROC = 0.89 ± 0.04 vs. 0.86 ± 0.05 vs. 0.83 ± 0.04).

Conclusions: The development of severe ROP might not only be influenced by oxygen exposure but also by its fluctuation, which provides direction for future study of pathophysiological factors associated with severe ROP development. Additionally, we demonstrated that multimodal neural networks can be a method to extract useful information from time-series data, which may be a valuable methodology for the investigation of other diseases using EHR data.

Financial Disclosure(s): Proprietary or commercial disclosure may be found in the Footnotes and Disclosures at the end of this article. *Ophthalmology Science* 2024;4:100417 © 2023 Published by Elsevier Inc. on behalf of the American Academy of Ophthalmology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Retinopathy of prematurity (ROP) is an eye disease that affects the retina of preterm babies. It is a leading cause of childhood blindness worldwide.[1,2] Blindness can occur from abnormal growth and development of blood vessels that lead to fibrovascular retinal detachment. Although there have been improvements in neonatal care, ROP is still a challenge for pediatric health care and needs early detection and intervention to prevent vision loss.[3] A comprehensive understanding of the risk factors and underlying mechanisms of ROP is crucial for developing effective screening strategies and treatment plans.

Our understanding of the relationship between oxygen exposure and the incidence and severity of ROP has evolved over time. Several risk factors related to oxygen have been identified, including the utilization of any supplemental oxygen, the highest oxygen concentration, the duration of oxygen administration, and prolonged mechanical ventilation.[1,4–10] The optimal range of oxygen saturation (SpO_2) remains controversial, despite numerous large randomized-controlled studies that have evaluated different target ranges.[11–15] Some studies found that the incidence and severity of ROP can be decreased by adjusting the delivered oxygen level through monitoring of SpO_2 targets and avoiding hyperoxemia,[16–19] although this is balanced against the risk of intermittent hypoxemia. Intermittent hypoxemia refers to repeated episodes of decreased levels of

oxygen in the blood, which may increase both overall morbidity and ROP risk.[19−21]

Our lack of precise understanding regarding the relationship between oxygen exposure and ROP risk may be due to the complexity of the relationship between oxygen delivery and blood-oxygen content. Because of the presence of lung disease and impaired oxygen transport, the relationship between inhaled oxygen delivered, measured by the fraction of inspired oxygen (FiO_2), and oxygen content in the blood, as most easily measured by blood SpO_2, is nonlinear. Moreover, SpO_2 is an imperfect measurement of the total oxygen concentration of the blood for an infant receiving supplemental oxygen because, when hemoglobin is fully saturated, the true partial pressure of oxygen in the arterial circulation may be much higher than in an infant with fully saturated hemoglobin without supplemental oxygen. The true partial pressure of oxygen in the blood is usually known only through blood gas measurement in an infant. Further complicating our understanding is that SpO_2 can vary from moment to moment, and FiO_2 may be adjusted accordingly, which in and of itself may be an ROP risk factor. For example, several studies have suggested that not only high oxygen exposure but also variability in oxygen exposure is a significant predictor of severe ROP.[22,23] There is support from preclinical data as well in the oxygen-induced retinopathy rat model, where McColm et al[24] found that a more severe phenotype could be induced by fluctuating exposure to high oxygen concentrations.

Advances in big data analytics and deep learning offer new methods to understand complex relationships in biological data. In particular, the integration of deep learning with electronic health record (EHR) data can not only offer insights into fundamental pathophysiological relationships between clinical variables but can also be utilized in clinical prediction models.[25−29] Moreover, deep learning models are flexible and capable of integrating different modalities of data, such as tabular data, free-text data, images, and time-series data, into a single predictive model.[30,31] The ability to process and utilize multimodal inputs enhances the model's performance and the accuracy of predictions.[32] The purpose of this study is to evaluate our gap in knowledge regarding time-dependent variables in the EHR and ROP and to compare prediction models using these variables compared to standard static variables (birth weight and gestational age [GA]).

Methods

Setting

The present study was conducted as part of a multicenter cohort study on ROP and was performed at the Oregon Health & Science University (OHSU), a large academic medical center in Portland, Oregon. In 2006, OHSU implemented an institution-wide EHR to handle all practice management, clinical documentation, order entry, medication prescribing, and billing. This study adhered to the tenets of the Declaration of Helsinki and was approved by the Institutional Review Board at OHSU (institutional review board #21403).

Dataset

The primary dataset consisted of infants whose parents or guardians provided informed consent to participate in the Imaging and Informatics in ROP study between January 2011 and October 2022. A small number of infants born prior to 2011 were also included if they met the other inclusion criteria for the study. Inclusion criteria for the study were as follows: (1) $\leq$ 32 0/7 weeks of GA and (2) complete oxygen administration (FiO_2) and saturation (SpO_2) data for a minimum of 25 of the first 30 days in the EHR. Patients were excluded if they had missing EHR data, were born after 32 0/7 weeks of GA, or were transferred to the OHSU after birth. The FiO_2 and SpO_2 data were extracted from the EHR data warehouse at OHSU. The majority of records (89%) had data intervals of 1 to 3 hours, while a few infants had more frequent recordings on specific days. Aggregated SpO_2 variables, such as daily mean, maximum, minimum, and coefficient of variation (COV), were calculated from these raw data to mitigate the impact of variability in the frequency of oxygen records. Daily aggregation smoothes out short-term fluctuations and reduces noise, facilitating the identification of meaningful trends. The diagnostic procedure for each retinal examination comprised a clinical bedside evaluation of the retinae, coupled with an image-based ROP diagnosis performed by 3 experienced ROP graders (J.P.C., M.F.C., and S.O.). The image-based diagnosis relied on all 5 retinal fundus image views and was based on the International Classification of Retinopathy of Prematurity criteria. The first group consisted of infants who either did not exhibit ROP or had mild ROP, which did not require treatment. The second group was comprised of infants who exhibited moderate to severe ROP (type 2 or more severe) and may or may not have received treatment.

Statistical Analysis

Continuous variables were expressed as mean and standard deviation, and a Welch's t-test was used to compare the difference between ROP severity groups. The distribution of categorical variables was described using frequency and percentage values, and group comparisons were performed using the chi-square tests. Mixed-effects linear models were used to compare the 2 groups in terms of the daily mean of SpO_2, daily COV of SpO_2, daily mean of inspired oxygen (FiO_2), daily COV of FiO_2, and the daily proportion of high oxygen exposure while accounting for the correlation among repeated measurements obtained from the same infant over time. A linear mixed model was initially fit for each variable with 5 predictors: GA, birth weight, day of life, ROP severity group, and an interaction term between the day of life and the ROP severity group. The significance of the interaction term was assessed, with a statistically significant result indicating that the relative effect of the ROP severity group changed over time, thereby making the generalization of group differences across time infeasible. For instances in which the interaction term was not significant, it was removed, and the model was refitted with GA, birth weight, the day of life, and the ROP severity group as the main predictors, thereby allowing for a longitudinal comparison between groups. Data analysis was conducted using the nlme package in R version 4.2.3 (R Foundation for Statistical Computing), with a 5% significance level applied for hypothesis testing.

Machine Learning and Deep Learning Models

Due to the nature of collinearity among the oxygen exposure and saturation data, we compared 2 popular machine learning models, random forest and support vector machine (SVM), to explore the possible oxygen parameters related to severe ROP onset. We
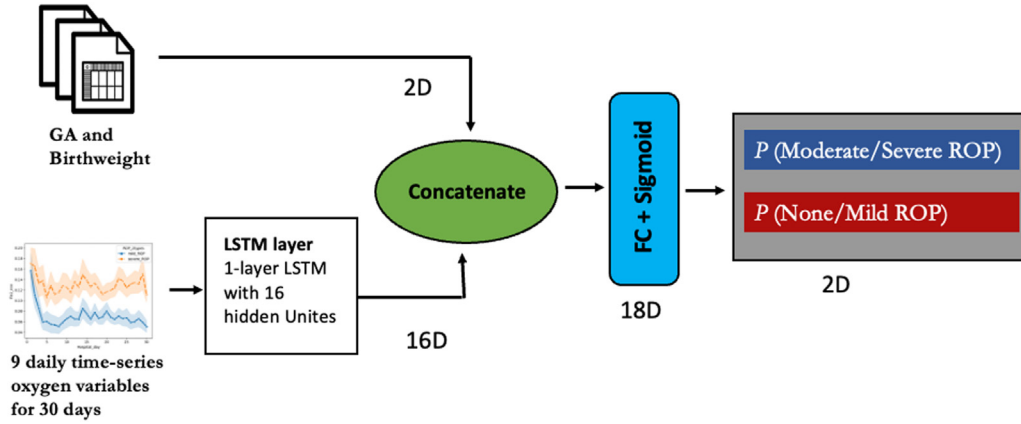
Figure 1. Architecture of multimodal long-short-term memory (LSTM) prediction model. Architecture for the multimodal-LSTM prediction model with 16 hidden units using oxygen time-series and static data. FC = fully connected layer; GA = gestational age; ROP = retinopathy of prematurity.

applied recursive feature selection to select the most relevant variables, including GA, average FiO_2, the COV of FiO_2, and the proportion of hyperoxemia. Next, we used GA to develop the first SVM model and used the 4 selected variables to develop the second SVM model. The fivefold cross-validation technique was employed to train models, and the outputs of each submodel were used to calculate average results. Cross-validation is a commonly used method to evaluate the performance of a predictive model on an independent dataset, which is a resampling procedure that iteratively trains and tests the model using different proportions of the dataset. The goal of cross-validation is to provide an estimation of the model's generalizability. A grid search was performed to optimize the hyperparameters of the models. These models were implemented using scikit-learn in the Python environment.

Lastly, we developed a multimodal deep learning model to extract information from time-series oxygen data and incorporate it with 2 static variables to improve the predictive model performance. The model was implemented using PyTorch in the Python environment. Figure 1 shows the model architectures of the multimodal neural networks. The long-short-term memory (LSTM) layers were used to capture the information from the 9 time-series oxygen data, including the daily proportion of normal FiO_2 (≤ 0.23), daily mean, maximum, minimum, and COV of FiO_2 and SpO_2. The static data included GA and birth weight. The neural network model architectures (hidden layers and directions of LSTM) and hyperparameters (optimizers, learning rates, batch size, and dropout rates) were optimized. The optimized neural network model was trained with Adam optimizer and used ReLU as the activation function. The batch size was 4 with a learning rate of 0.001. The parameters of the last hidden layer (dimensions = 16) were contacted with static data and connected with one more dense layer and the final out layer with a sigmoid function. Model performance was evaluated by computing the mean area under the receiver operating characteristic (AUROC) curve, precision, recall, and F1 score.

Results

Dataset Characteristics

This study included 230 infants who were born at OHSU and met the inclusion criteria, of whom 37 did not

Table 1. Baseline Demographic Characteristics

Characteristics	None/Mild ROP (n = 143) Mean (SD)	Moderate/Severe ROP (n = 87) Mean (SD)	P Value
Birth weight, gram	975.3 ± 241.8	716.2 ± 193.3	< 0.001*
Gestational age, week	27.3 ± 1.5	25.4 ± 1.3	< 0.001*
	N (%)	N (%)	
Sex			
Male	79 (55)	49 (56)	0.982
Female	64 (45)	38 (44)	0.982
Race			
White	110 (77)	67 (77)	1.0
Black	10 (7)	2 (2)	0.139
Asians	6 (4)	2 (2)	0.713
Others and unknown	17 (12)	16 (18)	0.18
Hispanic ethnicity	31 (22)	24 (28)	0.341

ROP = retinopathy of prematurity; SD = standard deviation.
Calculation based on Pearson chi-square test (sex, race, and Hispanic ethnicity) and Welch's t-test (birth weight and gestational age).
*P values were 2-sided, and $P \leq 0.05$ was considered statistically significant.
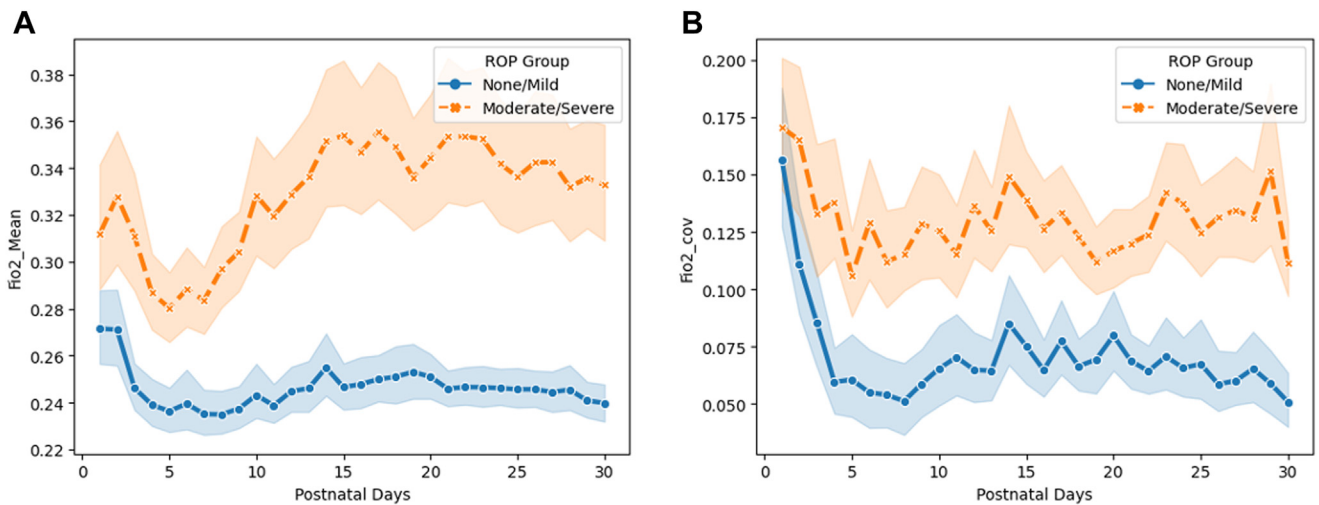
Figure 2. Descriptive characteristics of fraction of inspired oxygen (FiO_2) during the first 1 postnatal month. **A,** Daily average FiO_2 for 1 month. **B,** Daily average FiO_2 coefficient of variation (COV) for 1 month. ROP = retinopathy of prematurity.

develop ROP, 106 developed mild ROP, 39 developed type 2 ROP, and 48 developed treatment-requiring ROP (Table 1). The none/mild ROP subjects were grouped as mild or no ROP (n = 143), and moderate/severe ROP subjects were grouped as moderate or severe ROP (n = 87). Moderate to severe ROP was associated with lower birth weight ($P \leq 0.001$) and lower GA at birth ($P \leq 0.001$). No significant differences were observed for race, gender, and ethnicity. Birth weight and GA were taken into account in the adjusted statistical analyses.

Trends in FiO_2 and SpO_2

The trend of the average daily FiO_2 for the first month is shown in Figure 2A, while Figure 2B displays the daily COV of FiO_2 for the first month. The daily average FiO_2 percentages were significantly higher for patients with moderate to severe ROP compared to those with no or mild ROP, as shown in Table 2 (adjusted $P < 0.001$). The average FiO_2 percentage decreased for both groups during the first 5 days, then reached its peak at approximately 12

to 14 days, and slightly declined over the next 2 weeks for the group with severe ROP. The group with mild or no ROP displayed a relatively stable trend with FiO_2 values remaining below 25%. Similarly, the daily COV of FiO_2 was significantly higher for patients with severe ROP compared to those in the no or mild ROP group (adjusted $P < 0.001$). The daily COV of FiO_2 declined during the first week and fluctuated during the subsequent days.

Figure 3A shows the trends of average daily SpO_2, and Figure 3B displays the daily COV of SpO_2 for the first month. The daily average SpO_2 percentage was slightly higher but without statistical significance for patients with moderate or severe ROP (adjusted $P = 0.151$), as shown in Table 2. The average SpO_2 increased in the first 5 days and then declined until day 20; afterward, it remained steady or slightly increased. In contrast, there was no significant difference between the daily COV of SpO_2 between the 2 groups (adjusted $P = 0.58$). The daily COV of SpO_2 decreased in the first 5 days, then rose until day 20. After that, the daily COV of SpO_2 remained stable. Furthermore, in Table 2, we found that patients

Table 2. Longitudinal Variables between Study Groups Using Linear Mixed-Effect Models

Variables	Adjusted Slope Coefficient (Standard error)	Adjusted P Value
Daily average FiO_2	0.0385 (0.0070)	< 0.001*
Daily FiO_2 coefficient of variation	0.0382 (0.0056)	< 0.001*
Proportion of daily $FiO_2 \leq 0.23$	−0.0667 (0.0338)	0.043*
Daily average SpO_2	−0.5704 (0.959)	0.1509
Daily SpO_2 coefficient of variation	0.0032 (0.0058)	0.5801

FiO_2 = fraction of inspired oxygen; SpO_2 = oxygen saturation.
The adjusted slope coefficient demonstrates the average change in values between the 2 groups over time after controlling for birth weight and gestational age. The positive slope suggests that the none/mild retinopathy of prematurity group average value is less, and a negative slope indicates a greater average value for the none/mild retinopathy of prematurity group.
*P values were two-sided, and $P \leq .05$ was considered statistically significant.
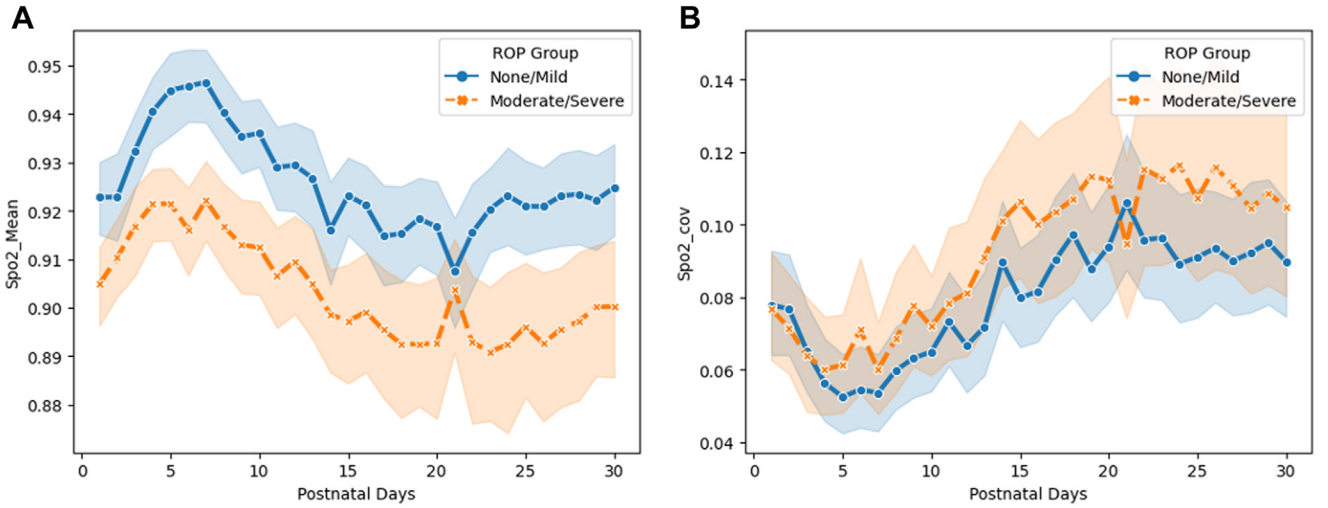
Figure 3. Descriptive characteristics of oxygen saturation (SpO$_2$) during the first 1 postnatal month. **A,** Daily average SpO$_2$ for 1 month. **B,** Daily average SpO$_2$ coefficient of variation (COV) for 1 month. ROP = retinopathy of prematurity.

with severe ROP had a lower proportion of time with ambient air FiO$_2$ compared to the subjects in the no or mild ROP group (adjusted $P = 0.043$).

Predictive Model Performance

The performance of the SVM models and multimodal neural networks for predicting severe ROP is shown in Table 3. The results of the fivefold cross-validation showed that the performance of the multimodal neural networks had the highest AUROC and F1 score, followed by the SVM model using GA and the average of 3 FiO$_2$ features, and SVM models trained solely on GA. The mean AUROC of the multimodal neural networks was 0.89 ± 0.04 compared to 0.86 ± 0.05 for the SVM with 4 variables and 0.83 ± 0.04 for the SVM with GA alone
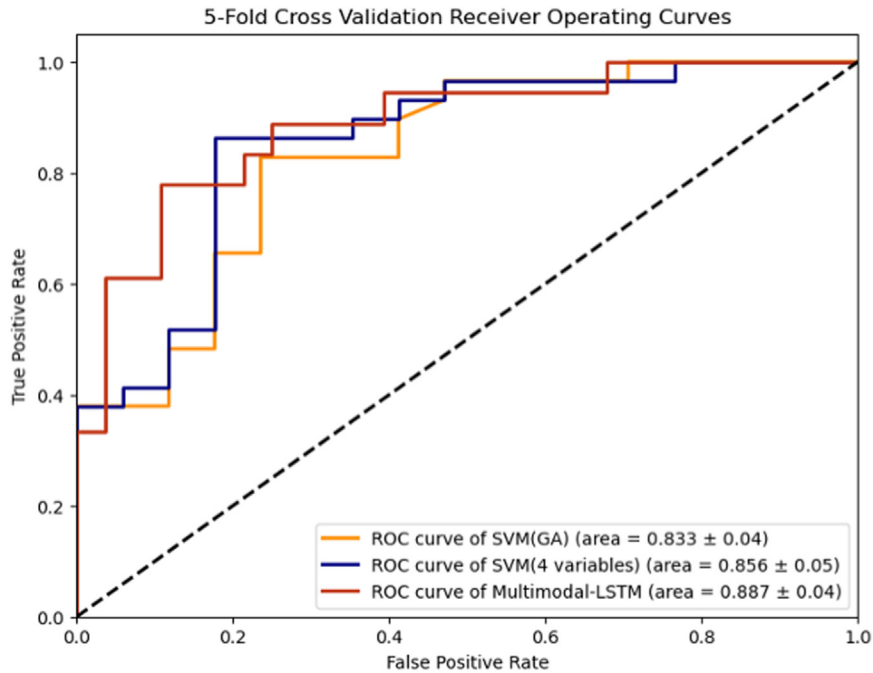

Figure 4. Receiver operating characteristic (ROC) curves for predictive models. ROC curves are shown for the 3 predictive models using different input data. 4 variables = 4 variables, including gestational age, daily average fraction of inspired oxygen (FiO$_2$), daily FiO$_2$ coefficient of variation, and proportion of daily FiO$_2 \leq 0.23$; GA = gestational age; LSTM = long-short-term memory; SVM = support vector machine.

Table 3. Performance Metrics for Prediction Models

		Surgical Outcomes Predictions			
	Model Inputs	AUROC	Precision	Recall	F1 Score
SVM-GA	GA	0.83	0.82	0.79	0.80
SVM-GA-FiO$_2$	GA + FiO$_2$*	0.86	0.82	0.82	0.82
MNN-LSTM	**T + S**	**0.89**	**0.87**	**0.84**	**0.85**

AUROC = area under the receiver operating characteristic curve; FiO$_2$ = fraction of inspired oxygen; FiO$_2$*, 3 variables including daily average FiO$_2$, daily FiO$_2$ coefficient of variation, and proportion of daily FiO$_2 \leq 0.23$; GA = gestational age; LSTM = long-short-term memory; MNN = multimodal neural network; S = static data (gestational age and birth weight); SVM = support vector machine; T = time-series oxygen variables for the first postnatal month. The bold in the table is maximum values of that evaluation metrics.

(Figure 4). Additionally, the multimodal neural networks demonstrated the highest F1 score at 0.85 ± 0.06, followed by the SVM with 4 variables at 0.82 ± 0.07 and the SVM with GA alone at 0.80 ± 0.06.

Discussion

In this study, we developed statistical, machine learning, and multimodal deep learning models to investigate the correlations between several quantitative oxygen parameters recorded in the EHR and the onset of severe ROP. The 2 key findings of our study were (1) that a significant positive correlation exists between both daily average FiO$_2$ and variability of FiO$_2$ with severe ROP, even after adjusting for birth weight and GA, and (2) that a multimodal deep learning model can capture richer information from time-series EHR data and provide better prediction results. These findings align with previous studies showing associations between variability in oxygen requirements being associated with outcomes such as bronchopulmonary dysplasia and overall mortality in neonates.[14,33,34]

The pathophysiological mechanism behind the relationship between oxygen fluctuation and end organ damage, such as severe ROP, is still not fully understood. Perelli et al[35] discussed one potential mechanism, demonstrating in a mouse model that transient neonatal exposure to excess oxygen disrupts the formation of the angiogenic astrocyte and suppresses astrocyte proliferation resulting in a disturbance of their spatial patterning. This disturbance caused vascular defects similar to those seen in infants with ROP. This study suggests that fluctuations in oxygen exposure can disrupt the molecular pathways that control astrocyte proliferation, causing an increase in astrocyte numbers and interfering with retinal angiogenesis, ultimately increasing the risk of developing severe ROP. Furthermore, other studies have suggested that fluctuating oxygen exposure may lead to oxidative stress, which might also increase the risk of severe ROP.[36,37]

However, achieving the goal of titrating FiO$_2$ to specific SpO$_2$ targets and limiting fluctuations in FiO$_2$ and SpO$_2$ remains challenging because of the large number of factors that affect SpO$_2$ and how quickly the infants' clinical state can change.[38] As such, improving the FiO$_2$−SpO$_2$ relationship remains a persistent challenge in neonatal care. Recent efforts have been made to develop methods for reducing FiO$_2$ and SpO$_2$ variations, including automated control systems[39] and feedback systems.[40] Nevertheless, further research is still needed to understand the pathobiology of oxygen fluctuations better and to aid in developing effective strategies for minimizing their effects. Our study contributes to this effort by highlighting the importance of monitoring oxygen exposure and, to the extent that it can be achieved, reducing unnecessary fluctuations in iatrogenic oxygen exposure to prevent severe ROP. In the future, further studies will be needed to explore the specific mechanisms related to iatrogenic oxygen fluctuations and the onset of severe ROP.

This study further demonstrated that multimodal neural networks could improve model performance by integrating different sources of input data. The model developed in this study incorporates 2 types of information, including static data (GA and birth weight) and time-series oxygen data. We used the LSTM layers to process and extract useful information from time-series EHR data. Long-short-term memory is a subtype of the recurrent neural network, which can retain long-term information and selectively forget irrelevant information, making it well-suited for extracting information from time-series EHR data.[41] Therefore, LSTM layers can identify temporal patterns and dependencies in the data, allowing the model to make more informed predictions compared to models that use static or aggregation data alone. In our study, the multimodal model demonstrated the highest AUROC at 0.89 ± 0.04, outperforming the SVM models with 4 variables and SVM models trained solely on GA. Our results suggest that utilizing the multimodal deep learning model could be a valuable method for developing a prediction model for the onset of severe ROP. This approach might be useful for other oxygen-related diseases.

This study has several limitations. First, our study was conducted in a single academic center, which could limit the generalizability of the results to other populations. To address this limitation, we intend to validate the models using a similar dataset from an external institute in the future. Second, the sample size of infants in this study was limited to those infants born at the single institution. We excluded infants transferred to OHSU to maintain the accuracy and completeness of oxygen data during the crucial first 30 days after birth. The time period between each oxygen data measurement was not standardized in this study; thus, we used the COV to represent the daily oxygen

fluctuation. In future studies, it would be ideal to use a fixed time stamp for oxygen data to overcome this limitation. Additionally, while birth weight and GA were adjusted for, illness severity was not included in the model. Furthermore, it is important to note that the feasibility of obtaining similar oxygen measurements in smaller hospitals with limited resources may pose unique challenges, but the use of daily aggregation and technological solutions, such as EHR integration or monitoring devices with cloud uploading, may help address potential resource gaps.

In conclusion, abundant time-series oxygen data from EHR can be easily extracted and utilized in predictive modeling for the onset of severe ROP. Our study indicates that not only oxygen exposure but also its fluctuation may play a role in the development of severe ROP, providing direction for future study of pathophysiological factors associated with oxygen exposure and fluctuation in severe ROP development. Additionally, we demonstrated that multimodal neural networks could be a method to extract useful information from time-series data, which may be a valuable methodology for the investigation of other diseases using EHR data.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process

During the preparation of this work, the author used Chat-GPT in order to improve readability and language. After using this tool, the authors reviewed and edited the content to ensure all statements were correct. The authors take full responsibility for the content of the publication.

Footnotes and Disclosures

[1] Department of Ophthalmology, Oregon Health & Science University, Portland, Oregon.

[2] Department of Neonatology, Oregon Health and Science University, Portland, Oregon.

[3] Department of Ophthalmology, University of Colorado (CU) School of Medicine, Denver, Colorado.

[4] Department of Electrical and Computer Engineering, Northeastern University, Boston, Massachusetts.

[5] Department of Ophthalmology and Visual Sciences, University of Illinois at Chicago, Chicago, Illinois.

[6] National Eye Institute, National Institutes of Health, Bethesda, Maryland.

[7] National Library of Medicine, National Institutes of Health, Bethesda, Maryland.

Abbreviations and Acronyms:

AUROC = area under the receiver operating characteristic curve; **COV** = coefficient of variation; **EHR** = electronic health record; **FiO2** = fraction of inspired oxygen; **GA** = gestational age; **LSTM** = long-short-term memory; **OHSU** = Oregon Health & Science University; **ROP** = retinopathy of prematurity; **SpO2** = oxygen saturation; **SVM** = support vector machine.

Correspondence:

J. Peter Campbell, MD, MPH, Department of Ophthalmology, Oregon Health & Science University, 3375 SW Terwilliger Boulevard, Portland, OR 97239. E-mail: campbelp@ohsu.edu.

References

1. Kim SJ, Port AD, Swan R, et al. Retinopathy of prematurity: a review of risk factors and their clinical significance. *Surv Ophthalmol*. 2018;63:618−637.
2. Yau GS, Lee JW, Tam VT, et al. Incidence and risk factors for retinopathy of prematurity in extreme low birth weight Chinese infants. *Int Ophthalmol*. 2015;35:365−373.
3. Blencowe H, Lawn JE, Vazquez T, et al. Preterm-associated visual impairment and estimates of retinopathy of prematurity at regional and global levels for 2010. *Pediatr Res*. 2013;74:35−49.

4. Akçakaya AA, Yaylali SA, Erbil HH, et al. Screening for retinopathy of prematurity in a tertiary hospital in Istanbul: incidence and risk factors. *J Pediatr Ophthalmol Strabismus*. 2012;49:21−25.

5. Akkoyun I, Oto S, Yilmaz G, et al. Risk factors in the development of mild and severe retinopathy of prematurity. *J AAPOS*. 2006;10:449−453.

6. Al-Amro SA, Al-Kharfi TM, Thabit AA, Al-Mofada SM. Risk factors for acute retinopathy of prematurity. *Compr Ther*. 2007;33:73−77.

7. Alajbegovic-Halimic J, Zvizdic D, Alimanovic-Halilovic E, et al. Risk factors for retinopathy of prematurity in premature born children. *Med Arch*. 2015;69:409.

8. Ying G-S, Quinn GE, Wade KC, et al. Predictors for the development of referral-warranted retinopathy of prematurity in the telemedicine approaches to evaluating acute-phase retinopathy of prematurity (e-ROP) study. *JAMA Ophthalmol*. 2015;133:304−311.

9. Pastro J, Toso BRGO. Influence of oxygen in the development of retinopathy of prematurity. *Rev Paul Enferm*. 2019;72:592−599.

10. Das PK, Paul SP, Halim MSPI, et al. Evaluation of SpO2, PaO2, FiO2 levels in developing retinopathy of prematurity. *Int J Contemp Pediatr*. 2021;8:1343.

11. Supplemental therapeutic oxygen for prethreshold retinopathy of prematurity (STOP-ROP), a randomized, controlled trial. I: primary outcomes. *Pediatrics*. 2000;105:295−310.

12. SUPPORT Study Group of the Eunice Kennedy Shriver NICHD Neonatal Research Network, Carlo WA, Finer NN, et al. Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med*. 2010;362:1959−1969.

13. Schmidt B, Whyte RK, Asztalos EV, et al. Effects of targeting higher vs lower arterial oxygen saturations on death or disability in extremely preterm infants: a randomized clinical trial. *JAMA*. 2013;309:2111−2120.

14. BOOST II United Kingdom Collaborative Group, BOOST II Australia Collaborative Group, BOOST II New Zealand Collaborative Group, et al. Oxygen saturation and outcomes in preterm infants. *N Engl J Med*. 2013;368:2094−2104.

15. Shukla A, Sonnie C, Worley S, et al. Comparison of biphasic vs static oxygen saturation targets among infants with retinopathy of prematurity. *JAMA Ophthalmol*. 2019;137:417−423.

16. Chow LC, Wright KW, Sola A; CSMC Oxygen Administration Study Group. Can changes in clinical practice decrease the incidence of severe retinopathy of prematurity in very low birth weight infants? *Pediatrics*. 2003;111:339−345.

17. Saugstad OD, Aune D. In search of the optimal oxygen saturation for extremely low birth weight infants: a systematic review and meta-analysis. *Neonatology*. 2011;100:1−8.

18. Anderson CG, Benitz WE, Madan A. Retinopathy of prematurity and pulse oximetry: a national survey of recent practices. *J Perinatol*. 2004;24:164−168.

19. Srivatsa B, Hagan JL, Clark RH, Kupke KG. Oxygenation factors associated with retinopathy of prematurity in infants of extremely low birth weight. *J Pediatr*. 2022;247:46−52.e4.

20. Di Fiore JM, Bloom JN, Orge F, et al. A higher incidence of intermittent hypoxemic episodes is associated with severe retinopathy of prematurity. *J Pediatr*. 2010;157:69−73.

21. Di Fiore JM, Kaffashi F, Loparo K, et al. The relationship between patterns of intermittent hypoxia and retinopathy of prematurity in preterm infants. *Pediatr Res*. 2012;72:606−612.

22. Cunningham S, Mclntosh N, Fleck B, Elton R. Transcutaneous oxygen levels in retinopathy of prematurity. *Lancet*. 1995;346:1464−1465.

23. York JR, Landers S, Kirby RS, et al. Arterial oxygen fluctuation and retinopathy of prematurity in very-low-birth-weight infants. *J Perinatol*. 2004;24:82−87.

24. McColm JR, Cunningham S, Wade J, et al. Hypoxic oxygen fluctuations produce less severe retinopathy than hyperoxic fluctuations in a rat model of retinopathy of prematurity. *Pediatr Res*. 2004;55:107−113.

25. Zhang D, Yin C, Zeng J, et al. Combining structured and unstructured data for predictive models: a deep learning approach. *BMC Med Inform Decis Mak*. 2020;20:1−11.

26. Si Y, Du J, Li Z, et al. Deep representation learning of patient data from Electronic Health Records (EHR): a systematic review. *J Biomed Inform*. 2021;115:103671.

27. Shickel B, Tighe PJ, Bihorac A, Rashidi P. Deep EHR: a survey of recent advances in deep learning techniques for electronic health record (EHR) analysis. *IEEE J Biomed Health Inform*. 2017;22:1589−1604.

28. Xiao C, Choi E, Sun J. Opportunities and challenges in developing deep learning models using electronic health records data: a systematic review. *Am Med Inform Assoc*. 2018;25:1419−1428.

29. Solares JRA, Raimondi FED, Zhu Y, et al. Deep learning for electronic health records: a comparative review of multiple deep neural architectures. *J Biomed Inform*. 2020;101:103337.

30. Stahlschmidt SR, Ulfenborg B, Synnergren J. Multimodal deep learning for biomedical data fusion: a review. *Brief Bioinform*. 2022;23:bbab569.

31. Summaira J, Li X, Shoib AM, et al. Recent advances and trends in multimodal deep learning: a review. *arXiv*. 2021. https://doi.org/10.48550/arXiv.2105.11087.

32. Kline A, Wang H, Li Y, et al. Multimodal machine learning in precision health: a scoping review. *NPJ Digit Med*. 2022;5:171.

33. Askie LM, Darlow BA, Davis PG, et al. Effects of targeting lower versus higher arterial oxygen saturations on death or disability in preterm infants. *Cochrane Database Syst Rev*. 2017;4:CD011190.

34. Askie LM, Darlow BA, Finer N, et al. Association between oxygen saturation targeting and death or disability in extremely preterm infants in the neonatal oxygenation prospective meta-analysis collaboration. *JAMA*. 2018;319:2190−2201.

35. Perelli RM, O'Sullivan ML, Zarnick S, Kay JN. Environmental oxygen regulates astrocyte proliferation to guide angiogenesis during retinal development. *Development*. 2021;148:dev199418.

36. Perrone S, Vezzosi P, Longini M, et al. Biomarkers of oxidative stress in babies at high risk for retinopathy of prematurity. *Front Biosci (Elite Ed)*. 2009;1:547−552.

37. Weinberger B, Laskin DL, Heck DE, Laskin JD. Oxygen toxicity in premature infants. *Toxicol Appl Pharmacol*. 2002;181:60−67.

38. Hagadorn JI, Furey AM, Nghiem T-H, et al. Achieved versus intended pulse oximeter saturation in infants born less than 28 weeks' gestation: the AVIOx study. *Pediatrics*. 2006;118:1574−1582.

39. Nair V, Loganathan P, Lal MK, Bachman T. Automated oxygen delivery in neonatal intensive care. *Front Pediatr*. 2022;10:915312.

40. Claure N, Bancalari E. Targeting arterial oxygen saturation by closed-loop control of inspired oxygen in preterm infants. *Clin Perinatol*. 2019;46:567−577.

41. Hua Y, Zhao Z, Li R, et al. Deep learning with long short-term memory for time series prediction. *IEEE Commun Mag*. 2019;57:114−119.