Development of artificial neural networks for early prediction of intestinal perforation in preterm infants

Joonhyuk Son1,5, Daehyun Kim2,5, Jae Yoon Na3,5, Donggoo Jung1, Ja-Hye Ahn3, Tae Hyun Kim4,* & Hyun-Kyung Park3,*

Intestinal perforation (IP) in preterm infants is a life-threatening condition that may result in serious complications and increased mortality. Early Prediction of IP in infants is important, but challenging due to its multifactorial and complex nature of the disease. Thus, there are no reliable tools to predict IP in infants. In this study, we developed new machine learning (ML) models for predicting IP in very low birth weight (VLBW) infants and compared their performance to that of classic ML methods. We developed artificial neural networks (ANNs) using VLBW infant data from a nationwide cohort and prospective web-based registry. The new ANN models, which outperformed all other classic ML methods, showed an area under the receiver operating characteristic curve (AUROC) of 0.8832 for predicting IP associated with necrotizing enterocolitis (NEC-IP) and 0.8797 for spontaneous IP (SIP). We tested these algorithms using patient data from our institution, which were not included in the training dataset, and obtained an AUROC of 1.0000 for NEC-IP and 0.9364 for SIP. NEC-IP and SIP in VLBW infants can be predicted at an excellent performance level with these newly developed ML models. https://github.com/kdhRiick2222/Early-Prediction-of-Intestinal-Perforation-in-Preterm-Infants.

Intestinal perforation (IP) in preterm infants is a life-threatening condition that requires urgent surgical intervention. Delayed diagnosis of IP can lead to serious complications and even death in certain patients. There are two major causes of IP in premature infants: necrotizing enterocolitis (NEC) and spontaneous intestinal perforation (SIP)1–4. NEC and SIP are acquired diseases, which means that IP occurs due to the actions of various risk factors. IP associated with NEC (NEC-IP) is an advanced stage of NEC in which NEC has progressed to a severe state. On the other hand, SIP can occur unexpectedly without preceding symptoms or signs5. Many previous studies have investigated the risk factors and pathogenesis of these two types of IP. According to the literature, possible risk factors include prematurity, low birth weight, perinatal use of medications (ibuprofen, steroids, indomethacin, and inotropic agents), maternal chorioamnionitis, patent ductus arteriosus (PDA), and high-grade intraventricular hemorrhage (IVH grade III or IV)5,7–20. Despite these efforts, due to the nature of retrospective studies with a small number of patients, there is thus far no consensus concerning risk factors. Furthermore, the pathogenesis of IP in premature infants is still not clearly understood; as a result, neonatologists and pediatric surgeons have difficulty predicting and preventing the occurrence of NEC-IP and SIP21–28.

Recently, artificial intelligence (AI) technologies have been used to solve various medical problems, including diagnosing diseases, choosing treatments, and predicting the risks and progression of diseases29–30. Machine learning (ML) algorithms, specifically artificial neural networks (ANNs), allow more accurate estimation of complex patterns through learning from training data and are able to predict which patients will develop a disease as an early and supportive diagnostic tool for clinicians. To date, only a few studies have used ML algorithms to predict IP in premature infants. Irles et al.11 reported an ML model predicting NEC-IP, and Lure et al.32 reported an ML model to differentiate NEC-IP and SIP prior to surgical interventions. Lin33 et al. reported an ML model to differentiate NEC-IP and SIP prior to surgical interventions. Lin33 et al. reported an
of dimension. Model 1 is a typical neural approach and can produce promising results when a large number of
dropout. It takes a 54-dimensional vector as an input and, after the application of all layers, renders feature vectors
layers; each layer is formed from a block (Fig. 1a) arranged with an activation function, batch normalization and
problem and facilitate stable training. Specifically, Model 1 is a binary classifier that is composed of 5 hidden
Proposed machine learning algorithms.

Baseline characteristics of the patients. NEC-IP intestinal perforation associated with necrotizing enterocolitis, SIP spontaneous intestinal perforation, SD standard deviation, PROM premature rupture of membranes, RDS respiratory distress syndrome, IVH intraventricular hemorrhage, PDA patent ductus arteriosus.

ML technique for individualized NEC risk scores using intestinal microbiota data. However, despite these efforts, studies concerning ML technologies in neonatal IP are still lacking, and there are no definitive tools to predict NEC-IP and SIP in advance. Therefore, our study aimed to develop our own ML models for predicting IP in very low birth weight (VLBW) infants using a nationwide cohort dataset and to compare the performance of the new models with that of classic ML models. Furthermore, to ensure that our novel predictive ML models were valid in real neonatal intensive care (NICU) settings, we tested new ML models with individual cases from our institution and assessed whether these new models achieved high predictive accuracy.

Results

Baseline characteristics of the patients. The baseline characteristics of the infants are shown in Table 1. Of the 12,555 patients, 521 patients had NEC-IP (4.1%), and 208 patients had SIP (1.7%). Three groups of VLBW infants were compared for analysis: (A) control group without NEC-IP or SIP, (B) patients with NEC-IP, and (C) patients with SIP. Factors such as low gestational age, low birth weight, need for resuscitation at delivery, respiratory distress syndrome (RDS) and administration of surfactant, use of certain medications (steroids, ibuprofen, and inotropes), hypotension, IVH grade III/IV, sepsis, and PDA on treatment were significantly associated with NEC-IP and SIP separately. For ANNs, it is well known that simply adding layers can lead to performance improvement in a previously published study31, but determined the hyperparameters (e.g., the number of channels) carefully and added more advanced techniques, such as batch normalization34 and drop-out35, to avoid the overfitting problem when the training dataset is insufficient. Therefore, we stacked the layers more deeply than the conventional models in diagnosis, as introduced in a previously published study31, but determined the hyperparameters (e.g., the number of channels) carefully and added more advanced techniques, such as batch normalization34 and drop-out35, to avoid the overfitting problem and facilitate stable training. Specifically, Model 1 is a binary classifier that is composed of 5 hidden layers; each layer is formed from a block (Fig. 1a) arranged with an activation function, batch normalization and dropout. It takes a 54-dimensional vector as an input and, after the application of all layers, renders feature vectors of dimension. Model 1 is a typical neural approach and can produce promising results when a large number of

|                        | A                      | B                      | C                      | p-value |
|------------------------|------------------------|------------------------|------------------------|---------|
| No perforation (n = 11,826) | NEC-IP (n = 521) | SIP (n = 208) | NEC-IP vs. SIP          |
| Gestational age (weeks, mean ± SD) | 28.58 ± 3.02 | 25.84 ± 2.41 | 26.25 ± 2.23 | <0.001  <0.001 |
| Birth weight (g, mean ± SD) | 1092.38 ± 284.05 | 820.03 ± 249.41 | 858.38 ± 248.15 | <0.001  <0.001 |
| Sex - male - n (%) | 5919 (50.1) | 285 (54.7) | 130 (62.5) | 0.038  <0.001 |
| Maternal chorioamnionitis - n (%) | 3426 (29.0) | 162 (31.1) | 64 (30.8) | 0.296  0.571 |
| PROM - n (%) | 4073 (34.4) | 196 (37.6) | 80 (38.5) | 0.135  0.227 |
| Antenatal steroid use - n (%) | 9341 (79.0) | 414 (79.5) | 169 (81.3) | 0.794  0.427 |
| Resuscitation at delivery - n (%) | 10,472 (88.6) | 489 (93.9) | 204 (98.1) | <0.001  <0.001 |
| RDS - n (%) | 9080 (76.8) | 487 (93.9) | 198 (95.2) | <0.001  <0.001 |
| Sulfactant use - n (%) | 9114 (77.1) | 489 (93.9) | 200 (96.2) | <0.001  <0.001 |
| Steroid use - n (%) | 2589 (21.9) | 198 (38.0) | 91 (43.8) | <0.001  <0.001 |
| Indomethacin use - n (%) | 54 (0.5) | 8 (1.5) | 2 (1.0) | 0.001  0.252 |
| Ibuprofen use - n (%) | 3539 (29.0) | 232 (44.5) | 83 (39.9) | <0.001  0.002 |
| Hypotension - n (%) | 2838 (24.0) | 325 (62.4) | 130 (62.5) | <0.001  <0.001 |
| Inotropic use - n (%) | 548 (4.6) | 59 (11.3) | 26 (12.5) | <0.001  <0.001 |
| IVH grade 3,4 - n (%) | 934 (7.9) | 146 (28.0) | 48 (23.1) | <0.001  <0.001 |
| Sepsis - n (%) | 2224 (18.8) | 266 (51.1) | 91 (43.8) | <0.001  <0.001 |
| PDA on medication - n (%) | 3615 (30.6) | 241 (46.3) | 87 (41.8) | <0.001  <0.001 |
| PDA ligation - n (%) | 1145 (9.7) | 142 (27.3) | 48 (23.1) | <0.001  <0.001 |

Table 1. Clinical characteristics of the patients. NEC-IP intestinal perforation associated with necrotizing enterocolitis, SIP spontaneous intestinal perforation, SD standard deviation, PROM premature rupture of membranes, RDS respiratory distress syndrome, IVH intraventricular hemorrhage, PDA patent ductus arteriosus.
datasets (e.g., NEC) are available. However, the outcomes from Model 1 can be vulnerable in real-world scenarios where the training dataset is insufficient.

Therefore, we attempted to further improve our baseline Model 1 to predict NEC-IP and SIP more accurately since the lack of data problem for NEC-IP and SIP is more serious than that for NEC. To alleviate this problem, we developed a new approach that transferred information from the network to predict NEC values to help estimate NEC-IP and SIP. Note that transfer learning in the field of deep learning is one of the most widely used approaches to solve lack of data problems. Therefore, we present additional models (Model 2 and Model 3) that can exploit information achieved from the NEC dataset to predict SIP and NEC-IP.

Specifically, Model 2 is composed of two different MLPs. One branch predicts NEC, and the other branch predicts either SIP or NEC-IP. Notably, at the 4th layer of the MLP branch for NEC-IP and SIP in Model 2, feature vectors from the third layer of the network for NEC are fed as an additional input. By concatenating the feature vectors from NEC, we can utilize the information for NEC in predicting NEC-IP/SIP.

Unlike that of Model 2, the network architecture of Model 3 is the same as that of Model 1. We employed conventional transfer learning to utilize information from NEC to estimate NEC-IP/SIP and fine-tuned the pretrained Model 1 to the specific NEC-IP/SIP datasets.

Comparison of performance between classic ML models and proposed ANN models. We provide prediction results in Table 2 to compare traditional ML models with our neural approaches. We observed that the proposed neural approach (Model 1) outperformed traditional learning-based methods in terms of area under the receiver operating characteristic curve (AUROC) scores for all cases (NEC, NEC-IP, and SIP).

Moreover, our extended networks (i.e., Model 2 and Model 3), which we designed to mitigate the lack of data problem for NEC-IP (521 positive cases) and SIP (208 positive cases), showed improved results in predicting NEC-IP and SIP over the baseline network (Model 1). In particular, compared to Model 1, which was trained on the NEC dataset, our proposed methods for NEC-IP and SIP (Model 2 and Model 3) exhibited improvements. Notably, Model 2 directly utilizes features distilled from Model 1, Model 3 fine-tunes Model 1 for the prediction of either NEC-IP or SIP, and Model 2 outperforms Model 3 in terms of AUROC, as shown in Table 3 and Fig. 2. The performance metrics of these models using balanced validation dataset are described in Supplementary Tables 1 and 2. This performance improvement achieved in Model 2 and Model 3 indicates that information extracted to predict NEC can also be used to predict NEC-IP and SIP more accurately. Moreover, the proposed direct feature distillation of Model 2 rather than the conventional fine-tuning approach (i.e., Model 3) can be a recommendable option for addressing problems with limited data.

Application of the new ANN models in a real clinical environment. To assess the feasibility of the algorithm in a real clinical environment, we tested our newly developed algorithms using the patient data from our institution, which were not included in the training dataset. A total of 57 VLBW infants who were born at our hospital between 2019 and 2020 were included in this test analysis. Among the three ANN models, Model 2 achieved the highest AUROC scores: 1.0000 for the prediction of NEC-IP and 0.9364 for the prediction of SIP (Table 4 and Fig. 3).

Discussion

Our novel ML algorithms predicted NEC-IP and SIP in VLBW infants with favorable AUROC scores, outperforming all other classic ML algorithms. One of our algorithms exhibited an AUROC score of 1.0000 for predicting NEC-IP and 0.9364 for predicting SIP in real clinical settings. Our study shows the integration of a vast nationwide dataset with ML, and the resulting model can be used to predict the possibility of specific medical conditions in patients who may not perfectly represent the signs and symptoms of the disease.

Among preterm infants in the NICU, various medical problems exist at the same time, and multidisciplinary collaboration is required to make medical decisions for these patients. IP is one of the most devastating medical conditions that occurs in the NICU. Early diagnosis, swift judgment, and prompt surgical intervention are required to prevent severe complications and poor outcomes. However, as we explained earlier, predicting the occurrence of IP is difficult. By running this algorithm, it is possible to analyze every preterm infant who is admitted to the NICU and predict each patient’s likelihood of developing NEC-IP and SIP. Thus, early predictions of these serious medical conditions could provide clinicians with a much more stable management environment, enabling them to make better treatment decisions.

In recent years, AI and big data have been increasingly integrated into medicine because it is difficult for the unaided human clinician to acquire all the latest published knowledge, as is required by modern evidence-based medicine. AI is an important resource for medical research in that it can efficiently process large amounts of data. Additionally, AI can produce consistent and unbiased results without fatigue. Several studies in healthcare research have reported sufficient or even better risk prediction by AI methods compared to existing models. Especially these days, in solving difficult problems with big data which have complex distribution, neural approaches exhibit excellent performance. Since neural networks can have non-linearity from adding layers and fine-tune parameters by transfer learning, they can get the upper hand in complex tasks. To sum up, ANN is an appropriate model for IP prediction, as it can handle imbalanced big data efficiently. However, even if the overall cohort is large, diseases with low prevalence always suffer from a lack of data. Due to the character of ML, it can produce excellent results only with a large amount of training data. Thus, applying ML to diseases with low prevalence remains a challenge. To overcome the data imbalance problem, several studies have applied data processing techniques such as oversampling and undersampling as we did in our ANN models.

To further improve the performance of our models, we modified them by adding another branch or pre-training it based on an algorithm that predicts other relevant disease with higher prevalence in order to help...
Figure 1. Architecture of machine learning models. (a) Model 1 is a baseline neural network based on the conventional multilayer perceptron (MLP) architecture. (b) Model 2 is composed of two different MLPs. One branch predicts NEC, and the other branch predicts either NEC-IP or SIP. The feature vectors from the 3rd layer of the network for NEC are concatenated with the 4th layer of the MLP branch for NEC-IP and SIP. (c) Model 3 has the same network architecture as that of Model 1. Pretrained Model 1 for NEC was further fine-tuned to estimate NEC-IP/SIP.
registry of VLBW infants. Their clinical data were collected from 74 participating NICUs across the country and evaluated using datasets (Table 5). Moreover, to facilitate the network training procedure, a data preprocessing step was set up. The study was approved by the Hanyang University Institutional Review Board (IRB No. 2013-06-025-043). Data obtained from the parents of each infant, and all methods were carried out following relevant guidelines. This study was conducted retrospectively for this study. Prior to participation in the KNN registry, informed consent was obtained. The main challenges of AI methods are due to the nature of self-extracted data from large datasets. Our study highlighted that it is necessary to customize the algorithm for each disease to apply an ML model in real clinical settings, especially if the disease is rare.

Although our study showed a favorable outcome, it had its share of limitations. First, a limited number of factors are included because only data collected from the Korean Neonatal Network (KNN) were used. It is expected that the collection of further IP-related data, such as clinical symptoms, vital signs, and radiologic findings, will enable the model to produce better results. In addition, the limitations of AI studies, such as representation, homogeneity, and accuracy, were observed in this study. Another limitation is that it is difficult to determine how AI methods generate results due to the nature of self-extracted data from large datasets.

In conclusion, we developed our own ANN models to predict IP early in VLBW infants, and these new models achieved higher accuracy than classic ML algorithms. To our knowledge, this is the first study to develop an ML model to predict both NEC-IP and SIP using nationwide VLBW infant data. In addition, the newly proposed ANN models showed excellent performance within real NICU clinical settings. When more clinical data, such as vital signs, radiologic findings, biomarkers, and laboratory results, are gathered, we believe that a more accurate ML model will be developed, thereby achieving early prediction of these serious medical conditions and better clinical outcomes for VLBW infants.

**Methods**

**Data collection.** We derived data from infants registered in the KNN, a nationwide prospective cohort registry of VLBW infants. Their clinical data were collected from 74 participating NICUs across the country and analyzed retrospectively for this study. Prior to participation in the KNN registry, informed consent was obtained from the parents of each infant, and all methods were carried out following relevant guidelines. This study was approved by the Hanyang University Institutional Review Board (IRB No. 2013-06-025-043).

The cohort comprised 12,555 VLBW infants born between January 5, 2013, and December 31, 2018, weighing less than 1500 g.

**Disease definitions.** NEC was defined according to Bell's modified staging grade ≥II. NEC-IP was diagnosed when patients with NEC underwent any kind of abdominal surgical intervention (peritoneal drainage or laparotomy). SIP was defined when the patients underwent surgical intervention due to IP and the surgeon found no predisposing causes, such as NEC, intestinal atresia, or meconium plug. The full list of 54 variables used in ML analysis is shown in Supplementary Table 3.

**Comparisons of baseline characteristics.** A total of 54 variables, including various maternal and perinatal factors, were collected for ML. Among them, 18 clinical factors that were proposed as possible risk factors for either NEC-IP or SIP in previous studies were analyzed using conventional statistical methods. Student’s t-test was performed to analyze the continuous variables, and the chi-squared test was used to analyze categorical variables. Statistical significance was set at \( P < 0.05 \). The Statistical Package for the Social Sciences version 22.0 for Windows software program (IBM Corp., Armonk, NY, USA) was used in all statistical analyses.

**Data preprocessing.** Our dataset was composed of 12,555 infants in total; we divided them into training and evaluation datasets (Table 5). Moreover, to facilitate the network training procedure, a data preprocessing step was set up.
First, to solve the missing data problem in the given dataset, we imputed the missing (null) values with plausible values. Specifically, input values were categorized into ordinal, continuous and categorical types. In the case of ordinal inputs, we imputed the null values with the mode (most frequently occurring) values, and in the case of continuous inputs, we imputed the null values with the mean values. Finally, missing values of categorical inputs were replaced with the median values. After recovering the data, we mapped the dataset between 0 and 1 by using min–max normalization. Finally, to mitigate the data imbalance problem (e.g., 11,703 negative and 852 positive cases for NEC in Table 5), which is a common intrinsic feature of disease datasets, we oversampled the smaller category (positive cases) and undersampled the bigger category (negative cases), as suggested in previous studies59,60.

Training. We used the binary cross entropy (BCE) loss to train Model 1, and Model 1 was separately trained to predict NEC, NEC-IP, and SIP. For Model 3, pretrained Model 1 for NEC was further fine-tuned, and the separately updated parameters using the BCE loss were used to predict NEC-IP and SIP.

|             | Model 1 | Model 2 | Model 3 |
|-------------|---------|---------|---------|
| AUROC       | 0.8128  | –       | –       |
| NEC         | 0.8665  | 0.8832  | 0.8692  |
| NEC-IP      | 0.8498  | 0.8797  | 0.8633  |
| F1-score    | 0.7701  | –       | –       |
| NEC         | 0.7181  | 0.8934  | 0.7273  |
| NEC-IP      | 0.8059  | 0.8204  | 0.8041  |

Table 3. Performance of the proposed ANN models in predicting NEC, NEC-IP, and SIP. ANN artificial neural network, NEC necrotizing enterocolitis, NEC-IP intestinal perforation associated with necrotizing enterocolitis, SIP spontaneous intestinal perforation, AUROC area under the receiver operating characteristic curve. *Scores were found with balanced validation dataset. Positive cases were oversampled.

|             | Model 1 | Model 2 | Model 3 |
|-------------|---------|---------|---------|
| AUROC       | 0.6745  | –       | –       |
| NEC         | 1.0000  | 1.0000  | 0.8704  |
| NEC-IP      | 0.9000  | 0.9364  | 0.8818  |
| F1-score    | 0.6903  | –       | –       |
| NEC         | 0.8571  | 0.9076  | 0.8710  |
| NEC-IP      | 0.7241  | 0.7179  | 0.7925  |

Table 4. Test results for 57 cases within a real NICU environment. NICU neonatal intensive care unit, AUROC area under the receiver operating characteristic curve, NEC necrotizing enterocolitis, NEC-IP intestinal perforation associated with necrotizing enterocolitis, SIP spontaneous intestinal perforation. *Scores were found with balanced validation dataset. Positive cases were oversampled.

Figure 2. Receiver operating characteristic curves of proposed ML models for (a) NEC prediction, (b) NEC-IP prediction, and (c) SIP prediction.
Unlike Model 1 and Model 3, Model 2 was trained in a two-stage manner. In the first stage, the left branch of the network (Fig. 1b) was trained for NEC with the BCE loss. Then, the BCE loss as well as the BCE for NEC in the left branch were used to jointly optimize the network for either NEC-IP or SIP. In both training steps, oversampling and undersampling technique were used to alleviate the data imbalance problem. We employed the two-stage training scheme rather than using the one-shot joint training approach, as we could retain more stability during the training from the two-stage approach.

To train the three proposed models (i.e., Model 1, Model 2, Model 3), the Adam optimizer with a learning rate of 0.0001 was used for loss minimization. We used a dropout rate of 0.2 and a batch size of 128. Instead of fixing the number of iterations, we stopped the training using the early stopping technique to avoid the overfitting

Figure 3. Receiver operating characteristic curves of proposed ML models from 57 test cases. (a) NEC prediction, (b) NEC-IP prediction, and (c) SIP prediction.

| Training dataset | Evaluation dataset |
|------------------|-------------------|
| NEC              | NEC-IP            | SIP               |
| Negative         | Positive          | Negative          | Positive | Total  |
| 10,000           | 770               | 1703              | 82       | 12,555 |
| 430              | 2035              | 91                |
| 170              | 2348              | 38                |

Table 5. Size of training and evaluation datasets. NEC necrotizing enterocolitis, NEC-IP intestinal perforation associated with necrotizing enterocolitis, SIP spontaneous intestinal perforation.

Figure 4. Flowchart of data processing. Input values were categorized into ordinal, continuous and categorical types. To solve the data imbalance problem, oversampling and undersampling technique were applied. Then, the data were normalized.
problem. When the loss increased seven times in growth, we stopped the training (i.e., early stop) and took the parameters just prior to the loss increase. The models were implemented in PyTorch with Python, and the Scikit-learn library was used to evaluate the results (e.g., AUROC, F1-Score, and ROC curve).

Data availability
According to the Korean Neonatal Network (KNN) Publication Ethics Policy, all information about patients is confidential. The information contained in the data must be protected as confidential, and only available to individuals who have access for the permitted research activity.

Received: 16 December 2021; Accepted: 7 July 2022
Published online: 15 July 2022

References
1. Hunter, C. J., Chokshi, N. & Ford, H. R. Evidence vs experience in the surgical management of necrotizing enterocolitis and focal intestinal perforation. J. Perinatol. 28(Suppl 1), S14-17 (2008).
2. Karla, K., Anttila, A., Iber, T., Pakarinen, M. & Koivusalo, A. Outcomes of surgery for necrotizing enterocolitis and spontaneous intestinal perforation in Finland during 1986–2014. J. Pediatr. Surg. 53(10), 1928–1932 (2018).
3. Vonghavivit, K. & Underwood, M. A. Intestinal perforation in the premature infant. J. Neonatal Perinatal Med. 10(3), 281–289 (2017).
4. Zozaya, C. et al. Neonatal and growth outcomes of extremely preterm neonates with necrotizing enterocolitis or spontaneous intestinal perforation. J. Pediatr. Surg. 56(2), 309–316 (2021).
5. Sahi, J. et al. Intestinal perforation in very preterm neonates: Risk factors and outcomes. J. Perinatol. 35(8), 595–600 (2015).
6. Pumberger, W., Mayr, M., Kohlhauser, C. & Weninger, M. Spontaneous localized intestinal perforation in very-low-birth-weight infants: A distinct clinical entity different from necrotizing enterocolitis. J. Am. Coll. Surg. 195(6), 796–803 (2002).
7. Ragouilliaux, C. J., Keeney, S. E., Hawkins, H. K. & Rowen, J. I. Maternal factors in extremely low birth weight infants who develop spontaneous intestinal perforation. Pediatrics 120(6), e1458–e1464 (2007).
8. Rayyan, M. et al. Risk factors for spontaneous localized intestinal perforation in the preterm infant. J. Matern. Fetal Neonatal Med. 31(19), 2617–2623 (2018).
9. Rose, A. T. & Patel, R. M. A critical analysis of risk factors for necrotizing enterocolitis. Semin. Fetal Neonat. Med. 23(6), 374–379 (2018).
10. Samuels, N., van de Graaf, R. A., de Jonge, R. C. J., Reiss, I. K. M. & Vermeulen, M. J. Risk factors for necrotizing enterocolitis in neonates: A systematic review of prognostic studies. BMC Pediatr. 17(1), 105 (2017).
11. Youn, Y. A., Kim, E. K. & Kim, S. Y. Necrotizing enterocolitis among very-low-birth-weight infants in Korea. J. Korean Med. Sci. 30(Suppl 1), S57–80 (2015).
12. Arnautovic, T. I. et al. Antenatal risk factors associated with spontaneous intestinal perforation in preterm infants receiving postnatal indomethacin. J. Pediatr. 232(59–64), e51 (2021).
13. Kandrapu, H. et al. Association of co-exposure of antenatal steroid and prophylactic indomethacin with spontaneous intestinal perforation. J. Pediatr. 235(34–41), e31 (2021).
14. Stavel, M. et al. Effect of prophylactic indomethacin administration and early feeding on spontaneous intestinal perforation in extremely low-birth-weight infants. J. Perinatol. 37(2), 188–193 (2017).
15. Yilmaz, Y. et al. Preeclampsia is an independent risk factor for spontaneous intestinal perforation in very preterm infants. J. Matern. Fetal Neonatal Med. 27(12), 1248–1251 (2014).
16. Houben, C. H. et al. Spontaneous intestinal perforation: The long-term outcome. Eur. J. Pediatr. Surg. 27(4), 346–351 (2017).
17. Prasad, U., Mohnani, A. & Hussain, N. Spontaneous intestinal perforation associated with premature twin infants. J. Neonatal Perinatal Med. 14(3), 403–409 (2021).
18. Maheshwari, A., Patel, R. M. & Christensen, R. D. Anemia, red blood cell transfusions, and necrotizing enterocolitis. Semin. Pediatr. Surg. 27(1), 47–51 (2018).
19. Wang, Z. L. et al. Risk factors of necrotizing enterocolitis in neonates with sepsis: A retrospective case-control study. Int. J. Immunopathol. Pharmacol. 34, 2058738420963818 (2020).
20. Talli, M. M. et al. Spontaneous intestinal perforation after oral ibuprofen treatment of patent ductus arteriosus in two very-low-birthweight infants. Acta Paediatr. 93(7), 999–1001 (2004).
21. Ndour, D. et al. Adverse events related to ibuprofen treatment for patent ductus arteriosus in premature neonates. Arch. Pediatr. 15(8), 452–455 (2020).
22. El Manouni El Hassani, S. et al. Predictive factors for surgical treatment in preterm neonates with necrotizing enterocolitis: A multicenter case-control study. Eur. J. Pediatr. 180(2), 617–625 (2021).
23. Eaton, S., Rees, C. M. & Hall, N. J. Current research on the epidemiology, pathogenesis, and management of necrotizing enterocolitis. Neonatology 111(4), 423–430 (2017).
24. Neu, J. Necrotizing enterocolitis: The mystery goes on. Neonatology 106(4), 289–295 (2014).
25. Frost, B. L., Modr, B. P., Jakic, T. & Caplan, M. S. New medical and surgical insights into neonatal necrotizing enterocolitis: A review. JAMA Pediatr. 171(1), 83–88 (2017).
26. Cho, S. X. et al. Characterization of the pathoimmunology of necrotizing enterocolitis reveals novel therapeutic opportunities. Nat. Commun. 11(1), 5794 (2020).
27. Krittawanong, C. et al. Future direction for using artificial intelligence to predict and manage hypertension. Curr. Hypertens. Rep. 20(9), 75 (2018).
28. Obermeyer, Z. & Emanuel, E. J. Predicting the future: Big data, machine learning, and clinical medicine. N. Engl. J. Med. 375(13), 1216–1219 (2016).
29. Zhang, Z. et al. Derivation and validation of an ensemble model for the prediction of agitation in mechanically ventilated patients maintained under light sedation. Crit. Care Med. 49(3), e279–e287 (2021).
30. Na, I. Y. et al. Artificial intelligence model comparison for risk factor analysis of patent ductus arteriosus in nationwide very low birth weight infants cohort. Sci. Rep. 11(1), 1–10 (2021).
31. Irel, C. et al. Estimation of neonatal intestinal perforation associated with necrotizing enterocolitis by machine learning reveals new key factors. Int. J. Environ. Res. Public Health. 15(11), 2509 (2018).
32. Lu, A. C. et al. Using machine learning analysis to assist in differentiating between necrotizing enterocolitis and spontaneous intestinal perforation: A novel predictive analytic tool. J. Pediatr. Surg. 56(10), 1703–1710 (2020).
33. Lin, Y. C., Salleb-Aouissi, A. & Hooven, T. A. Interpretable prediction of necrotizing enterocolitis from machine learning analysis of premature infant stool microbiota. BMC Bioinform. 23(1), 104 (2022).
34. Joffe, S. & Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. in Proceedings of the 32nd International Conference on Machine Learning. vol 37, 448–456 (PMLR, 2015).
Acknowledgements

This research was supported by funds (2019-ER7103-02#) from Research of the Korea National Institute of Health and research funds from Hanyang University MEB (Global Center for Developmental Disorders, HY-202100000002865), an Institute of Information & Communications Technology Planning & Evaluation (ITTP) grant funded by the Korean government (MSIT) [No. 2020-0-01373, Artificial Intelligence Graduate School Program (Hanyang University)], and the Bio & Medical Technology Development Program of the National Research Foundation of Korea (NRF) funded by the Korean government (MSIT) [Grant Number NRF-2019M3ESD1A01069363].

Author contributions

H.K.P and T.H.K. conceptualized, designed, and supervised the study. J.S. and J.H.A. were involved in data collection and initial statistical analysis. T.H.K., D.K., and D.J. performed the computational analyses, development of the algorithms, and interpretation of the data. J.S., D.K. and J.Y.N. were involved in the writing, and editing of the paper. All authors reviewed and approved the revised manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-022-16273-5.

Correspondence and requests for materials should be addressed to T.H.K. or H.-K.P.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
