Supplemental Material

This appendix presents additional information

Supplement to: A simple and interpretable severe intraventricular hemorrhage prediction model for extremely low birth weight infants using machine learning

Corresponding author

Felipe Yu Matsushita, MD – felipe.matsushita@hc.fm.usp.br

Department of Pediatrics,

Instituto da Criança, Faculty of Medicine, University of São Paulo

Av. Dr. Enéas Carvalho de Aguiar, 647

São Paulo, São Paulo, Brazil, 05403901

ORCID: 0000-0002-5106-7497

**Supplementary material**

**eFigure 1. Decision-tree model**



Variables included in the decision tree: Inotropic therapy and C-section delivery. Number in node represents percentage of event to occur inside the corresponding group. Percentage of sIVH in patients without inotropic necessity = 11%. Percentage of sIVH in patients with inotropic necessity and vaginal delivery = 85%.

**eTable 1. Stepwise logistic regression using all features**

|  |  |  |  |
| --- | --- | --- | --- |
| Feature | Odds Ratio | 95% CI | P value |
| 5-minute APGAR score | 0.815 | 0.646 – 1.021 | 0.075 |
| Delivery | 0.153 | 0.041 – 0.5262 | **0.003** |
| Antenatal corticoid | 0.448 | 0.146 – 1.302 | 0.145 |
| Highest pCO2 | 1.031 | 1.003 – 1.062 | **0.031** |
| Inotrope | 6.787 | 2.39 – 21.13 | **<0.001** |

**eFigure 2. Decision-tree model**



Variables included in decision model: delivery type, highest pCO2, and inotropic therapy. Number in node represents percentage of event to occur inside the corresponding group. Percentage of sIVH in patients with inotropic necessity and vaginal delivery = 85%. Percentage of sIVH in patients with inotropic necessity, c-section delivery and highest pCO2 lower than 44mmHg = 11%.

**eFigure 3. Correlation matrix analyzing all variables**



**eTable 2. Logistic regression after correlation matrix (AIC 115.4 / AUC 0.739)**

|  |  |  |  |
| --- | --- | --- | --- |
| Features | Odds Ratio | 95% CI | P value |
| Lowest pH | 0.009 | 0.00003 – 1.963 | 0.092 |
| Gestational age | 0.835 | 0.543 – 1.252 | 0.390 |
| Lowest mean blood pressure | 1.051 | 0.887 – 1.250 | 0.557 |
| 5-minute APGAR score | 0.574 | 0.388 – 0.795 | **0.002** |
| Delivery | 0.029 | 0.002 – 0.245 | **0.002** |
| Birth weight | 1.009 | 1.003 – 1.015 | **0.002** |
| Lowest temperature | 1.255 | 0.565 – 2.842 | 0.575 |
| Antenatal corticoid | 0.851 | 0.211 – 3.455 | 0.818 |
| Highest mean blood pressure | 0.939 | 0.863 – 1.012 | 0.119 |
| Worst Platelet | 0.980 | 0.962 – 0.994 | **0.016** |
| Gender | 0.507 | 0.134 – 1.760 | 0.294 |
| Lowest pCO2 | 1.030 | 0.946 – 1.124 | 0.491 |
| Diuresis | 0.922 | 0.570 – 1.458 | 0.731 |
| Positive Blood Culture | 0.084 | 0.002 – 1.938 | 0.145 |
| C-reactive protein | 0.970 | 0.924 – 1.011 | 0.178 |
| Chorioamnionitis | 0.280 | 0.027 – 2.389 | 0.260 |
| Pneumothorax | 0.085 | 0.002 – 1.753 | 0.131 |
| hsPDA | 0.602 | 0.104 – 3.355 | 0.558 |
| Inotrope | 33.44 | 6.04 – 299.6 | **<0.001** |

**eFigure 4. Decision-tree model**



Variables included in decision tree model: 5-minute APGAR score, delivery, birth weight, worst platelet, and inotropic therapy. Number in node represents percentage of event to occur inside the corresponding group. Percentage of sIVH in patients without inotropes = 11%. Percentage of sIVH in patients with inotropic therapy necessity, c-section delivery and platelet lower than 60.000mm3 = 67%.

**eTable 3. Stepwise logistic regression after correlation matrix (AIC 101 / AUC 0.768)**

|  |  |  |  |
| --- | --- | --- | --- |
| Feature | Odds Ratio | 95% CI | P value |
| Lowest pH | 0.035 | 0.0005 – 2.100 | 0.110 |
| 5-minute APGAR score | 0.637 | 0.481 – 0.818 | **<0.001** |
| Delivery | 0.048 | 0.007 – 0.248 | **<0.001** |
| Birth Weight | 1.006 | 1.002 – 1.011 | **0.003** |
| Worst Platelet | 0.983 | 0.968 – 0.996 | **0.019** |
| Positive Blood Culture | 0.063 | 0.002 – 1.112 | 0.068 |
| Inotrope | 10.92 | 3.06 – 47.98 | **<0.001** |

**eTable 4. Logistic regression using confirmed important variables using Boruta algorithm**

|  |  |  |  |
| --- | --- | --- | --- |
| Feature | Odds Ratio | 95% CI | P value |
| Lowest pH | 6.51 | 0.00008 - 10619 | 0.749 |
| Inotrope | 6.53 | 1.87 – 25.7 | **0.004** |
| Gestational age | 0.938 | 0.695 – 1.245 | 0.666 |
| Highest pCO2 | 1.04 | 0.989 – 1.10 | 0.132 |
| Lowest base excess | 0.941 | 0.750 – 1.172 | 0.595 |
| Delivery | 0.153 | 0.03 – 0.60 | **0.008** |
| Fluid bolus | 1.003 | 0.829 – 1.234 | 0.971 |
| Lowest diastolic blood pressure | 0.969 | 0.789 – 1.160 | 0.749 |
| Lowest mean blood pressure | 1.049 | 0.871 – 1.283 | 0.619 |

**eFigure 5. Decision-tree model**



Variables included in decision tree model: Lowest pH, inotropic therapy, gestational age, highest pCO2, Lowest base excess, delivery, fluid bolus necessity, lowest mean blood pressure, and lowest diastolic blood pressure. Number in node represents percentage of event to occur inside the corresponding group. Percentage of sIVH in patients with pH > 7.2 = 6%. Percentage of sIVH in patients with pH < 7.2, vaginal delivery and pCO2 > 54mmhg = 100%.

**eFigure 6. Decision-tree model**



Variable included in decision tree model: CRIB II score. Number in node represents percentage of event to occur inside the corresponding group. Percentage of sIVH in patients with CRIB score < 17 = 20%. Percentage of sIVH in patients with CRIB score equal or higher than 17 = 86%.

**eFigure 7. Decision-tree model**



Variable included in decision tree model: Gestational age, mechanical ventilation antenatal corticoid, 5-minute APGAR, birth weight, cesarean section, male gender. Number in node represents percentage of event to occur inside the corresponding group. Percentage of sIVH in patients with gestational age equal or greater than 25 weeks and 5-minute APGAR score < 3 = 57%.