

FELIPE YU MATSUSHITA

Lactato como preditor de mortalidade em recém-nascidos pré-termo de extremo baixo peso usando inteligência artificial

Tese apresentada à Faculdade de Medicina da Universidade de São Paulo para obtenção do título de Doutor em Ciências

Programa de Pediatria

Orientadora: Prof^a. Dr^a. Vera Lucia Jornada Krebs

São Paulo

2022

FELIPE YU MATSUSHITA

Lactato como preditor de mortalidade em recém-nascidos pré-termo de extremo baixo peso usando inteligência artificial

Tese apresentada à Faculdade de Medicina da Universidade de São Paulo para obtenção do título de Doutor em Ciências

Programa de Pediatria

Orientadora: Prof^a. Dr^a. Vera Lucia Jornada Krebs

São Paulo

2022

Dados Internacionais de Catalogação na Publicação (CIP)

Preparada pela Biblioteca da
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Matshushita, Felipe Yu
Lactato como preditor de mortalidade em recém-nascidos pré-termo de extremo baixo peso usando inteligência artificial / Felipe Yu Matshushita. -- São Paulo, 2022.

Tese (doutorado) -- Faculdade de Medicina da Universidade de São Paulo.
Programa de Pediatria.
Orientadora: Vera Lucia Jornada Krebs.

Descritores: 1. Recém-nascido 2. Recém-nascido prematuro 3. Ácido láctico 4. Mortalidade 5. Inteligência artificial 6. Aprendizado de máquina

USP/FM/DBD-418/22

Responsável: Erinalva da Conceição Batista, CRB-8 6755

DEDICATÓRIA

À Jéssica, minha esposa e fonte de inspiração;

Aos meu pais, Fumihiro e Cecília, Francisco, exemplos de perseverança

À minha irmã, Victoria, companheira de corpo e de alma.

AGRADECIMENTOS

À Dra. Vera Krebs, minha orientadora, pela dedicação de tempo e sabedoria para a realização desse projeto.

A todas as assistentes do Centro Neonatal do Instituto da Criança, pela minha formação acadêmica e pelo apoio pessoal.

Essa dissertação está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: adaptado de *International Committee of Medical Journals Editors* (Vancouver)

Universidade de São Paulo. Faculdade de Medicina. Serviço de Biblioteca e Documentação. *Guia de apresentação de dissertações, teses e monografias*. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 3a ed. São Paulo: Serviço de Biblioteca e Documentação; 2011.

Abreviaturas dos títulos dos periódicos de acordo com *List of Journals Indexed in Index Medicus*.

SUMÁRIO

Lista de abreviaturas

Lísta de símbolos

Lista de tabelas

Resumo

Abstract

1	INTRODUÇÃO	1
2	OBJETIVOS.....	5
3	MÉTODOS.....	7
	3.1 Casuística	8
	3.2 Métodos.....	9
	3.3 Análise estatística	12
	3.4 Aprovação por Comissão de Ética	12
4	RESULTADOS	13
5	DISCUSSÃO	19
6	CONCLUSÕES	23
7	ANEXOS.....	25
8	REFERÊNCIAS.....	30

LISTA DE ABREVIATURAS

ML	- Machine Learning
ROC	- Receiver operating characteristic
RN	- Recém-nascido
PRISMA	- Preferred Reporting Items for Systematic Reviews and Meta-Analysis
PIG	- Pequeno para Idade Gestacional
AIG	- Adequado para Idade Gestacional
GIG	- Grande para Idade Gestacional
PC	- Perímetro Cefálico
PN	- Peso de Nascimento
IC	- Intervalo de Confiança
CRIB-II	- Clinical Risk Index for Babies
BE	- Base Excess
IQR	- Intervalo Interquartil
IG	- Idade Gestacional
n	- Número
UTIN	- Unidade de Tratamento Intensivo Neonatal
FO	- Ferida Operatória
ICrHCFMUSP	- Instituto da Criança e do Adolescente do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo
SNAPPE-II	- Neonatal Acute Physiology Perinatal Extension-II

LISTA DE SÍMBOLOS

<	- Menor que
>	- Maior que
≤	- Menor ou igual
≥	- Maior ou igual
%	- Porcentagem
+	- Adição
-	- Subtração
=	- Igual a
/	- Divisão
G	- Gramas
Kg	- Kilogramas
p	- Índice de probabilidade
N	- Tamanho da amostra
β	- Coeficiente de regressão
Cm	- Centímetros
M	- Masculino
F	- Feminino
S	- Sim
N	- Não

LISTA DE TABELAS

Tabela 1 – Características gerais e demográficas – Página 15

Tabela 2 – Desempenho de previsão de mortalidade em 24 horas de modelos de Machine Learning usando gasometria arterial e lactato – Página 16

Tabela 3 – Desempenho de previsão de mortalidade em 24 horas de modelos de Machine Learning usando gasometria arterial sem lactato – Página 17

RESUMO

Matsushita FY. *Lactato como preditor de mortalidade em recém-nascidos pré-termo de extremo baixo usando inteligência artificial*. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2022.

Introdução: A mortalidade neonatal representa o maior componente da mortalidade infantil. Portanto, modelos preditivos para prever o risco de óbito em neonatos são altamente recomendados, especialmente entre recém-nascidos pré-termo. Apesar da relevância do tema, há poucos modelos utilizando critérios objetivos para avaliar o risco de óbito em tempo real nesta população. **Objetivo:** Avaliar a utilidade do lactato sérico como preditor de mortalidade em recém-nascidos pré-termo com peso de nascimento inferior a 1000 gramas. **Métodos:** Estudo de coorte retrospectivo, analisando recém-nascidos pré-termo com peso de nascimento menor que 1000 gramas admitidos em unidade de terapia intensiva neonatal terciária de centro único em São Paulo, Brasil, no período de seis anos. Foram incluídos todos os recém-nascidos com peso de nascimento < 1000 g, com pelo menos uma gasometria arterial com dosagem simultânea de lactato sérico coletados durante a internação. Foram excluídos os recém-nascidos que faleceram em sala de parto, recém-nascidos portadores de malformação maior e aqueles transferidos para outro serviço antes da alta. Foram utilizadas técnicas de inteligência artificial, com análise dos parâmetros preditivos através de algoritmos de *Machine Learning*. As variáveis analisadas foram gasometria arterial (pH, pCO₂, HCO₃ e *base excess*), lactato sérico e características clínicas (peso, idade gestacional, pequeno para a idade gestacional, escore CRIB II, sexo feminino, Apgar, via de parto, gemelaridade, corticóide antenatal, intubação endotraqueal na sala de parto, necessidade de epinefrina na sala de parto, corioamnionite e menor temperatura nas primeiras 12 horas de vida). Foi comparada a *performance* de três métodos diferentes de *Machine Learning*, para avaliar o risco de óbito em 24 horas: Regressão Logística, *Extreme Gradient Boosting* e *AutoML Tables*. Os dados foram randomicamente divididos em dois subgrupos: um subgrupo de treinamento

(80%) para aperfeiçoamento de hiperparâmetros para criação de um modelo plausível e um subgrupo de teste (20%) para avaliar a *performance* do modelo. Após a determinação dos hiperparâmetros ideais foram calculados os valores de precisão, a área sob a curva (ROC), *recall* e acurácia. Todas as análises foram realizadas usando Python versão 3.6.9 e Google Cloud Platform. **Resultados:** Foram incluídos 257 neonatos. A mediana da idade gestacional foi de 27,1 (26 – 29,1) semanas e a mediana do peso ao nascer foi de 746 (600 – 880) gramas. Houve 1.932 amostras de gasometria com níveis séricos de lactato correspondentes. O modelo Extreme Gradient Boosting com lactato obteve a maior área sob a curva (0.898). Base excesso, lactato e pH foram, em ordem de importância, as variáveis mais importantes associadas a mortalidade em 24 horas. **Conclusão:** Em recém-nascidos pré-termo de extremo baixo peso, a incorporação de medidas de lactato e gasometria arterial em modelos preditivos de mortalidade pode melhorar a estratificação de risco em tempo real. Os modelos tradicionais de classificação logística parecem ser superados pelos algoritmos de aprendizado de máquina. Demonstramos que os modelos de algoritmos de aprendizado de máquina *Extreme Gradient Boosting* pode ser ferramenta de apoio importante na tomada de decisões mais precisas e rápidas e na estratificação de risco clínico em unidade de terapia intensiva neonatal.

Palavras-chave: Recém-Nascido, Recém-nascido prematuro, Ácido láctico, mortalidade, inteligência artificial, aprendizado de máquina

ABSTRACT

Matsushita FY. Risk prediction model for 24-hour mortality in preterm infants using lactate and blood gas analysis: A machine learning approach. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2022.

Introduction: Neonatal mortality represents the largest component of child mortality. Therefore, predictive models to predict the risk of death in neonates are highly recommended, especially among preterm neonates. Despite the relevance of the topic, there are few models using objective criteria to assess the risk of death in real time in this population. **Objective:** To evaluate the practicability of serum lactate as a predictor of mortality in preterm newborns with birth weight below 1000 grams. **Methods:** A retrospective cohort study analyzing preterm newborns with birth weight less than 1000 grams admitted to a single-center tertiary neonatal intensive care unit in São Paulo, Brazil, over a six-year period. All newborns with birth weight less than 1000 grams, with at least one arterial blood gas analysis with simultaneous measurement of serum lactate collected during hospitalization were included. Newborns who died in the delivery room, newborns with major malformations and those transferred to another service before discharge were excluded. Artificial intelligence techniques were used, with analysis of predictive models through Machine Learning algorithms. The variables analyzed were arterial blood gas analysis (pH, pCO₂, HCO₃ and base excess), serum lactate and clinical characteristics (weight, gestational age, small for gestational age, CRIB II score, female sex, Apgar, mode of delivery, twin birth, antenatal corticosteroids, endotracheal intubation in the delivery room, epinephrine in the delivery room necessity, chorioamnionitis and lowest temperature in the first 12 hours of life). The performance of three different Machine Learning methods to assess the risk of death in 24 hours was compared: Logistic Regression, Extreme Gradient Boosting and AutoML Tables. The data were randomly divided into two subgroups: a training subgroup (80%) to improve hyperparameters to create a plausible model and a test subgroup (20%) to evaluate the model's performance. After determining the ideal hyperparameters,

precision values, area under the curve (ROC), recall and accuracy were calculated. All analyzes were performed using Python version 3.6.9 and Google Cloud Platform. **Results:** 257 neonates were included. The median gestational age was 27.1 (26 – 29.1) weeks and the median birth weight was 746 (600 – 880) grams. There were 1932 blood gas samples with corresponding serum lactate levels. The Extreme Gradient Boosting model with lactate had the highest area under the curve (0.898). Base excess, lactate and pH were, in order of importance, the most important variables associated with 24-hour mortality. **Conclusion:** In extremely low birth weight preterm infants, the incorporation of lactate and arterial blood gas measurements into predictive models of mortality can improve real-time risk stratification. Traditional logistic classification models seem to be surpassed by machine learning algorithms. We demonstrated that Extreme Gradient Boosting machine learning algorithm models can be an important support tool in making more accurate and faster decisions and in clinical risk stratification in a neonatal intensive care unit.

Key words: newborn, preterm, mortality, artificial intelligence, machine learning

1 INTRODUÇÃO

Embora tenha havido uma importante redução da mortalidade infantil nas últimas décadas, observa-se que os óbitos neonatais ainda representam o maior componente da mortalidade infantil. Este comportamento está relacionado às elevadas taxas de prematuridade e baixo peso ao nascer (1,2).

Uma análise sistemática de todos os dados de domínio público sobre mortalidade neonatal obtidos em bancos de dados das Nações Unidas para Estimativa de Mortalidade Infantil, mostrou que os neonatos representam a parcela mais vulnerável desta população em todo o mundo. Os autores utilizaram um modelo de regressão bayesiano, para analisar dados de diferentes fontes, obtidos em sistemas de registro vital, sistemas de registro de amostra e pesquisas domiciliares, para estimar a taxa de mortalidade neonatal de cada país entre 1990 e 2017. Neste período a mortalidade neonatal global diminuiu 51%, passando de 36,6 mortes/1000 nascidos vivos em 1990, para 18 mortes/1000 nascidos vivos em 2017, sendo o número estimado de mortes neonatais durante o mesmo período de 5,0 milhões e 2,5 milhões, respectivamente (3).

Entre os recém-nascidos pré-termo de extremo baixo peso (<1000g) algumas condições podem aumentar significativamente o risco de desfecho desfavorável, como o quadro clínico de hipoperfusão tecidual e choque. Apesar da relevância do tema, ainda não há consenso no manejo de instabilidade hemodinâmica em recém-nascidos. Sua avaliação é extremamente difícil, pois os achados clínicos possuem pouca correlação com a mortalidade. Por exemplo, a pressão arterial é um parâmetro indispensável na avaliação de choque nas crianças e adultos, porém, até o momento, não existe uma definição de hipotensão na população neonatal (4). Neste contexto, o lactato sérico poderia ser uma variável objetiva adicional para avaliação hemodinâmica nesta população.

Por ser um produto do metabolismo anaeróbico, a hiperlactatemia resultante do desbalanço entre a oferta e demanda de oxigênio, pode ser considerada um marcador de hipoxemia tecidual (5,6). Em pacientes com sepse em situações de instabilidade hemodinâmica este exame é um biomarcador bastante utilizado como preditor potencial de mortalidade. A avaliação do lactato sérico faz parte do diagnóstico de choque séptico em adultos, conforme o consenso internacional de sepse. Embora classicamente o lactato seja analisado no contexto de sepse, foi demonstrado em adultos que este marcador pode ser utilizado em qualquer situação de hipoperfusão tecidual, associada ou não à infecção (7,8,9,10).

No Brasil recentemente foi realizada uma revisão sistemática e metanálise (11) para avaliar a associação entre lactato sérico e morbimortalidade. Os autores seguiram recomendações baseadas no PRISMA (Itens de Relatório Preferenciais para Revisões Sistemáticas e Meta-análises) e *Cochrane Center for Reviews and Dissemination*. A estratégia de busca foi desenvolvida de acordo com as recomendações do PRESS (Revisão por Pares de Estratégias de Busca Eletrônica). Foram incluídos 49 estudos observacionais e 14 estudos de acurácia de testes diagnósticos. O risco de viés foi avaliado pela Escala de Newcastle-Ottawa para estudos observacionais e pela ferramenta QUADAS-2 para estudos de acurácia de testes diagnósticos. Entre 3184 publicações, 63 estudos preencheram todos os critérios de elegibilidade totalizando 46069 neonatos. Foi demonstrada associação entre níveis mais elevados de lactato sérico e mortalidade (diferença média padrão, -1,09 [IC 95%, -1,46 a -0,73]).

Embora o lactato sérico constitua um marcador importante de hipoperfusão em pacientes criticamente doentes, sua utilidade em identificar e estratificar neonatos com maior risco de eventos adversos ainda não foi estabelecida. Existem poucos estudos em recém-nascidos, especialmente em pré-termos com peso de nascimento inferior a 1000g. Estes neonatos constituem uma população de alto risco para óbito, sendo altamente recomendada a avaliação de fatores associados à mortalidade durante a internação hospitalar.

A hipótese do presente estudo é que existe associação entre o lactato sérico e a taxa de mortalidade em recém-nascidos de extremo baixo peso. Assim como em crianças e adultos, o lactato pode ser usado como preditor de mortalidade nestes neonatos.

2 OBJETIVOS

O presente estudo tem como objetivo:

Avaliar o lactato sérico como preditor de mortalidade em recém-nascidos pré-termo com peso de nascimento inferior a 1000 gramas usando técnicas de inteligência artificial e *Machine Learning*.

3 MÉTODOS

3.1 Casuística

Modelo do Estudo

Foi realizado um estudo coorte retrospectivo, analisando recém-nascidos pré-termo com peso de nascimento menor que 1000 gramas nascidos no Centro Neonatal do Instituto da Criança e do Adolescente do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo HCFMUSP, centro de referência nacional de gestação de alto risco em São Paulo, Brasil, entre 2012 e 2017.

Critérios de inclusão

Todos os recém-nascidos com peso de nascimento inferior a 1000 gramas, admitidos no Centro Neonatal do Instituto da Criança e do Adolescente do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo entre 01 de janeiro de 2012 e 31 de dezembro de 2017, com pelo menos uma gasometria arterial com lactato pareado coletado durante a internação foram incluídos no estudo.

Critérios de exclusão

Foram excluídos os recém-nascidos que faleceram em sala de parto, recém-nascidos portadores de malformação maior, cardiopatia complexa e aqueles transferidos para outro serviço antes da alta hospitalar.

A idade gestacional foi determinada segundo a ultrassonografia fetal realizada até a vigésima semana de gestação ou, na falta deste exame, através da data da última menstruação.

A avaliação da gravidade clínica foi feita através do escore CRIB II (Clinical Risk Index for Babies).

Para a dosagem de gasometria arterial e lactato foi coletado o volume mínimo de 0,5 mL de sangue. O lactato sérico foi analisado em mmol/L.

Os exames foram realizados no aparelho ABL 800 FLEX®, método de membrana de íon seletivo.

3.2 Métodos

As características de base incluídas foram:

- Idade gestacional (em semanas)
- Peso de nascimento (em gramas)
- Escore CRIB II (Clinical Risk Index For Babies)

Pequeno para Idade Gestacional, definido como peso de nascimento menor que p10 pela curva de Fenton (14)

- Sexo feminino
- Apgar de 5º minuto
- Parto vaginal
- Gemelaridade
- Corticóide antenatal
- Intubação endotraqueal na sala de parto
- Necessidade de epinefrina na sala de parto
- Corioamnionite
- Menor temperatura nas primeiras 12 horas de vida

Parâmetros preditivos:

Foram introduzidas sete variáveis nos algoritmos de *Machine Learning* (12), sendo incluídos os seguintes parâmetros:

- Gasometria arterial (pH, pCO₂, HCO₃ e *base excess*);

- Lactato sérico (analisado em mmol/L).
- Dias de vida
- Idade gestacional corrigida

Modelos de *Machine Learning*

Foi comparada a performance de três métodos diferentes de *Machine Learning*, para avaliar o risco de óbito em 24 horas:

- Regressão Logística
- *Extreme Gradient Boosting*
- *AutoML Tables*

Os dados foram randomicamente divididos em dois subgrupos: um subgrupo de treinamento (80%) para aperfeiçoamento de hiperparâmetros para criação de um modelo plausível e um subgrupo de teste (20%) para avaliar a performance do modelo.

Hiperparâmetros são parâmetros extras que precisam ser estabelecidos antes do processo de aprendizado do modelo para melhorar a performance do algoritmo. Todos os casos com falta de dados foram excluídos da análise.

A classificação de regressão logística é uma técnica de aprendizado de máquina que prevê resultados usando variáveis dependentes e uma função logit. É frequentemente usada como base de comparação para classificações binárias devido à simplicidade de sua construção e alto desempenho (15,16). *Extreme Gradient Boosting* é uma técnica que utiliza modelos fracos que são adicionados um de cada vez e ajustados para corrigir os erros de previsão feitos pelos modelos anteriores (18). Para a *Extreme Gradient Boosting* foram usados os seguintes hiperparâmetros:

- `LEARN_RATE = 0.1`
- `BOOSTER_TYPE = 'GBTREE'`
- `MAX_TREE_DEPTH = 3`

- SUBSAMPLE = 0.85
- EARLY_STOP = TRUE

AutoML Tables é um produto desenvolvido pela plataforma Google Cloud que automaticamente treina para múltiplos modelos de Machine Learning (linear, feedforward deep neural network, gradiente boosted decision tree, AdaNet, Ensembles) e determina qual o melhor modelo. O desfecho primário foi óbito em 24 horas após a coleta da gasometria arterial com lactato.

Após a determinação dos hiperparâmetros ideais foram determinados, foram calculados os valores de precisão, área sob a curva (ROC), recall e acurácia. Todas as análises foram realizadas usando Python versão 3.6.9 e Google Cloud Platform (21).

3.3 Análise estatística

As variáveis contínuas foram testadas para normalidade usando o teste de Kolmogorov-Smirnov. Os resultados estão apresentados como números com proporções (%), médias com desvio padrão ou mediana com intervalo interquartil. Para comparação de características demográficas, foi utilizado o teste do qui-quadrado para variáveis categóricas e teste de Mann-Whitney para variáveis contínuas. Foi considerado significativo o valor de $p < 0,05$.

3.4 Aprovação por Comissão de Ética

O estudo foi aprovado pela Comissão de Ética para Análise de Projetos de Pesquisa do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, CAAE 15762719.6.0000.0068, em 12/06/2019.

4 RESULTADOS

Foram incluídos 257 neonatos que tiveram pelo menos uma gasometria e um lactato sérico pareado durante a internação na unidade de cuidados intensivos neonatal. As características basais são apresentadas na Tabela 1. A mediana da idade gestacional foi de 27,1 (26 – 29,1) semanas e a mediana do peso ao nascer foi de 746 (600 – 880) gramas. Na população do estudo, houve 109 pacientes que evoluíram a óbito, representando uma taxa de mortalidade de 42%. Houve 1.932 amostras de gasometrias arteriais com níveis séricos de lactato correspondentes.

Todos os modelos demonstraram boa precisão (91 – 94%) e resultados AUROC (0,807 – 0,898) ao usar medidas de gasometria com lactato como variáveis. No entanto, a *recall* (9 - 29%) foi baixa. O algoritmo Extreme Gradient Boosting obteve a maior pontuação AUROC (0,898), acurácia (94,1%) e precisão (87,5%) (Tabela 2). Em seguida, foi utilizado o AutoML Tables para determinar a importância de cada parâmetro associado à mortalidade em 24 horas, e os três principais parâmetros foram, em ordem: base excess, lactato e níveis de pH (Figura 1).

Quando o lactato foi removido dos modelos de aprendizado de máquina, a pontuação AUROC do modelo de Extreme Gradiente Boosting caiu consideravelmente (0,807) (Tabela 3).

Foi criado então um modelo de predição usando o algoritmo Extreme Gradient Boosting, que obteve a maior pontuação AUROC, acurácia e precisão (Figura 2) usando o Google Cloud Platform. Este modelo de predição permite que o usuário informe os parâmetros da gasometria (pH, pCO₂, HCO₃ e base excess), dias de vida do recém-nascido, idade gestacional corrigida e lactato, e o algoritmo calcula, automaticamente, a probabilidade de óbito nas próximas 24 horas.

Tabela 1. Características gerais e demográficas

Variável	Total (n = 257)	Sobrevida (n = 148)	Óbito (n = 109)	P-value
Idade Gestacional (semanas)	27.1	28	26.3	< 0.001
mediana (IQR)	(26-29.1)	(26.5-30.1)	(25-27.4)	
Peso de Nascimento (g)	746	822	610	< 0.001
mediana (IQR)	(600- 880)	(700-940)	(530-760)	
CRIB II escore	12	11	14	< 0.001
mediana (IQR)	(10-14)	(9-12)	(12-15)	
Pequeno para Idade Gestacional	125	73	52	0.798
n (%)	(48.6)	(49.3)	(47.7)	
Sexo feminino	123	74	49	0.424
n (%)	(47.9)	(50)	(45)	
Apgar no 5º minuto de vida	8	8 (7 – 9)	8	< 0.001
mediana (IQR)	(6 – 9)		(5-8)	
Parto vaginal	41	13 (8.8)	28	< 0.001
n (%)	(16)		(25.7)	
Gemelaridade	77	33	44	0.002
n (%)	(30)	(22.3)	(40.4)	
Corticóide antenatal, n (%)	129	79	51	0.234
	(50.2)	(53.4)	(46.8)	
Intubação na sala de parto, n (%)	167	81	86	< 0.001

	(65)	(54.7)	(78.9)	
Adrenalina na sala de parto	28	11	17	0.038
n (%)	(10.9)	(7.4)	(15.6)	
Corioamnionite	29	12	17	0.061
n (%)	(11.3)	(8.1)	(15.6)	
Menor temperatura até 12h de vida	35	35	34.5	< 0.001
mediana (IQR)	(34.1-35.8)	(34.5-35.8)	(33.8-35.2)	

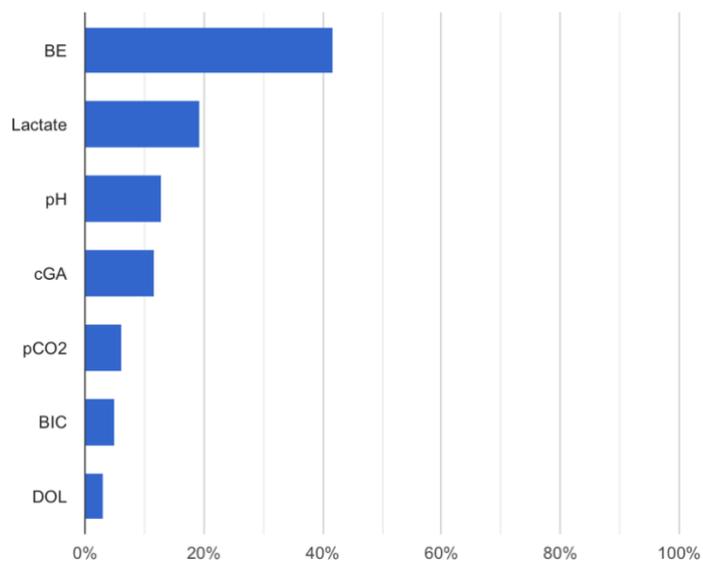
Tabela 2. Desempenho de previsão de mortalidade em 24 horas de modelos de Machine Learning usando gasometria arterial e lactato

		Accuracy	Precision	Recall	AUROC
Extreme	Gradient	0.941	0.875	0.25	0.898
Boosting					
AutoML Tables		0.933	0.833	0.29	0.826
Regressão Logística		0.913	0.375	0.09	0.807

Tabela 3. Desempenho de previsão de mortalidade em 24 horas de modelos de Machine Learning usando gasometria arterial sem lactato

	Accuracy	Precision	Recall	AUROC
AutoML Tables	0.938	1.000	0.294	0.857
Regressão Logística	0.940	0.600	0.130	0.848
Extreme Gradient Boosting	0.942	0.666	0.173	0.807

Figura 1. Importância das variáveis (em %) para predição de mortalidade



BE: Base excess; cGA: Corrected gestational age; BIC: HCO_3 ; DOL: Days of life

Figura 2. Calculando o risco de óbito (%) em 24 horas através dos parâmetros de gasometria arterial e lactato em modelo de Machine Learning

Test your model with sample input data

Request an online prediction by sending your input data instances as a JSON object.

[Learn how to format input data](#)

```
{"instances": [[7.3, 15.0, 55.0, -5.0, 2, 15.0, 30.0]]}
```

TEST

```
{
  "predictions": [
    [
      0.036613889038562775,
      0.9633861780166626
    ]
  ]
}
```

5 DISCUSSÃO

Em Unidade de Tratamento Intensivo Neonatal (UTIN), a estimativa precisa da mortalidade é uma ferramenta valiosa que auxilia os profissionais de saúde. No entanto, a maioria dos modelos preditivos de mortalidade em recém-nascidos pré-termo limita-se à admissão na UTIN. Para avaliar o risco de mortalidade na admissão na UTIN, o *Score for Neonatal Acute Physiology Perinatal Extension-II* (SNAPPE-II) inclui sinais vitais, laboratoriais e características basais (22). O escore *Clinical Risk for Infants and Babies* (CRIB-II) considera sexo, peso ao nascer, idade gestacional, temperatura e excesso de base para avaliar o risco de mortalidade na admissão na UTIN (13). Recentemente foram propostos outros modelos que predizem a mortalidade neonatal após a admissão, como o *Transport Risk Index of Physiologic Stability Version II* (TRIPS-II) (23), o *Neonatal Mortality Risk Among Neonates Weighing 2000 g or Less* (NMR-2000) (24) e o *Preterm Infants Survival Assessment* (PISA) (25). No entanto, eventos dinâmicos na unidade de tratamento intensivo podem ter impacto no risco inicial.

A combinação de características de linha de base e recursos em tempo real fornece uma avaliação mais individualizada do risco de mortalidade, que evolui conforme as mudanças que ocorrem no estado clínico do paciente. Para superar esse problema, Feng J et al. (21) Lee J et al.(23) Jaskari J et al.(26), criaram modelos preditivos de mortalidade baseados em dados de sinais vitais coletados durante a permanência na UTIN. No entanto, os sinais vitais como frequência respiratória, frequência cardíaca e pressão arterial podem variar amplamente entre os neonatos e não há consenso sobre o que constitui um intervalo de referência normal (27,28). Além disso, no estudo de Lee J et al. (23) as variáveis basais (peso ao nascer e idade gestacional ao nascimento) foram mais importantes do que as leituras de sinais vitais.

Os resultados obtidos no presente estudo demonstram que a utilização de amostras de gasometria arterial e lactato em algoritmos *de Machine Learning*

como Extreme Gradient Boosting podem ser utilizados para ajudar a prever a mortalidade em 24 horas em recém-nascidos de extremo baixo peso.

Desta forma, propomos um modelo preditivo de mortalidade utilizando recursos laboratoriais objetivos em tempo real, para aplicação em recém-nascidos pré-termo de alto risco para óbito. Até o momento, na literatura consultada por nós, não há publicações sobre este tipo de modelo.

Os valores de lactato e gasometria podem ser obtidos de forma rápida e simples. Essa avaliação é atualmente uma prática clínica padrão em unidades de tratamento intensivo e não requer esforço adicional. Concordamos com von Auenmueller KI et al (29), quando afirmam que a falta de viés do examinador, que contribui para a objetividade, é uma vantagem sobre o exame clínico (29). O presente estudo combina a capacidade preditiva dos modelos de aprendizado de máquina com a objetividade de uma gasometria arterial e lactato no sangue, que é um teste simples, prontamente disponível e amplamente utilizado.

Os resultados obtidos indicam que os algoritmos de aprendizado de máquina, podem apresentar vantagens em relação ao modelo tradicional de classificação logística na estratificação de risco clínico em Unidade de Tratamento Intensivo Neonatal. A melhor avaliação de risco em tempo real em recém-nascidos pré-termo de extremo baixo peso criticamente doentes é altamente desejável, considerando a vulnerabilidade desta população.

A pesquisa tem algumas limitações, que devem ser observadas. Primeiramente, trata-se de um estudo retrospectivo de centro único, fato que pode interferir na generalização dos achados. Em segundo lugar, conjuntos de dados maiores auxiliam os modelos preditivos de aprendizado de máquina, podendo ser necessário um estudo com maior número de pacientes. Vale a pena notar que o uso da gasometria para prever a mortalidade é um problema desequilibrado, pois o número de amostras de gases no sangue é muito superior ao número de eventos (morte). Em dados não balanceados, podemos adquirir uma alta precisão e pontuação AUC simplesmente prevendo que todas as observações pertencem à classe majoritária. Por último, é importante notar que

nossos modelos possuíram alta precisão, porém com baixo *recall*, portanto não devem ser utilizados para rastreamento.

6 CONCLUSÕES

O estudo permitiu chegar às seguintes conclusões:

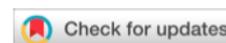
- A incorporação de medidas de lactato e gasometria arterial em modelos preditivos de mortalidade pode melhorar a estratificação de risco em tempo real em recém-nascidos pré-termo de extremo baixo peso.
- Os modelos tradicionais de classificação logística parecem ser superados por algoritmos de aprendizado de máquina.
- O modelo *Extreme Gradient Boosting* parece ser uma ferramenta de apoio a decisões precisas, rápidas e robustas para estratificação de risco clínico desta população na Unidade de Tratamento Intensivo Neonatal.

7 ANEXOS

Anexo A – Publicação desta tese na Revista F1000 Research

F1000Research

F1000Research 2022, 11:444 Last updated: 17 AUG 2022



RESEARCH ARTICLE

Risk prediction model for 24-hour mortality in preterm infants using lactate and blood gas analysis: A machine learning approach and retrospective cohort study [version 1; peer review: awaiting peer review]

Felipe Yu Matsushita , Vera Lúcia Jornada Krebs, Werther Brunow de Carvalho

Pediatrics, Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, São Paulo, São Paulo, 05403-000, Brazil

V1 First published: 20 Apr 2022, 11:444
<https://doi.org/10.12688/f1000research.110711.1>
 Latest published: 20 Apr 2022, 11:444
<https://doi.org/10.12688/f1000research.110711.1>

Open Peer Review

Approval Status *AWAITING PEER REVIEW*

Any reports and responses or comments on the article can be found at the end of the article.

Abstract

Background: This study aimed to evaluate the performance of machine learning algorithms using lactate and arterial blood gas parameters to predict the imminent risk of death in extremely low birth weight infants.

Methods: A retrospective cohort study analyzing preterm infants with birth weight less than 1000 grams in a single-center tertiary neonatal intensive care unit in São Paulo, Brazil, between 2012 and 2017 was carried out. We included all infants with at least one arterial blood gas analysis with paired serum lactate. To assess 24-hour mortality risk, we conducted three machine learning algorithms (Logistic Regression, Extreme Gradient Boosting, and AutoML Tables).

Results: We analyzed 1932 blood gas samples with matched lactate measurements. Our study population had a median gestational age of 27.1 (26 – 29.1) weeks and a median birth weight of 746 (600 – 880) grams. The Extreme Gradient Boosting model with lactate achieved the highest area under the receiver operating characteristic (AUROC) of 0.898. Base excess, lactate, and pH were, in order of importance, the most important features associated with 24-hour mortality.

Conclusions: Incorporating lactate and blood gas samples into real-time mortality predictive models may aid to identify those preterm infants with a higher risk of death.

Keywords

preterm, machine learning, artificial intelligence, prediction, mortality

Anexo B – Publicação na European Journal of Pediatrics - 2022

European Journal of Pediatrics (2022) 181:1085–1097
<https://doi.org/10.1007/s00431-021-04298-3>

ORIGINAL ARTICLE



Identifying clinical phenotypes in extremely low birth weight infants—an unsupervised machine learning approach

Felipe Yu Matsushita¹ · Vera Lúcia Jornada Krebs¹ · Werther Brunow de Carvalho¹

Received: 1 September 2021 / Revised: 14 October 2021 / Accepted: 16 October 2021 / Published online: 3 November 2021
© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2021

Abstract

There is increasing evidence that patient heterogeneity significantly hinders advancement in clinical trials and individualized care. This study aimed to identify distinct phenotypes in extremely low birth weight infants. We performed an agglomerative hierarchical clustering on principal components. Cluster validation was performed by cluster stability assessment with bootstrapping method. A total of 215 newborns (median gestational age 27 (26–29) weeks) were included in the final analysis. Six clusters with different clinical and laboratory characteristics were identified: the “Mature” (Cluster 1; $n = 60$, 27.9%), the mechanically ventilated with “adequate ventilation” (Cluster 2; $n = 40$, 18.6%), the mechanically ventilated with “poor ventilation” (Cluster 3; $n = 39$, 18.1%), the “extremely immature” (Cluster 4; $n = 39$, 18.1%), the neonates requiring “Intensive Resuscitation” in the delivery room (Cluster 5; $n = 20$, 9.3%), and the “Early septic” group (Cluster 6; $n = 17$, 7.9%). In-hospital mortality rates were 11.7%, 25%, 56.4%, 61.5%, 45%, and 52.9%, while severe intraventricular hemorrhage rates were 1.7%, 5.3%, 29.7%, 47.2%, 44.4%, and 28.6% in clusters 1, 2, 3, 4, 5, and 6, respectively ($p < 0.001$).

Conclusion: Our cluster analysis in extremely preterm infants was able to characterize six distinct phenotypes. Future research should explore how better phenotypic characterization of neonates might improve care and prognosis.

What is Known:

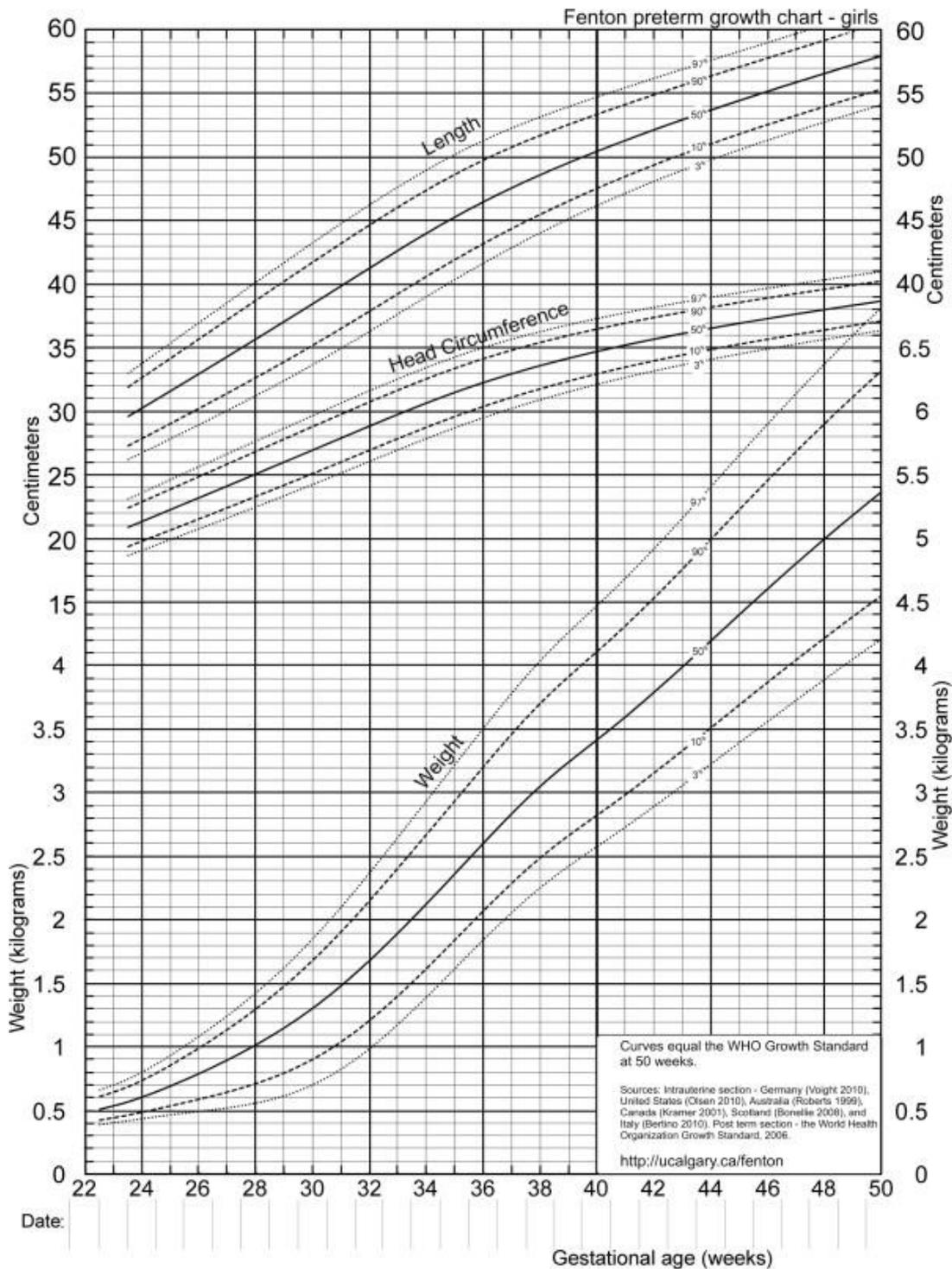
- Patient heterogeneity is becoming more acknowledged as a cause of clinical trial failure.
- Machine learning algorithms can find patterns within a heterogeneous group.

What is New:

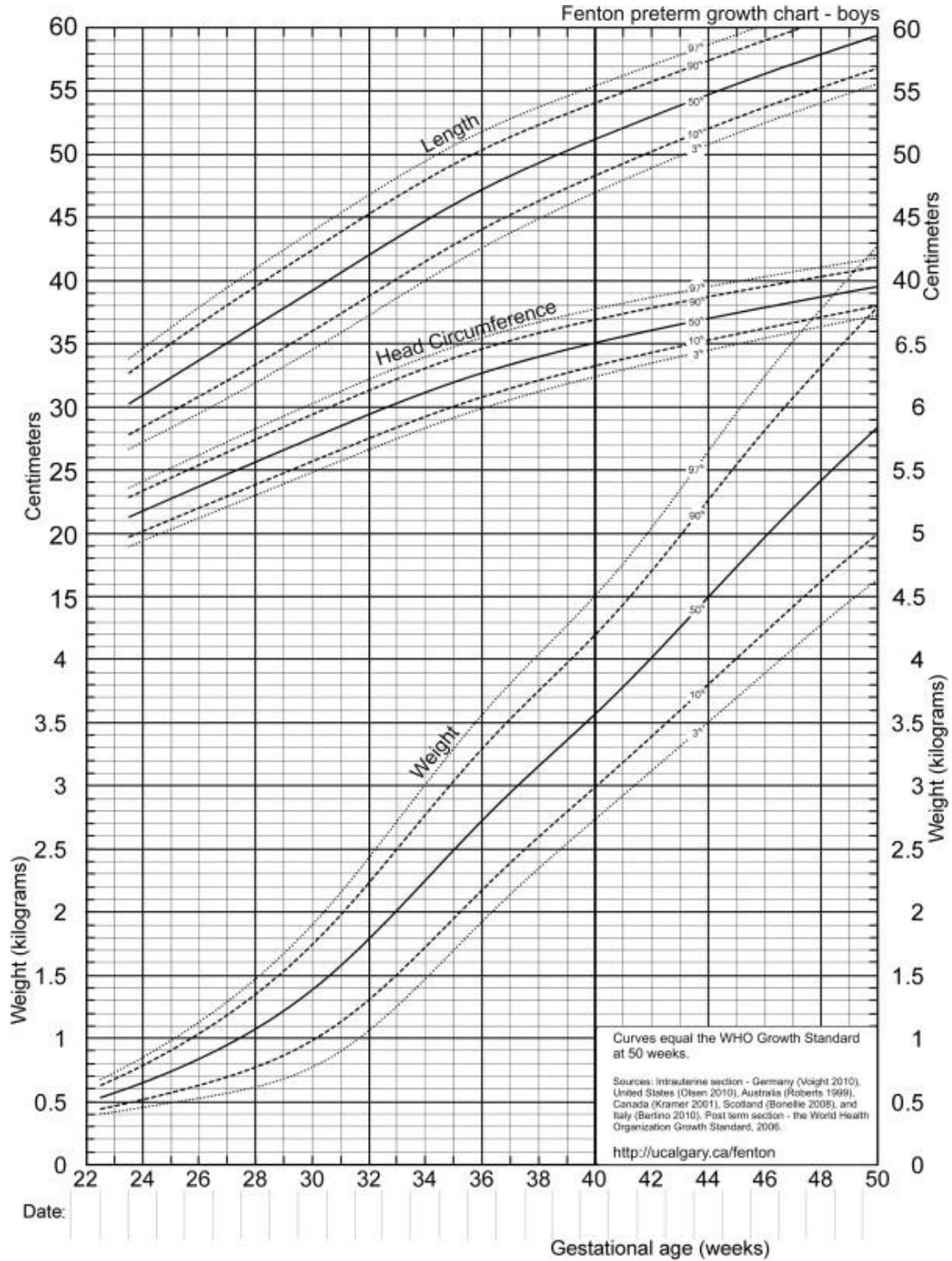
- We identified six different phenotypes of extremely preterm infants who exhibited distinct clinical and laboratorial characteristics.

Keywords Preterm · Extremely low birth weight infants · Critical illness · Phenotypes · Machine learning · Algorithms · Cluster analysis

Anexo C - Curva de crescimento de Fenton 2013 para meninas



Anexo D - Curva de crescimento de Fenton 2013 para meninos



8 REFERÊNCIAS

1. Sankar MJ, Natarajan CK, Das RR, Agarwal R, Chandrasekaran A, Paul VK. When do newborns die? A systematic review of timing of overall and cause-specific neonatal deaths in developing countries. *J Perinatol*. 2016;36 Suppl 1(Suppl 1):S1-S11. doi:10.1038/jp.2016.27.
2. United Nations Inter-agency Group for Child Mortality Estimation (UN IGME), 'Levels & Trends in Child Mortality: Report 2020, Estimates developed by the United Nations Inter-agency Group for Child Mortality Estimation', United Nations Children's Fund, New York, 2020.
3. Hug L, Alexander M, You D, Alkema L; UN Inter-agency Group for Child Mortality Estimation. National, regional, and global levels and trends in neonatal mortality between 1990 and 2017, with scenario-based projections to 2030: a systematic analysis. *Lancet Glob Health*. 2019 Jun;7(6):e710-e720. doi: 10.1016/S2214-109X(19)30163-9. Erratum in: *Lancet Glob Health*. 2019 Sep;7(9):e1179. PMID: 31097275; PMCID: PMC6527519.
4. Matsushita FY, Krebs VLJ, Carvalho WB De Neonatal Hypotension: What Is the Efficacy of Each Anti-Hypotensive Intervention? A Systematic Review. doi: 10.1007/s40746-019-00175-2.
5. Haas SA, Lange T, Saugel B, Petzoldt M, Fuhrmann V, Metschke M, Kluge S (2016) Severe hyperlactatemia, lactate clearance and mortality in unselected critically ill patients. *Intensive Care Med* 42:202–210.
6. Saugel B, Trepte CJ, Heckel K, Wagner JY, Reuter DA (2015) Hemodynamic Management of Septic Shock: Is It Time for "individualized Goal-Directed Hemodynamic Therapy" and for Specifically Targeting the Microcirculation? *Shock* 43:522–529.

7. Hernandez G, Bellomo R, Bakker J. The ten pitfalls of lactate clearance in sepsis. *Intensive Care Med.* 2019 Jan;45(1):82-85. doi: 10.1007/s00134-018-5213-x. Epub 2018 May 12. PMID: 29754310; PMCID: PMC6334727.
8. Andersen LW, Mackenhauer J, Roberts JC, Berg KM, Cocchi MN, Donnino MW (2013) Etiology and therapeutic approach to elevated lactate levels. *Mayo Clin Proc* 88:1127–1140.
9. Levy MM, Evans LE, Rhodes A. The Surviving Sepsis Campaign Bundle: 2018 update. *Intensive Care Med.* 2018 Jun;44(6):925-928. doi: 10.1007/s00134-018-5085-0. Epub 2018 Apr 19. PMID: 29675566.
10. Davis AL, Carcillo JA, Aneja RK, et al (2017) The American College of Critical Care Medicine Clinical Practice Parameters for Hemodynamic Support of Pediatric and Neonatal Septic Shock : Executive Summary. doi: 10.1097/PCC.0000000000001259.
11. Matsushita FY, Krebs VLJ, De Carvalho WB. Associação entre lactato sérico e morbimortalidade em recém-nascidos: uma revisão sistemática e meta-análise. 25º Congresso Brasileiro de Perinatologia, 1 a 4 de dezembro de 2021, São Paulo, Brasil, apresentação de poster.
12. Subudhi S, Verma A, Patel AB, Hardin CC, Khandekar MJ, Lee H, McEvoy D, Stylianopoulos T, Munn LL, Dutta S, Jain RK. Comparing machine learning algorithms for predicting ICU admission and mortality in COVID-19. *NPJ Digit Med.* 2021 May 21;4(1):87. doi: 10.1038/s41746-021-00456-x. PMID: 34021235; PMCID: PMC8140139.
13. Parry G, Tucker J, Tarnow-Mordi W; UK Neonatal Staffing Study Collaborative Group. CRIB II: an update of the clinical risk index for babies score. *Lancet.* 2003 May 24;361(9371):1789-91. doi: 10.1016/S0140-6736(03)13397-1. PMID: 12781540.

14. Fenton TR, Kim JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr* 13:59.
15. Biau G, Scornet E. A random forest guided tour. *Test*. 2016;25(2):197–227.
16. LaValley MP. Logistic regression. *Circulation*. 2008;117(18):2395–9.
17. Noble WS. What is a support vector machine? *Nat Biotechnol*. 2006;24(12):1565–7.
18. Friedman JH. Greedy function approximation: A gradient boosting machine. *Ann Stat*. 2001;29(5):1189–232.
19. Mucherino A, Papajorgji P., Pardalos P. k-Nearest Neighbor Classification. *Data Min Agric*. 2009;34:83–106.
20. Lee J, Cai J, Li F, Vesoulis ZA. Predicting mortality risk for preterm infants using random forest. *Sci Rep*. 2021;11(1):1–9.
21. Feng J, Lee J, Vesoulis ZA, Li F. Predicting mortality risk for preterm infants using deep learning models with time-series vital sign data. *npj Digit Med*. 2021;4(1):1–8.
22. Harsha SS, Archana BR. SNAPPE-II (score for neonatal acute physiology with perinatal extension-II) in predicting mortality and morbidity in NICU. *J Clin Diagnostic Res*. 2015;9(10):SC10–2.
23. Lee S, Aziz K, Dunn M, Clarke M, Kovacs L, Ojah C, et al. Transport risk index of physiologic stability, version II (TRIPS-II): A simple and practical neonatal illness severity score. *Am J Perinatol*. 2013;30(5):395–400.

24. Medvedev MM, Brotherton H, Gai A, Tann C, Gale C, Waiswa P, Elbourne D, Lawn JE, Allen E. Development and validation of a simplified score to predict neonatal mortality risk among neonates weighing 2000 g or less (NMR-2000): an analysis using data from the UK and The Gambia. *Lancet Child Adolesc Health*. 2020 Apr;4(4):299-311. doi: 10.1016/S2352-4642(20)30021-3. Epub 2020 Feb 28. PMID: 32119841; PMCID: PMC7083247.
25. Podda M, Bacciu D, Micheli A, Bellù R, Placidi G, Gagliardi L. A machine learning approach to estimating preterm infants survival: development of the Preterm Infants Survival Assessment (PISA) predictor. *Sci Rep*. 2018;8(1):1–9.
26. Jaskari J, Myllarinen J, Leskinen M, Rad AB, Hollmen J, Andersson S, et al. Machine Learning Methods for Neonatal Mortality and Morbidity Classification. *IEEE Access*. 2020;8:123347–58.
27. Paliwoda M, New K, Davies M, Bogossian F. Physiological vital sign ranges in newborns from 34 weeks gestation: A systematic review. *Int J Nurs Stud*. 2018;77(May 2017):81–90.
28. Manja V, Lakshminrusimha S, Cook DJ. Oxygen saturation targetrange for extremely preterm infants: A systematic review and meta-analysis. *JAMA Pediatr*. 2015;169(4):332–40.
29. von Auenmueller KI, Christ M, Sasko BM, Trappe H-J. The Value of Arterial Blood Gas Parameters for Prediction of Mortality in Survivors of Out-of-hospital Cardiac Arrest. *J Emerg Trauma Shock*. 2017;10(3):134–9.

