



PREDIÇÃO DE NASCIMENTO PRÉ-TERMO EXTREMO EM DUAS CIDADES BRASILEIRAS

**SÃO LUÍS, MA
MAIO – 2026**

EMMANUELLE NOVAES DE VASCONCELOS BRITO

**PREDIÇÃO DE NASCIMENTO PRÉ-TERMO EXTREMO EM DUAS CIDADES
BRASILEIRAS**

Dissertação apresentada ao Programa de Pós-Graduação em Saúde Coletiva da Universidade Federal do Maranhão como requisito parcial à obtenção do título de Mestre em Saúde Coletiva.

Orientadora: Prof^ª Dr^a Erika Bárbara Abreu F. Thomaz

Coorientador: Prof. Dr. Bruno Feres de Souza

SÃO LUÍS, MA

MAIO – 2026

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Novaes de Vasconcelos Brito, Emmanuelle.

PREDIÇÃO DE NASCIMENTO PRÉ-TERMO EXTREMO EM DUAS
CIDADES BRASILEIRAS / Emmanuelle Novaes de Vasconcelos
Brito. - 2026.
183 f.

Coorientador(a) 1: Bruno Feres de Souza.

Orientador(a): Erika Bárbara Abreu Fonseca Thomaz.

Dissertação (Mestrado) - Programa de Pós-graduação em
Saúde Coletiva/ccbs, Universidade Federal do Maranhão, São
Luís, 2026.

1. Saúde Materno-infantil. 2. Coorte de Nascimento.
3. Recém-nascido Prematuro. 4. Prematuridade Extrema. 5.
Aprendizado de Máquina. I. Bárbara Abreu Fonseca Thomaz,
Erika. II. Feres de Souza, Bruno. III. Título.

PREDIÇÃO DE NASCIMENTO PRÉ-TERMO EXTREMO EM DUAS CIDADES BRASILEIRAS

Emmanuelle Novaes de Vasconcelos Brito

Dissertação aprovada em 26 de maio de 2026 pela banca examinadora constituída dos seguintes membros:

Banca Examinadora:

Prof^ª Dr^ª Erika Bárbara Abreu F. Thomaz
Orientadora
Universidade Federal do Maranhão (UFMA)

Prof. Dr. Bruno Feres de Souza
Coorientador
Universidade Federal do Maranhão (UFMA)

Prof^ª Dr^ª Maria da Conceição Pereira Saraiva
Examinador Externo Titular
Universidade de São Paulo (USP)

Prof^ª Dr^ª Alcione Miranda dos Santos
Examinador Interno Titular
Universidade Federal do Maranhão (UFMA)

“A persistência é o menor caminho do êxito.”

Charles Chaplin

AGRADECIMENTOS

A Deus, pela presença constante em minha vida, pela sabedoria concedida e pela força necessária para superar os desafios que surgiram ao longo desta trajetória acadêmica.

À minha família, especialmente ao meu esposo Ronilson, pelo apoio incondicional, compreensão e incentivo contínuo. Às minhas filhas Marcela e Isabela, pela paciência e por compreenderem minhas ausências durante o período de dedicação à esta pesquisa, que também é uma conquista deles.

Aos meus pais, pelo amor, pelos valores transmitidos e pelo exemplo de esforço, ética e perseverança. Agradeço por sempre acreditarem em meu potencial e por me motivarem a buscar a realização dos meus objetivos.

À minha orientadora, professora Dr^a Erika Bárbara Abreu F. Thomaz, exemplo de profissionalismo, ética e sensibilidade humana, manifesto minha sincera gratidão pela orientação competente, pela disponibilidade constante, pelas valiosas contribuições e pelo apoio em todas as etapas desta pesquisa. Seu comprometimento, paciência e incentivo foram fundamentais para a consolidação desta dissertação e para o meu crescimento acadêmico e pessoal.

Ao meu coorientador, professor Dr^o Bruno Feres de Souza, expressei meu reconhecimento e apreço pela dedicação, pelas contribuições relevantes, pela atenção e pelo acompanhamento cuidadoso durante o desenvolvimento desta dissertação. Sua experiência e apoio técnico-científico foram essenciais para o aprimoramento desta pesquisa.

À Universidade Federal do Maranhão (UFMA) e ao Programa de Pós-Graduação em Saúde Coletiva (PPGSC), pela oportunidade de formação e pelo ambiente propício à produção científica. Aos docentes, pela dedicação e pelos conhecimentos compartilhados; à coordenação, à secretaria em especial Malany e aos demais colaboradores, pelo suporte e atenção ao longo do curso; e à Universidade de São Paulo (USP) pela co-gestão das coortes BRISA, fonte de dados deste trabalho.

Aos coordenadores das coortes BRISA, pelo fornecimento do banco de dados e pela colaboração que possibilitou a realização desta pesquisa. Agradeço a todos os profissionais e participantes envolvidos pelo compromisso e pela contribuição ao desenvolvimento científico.

Às instituições financiadoras e de apoio a esta pesquisa: Departamento de Ciência e Tecnologia (DECIT) do Ministério da Saúde; Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (FAPEMA); Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP); e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

Aos colegas e amigos que estiveram presentes nesta jornada, pela parceria, pelas trocas de experiência e pelo apoio mútuo que tornaram o percurso mais leve e enriquecedor.

A todos que, de forma direta ou indireta, contribuíram para a concretização desta dissertação, registro meus sinceros agradecimentos.

LISTA DE FIGURAS

Dissertação

Figura 1. Exemplo de nomenclatura dos diferentes grupos de aprendizado de máquina.....	29
Figura 2. Exemplo das etapas do desenvolvimento de modelos supervisionados de algoritmos de aprendizado de máquina	37
Figura 3. Fluxograma das coortes de pré-natal e nascimento, São Luís – MA, 2010.....	40
Figura 4. Fluxograma das coortes de pré-natal e nascimento, Ribeirão Preto – SP, 2010.....	41
Figura 5. Exemplo de Técnica de validação cruzada (5-folds cross-validation).....	47
Figura 6. Exemplo de uma matriz de confusão para um problema de classificação	49
Figura 7. Exemplo de Curva ROC	50

Artigo

Figura 1. Fluxograma da amostra das coortes BRISA São Luís, MA e Ribeirão Preto, SP (2010).	63
Figura 2. Curva ROC dos modelos de aprendizado de máquina utilizados na predição do NPT extremo nas Coortes Brisa de 2010, São Luís/MA e Ribeirão Preto/SP, com “ampt” (esquerda) e sem “ampt” (direita).....	77
Figura 3. Importância dos preditores de nascimento pré-termo extremo com base nos valores de Shapley, com variável ameaça de parto prematuro na gestação atual (ampt) SL-RP (knn); SL(svmRadial); RP (ranger).....	79
Figura 4. Importância dos preditores de nascimento pré-termo extremo com base nos valores de Shapley, excluindo a variável “ampt” SL-RP (nnet); SL (gam); RP (Rborist).	81

LISTA DE TABELAS

Tabela 1. Exemplo da distribuição das variáveis preditoras e desfecho de pré-natal e nascimento da Coorte BRISA de 2010, São Luís/MA e Ribeirão Preto/SP.	70
Tabela 2. Comparação do desempenho dos modelos de aprendizado de máquina utilizados na predição do NPT extremo nas Coortes Brisa de 2010, São Luís/MA e Ribeirão Preto/SP, incluindo a variável ampt(ameaça de parto prematuro na gestação atual).....	72
Tabela 3. Comparação do desempenho dos modelos de aprendizado de máquina utilizados na predição do NPT extremo nas Coortes Brisa de 2010, São Luís/MA e Ribeirão Preto/SP, excluindo a variável ampt.....	75

LISTA DE QUADROS

Dissertação

Quadro 1. Métricas de desempenho utilizadas em aprendizado de máquina de acordo com o tipo de problema	31
Quadro 2. Matriz de estudos, relacionando autores e seus principais achados com uso de AM em NPT e NPT extremo	32
Quadro 3. Exemplos da configuração das variáveis do estudo	177

Artigo

Quadro 1. Exemplo da configuração das variáveis do estudo	65
--	----

LISTA DE ABREVIATURAS E SIGLAS

NPT extremo — Nascimento pré-termo extremo

IA — Inteligência artificial

AM — Aprendizado de Máquina

BRISA — Estudo de Coortes Brasileiras de Pré-natal e Nascimento de Ribeirão Preto e São Luís

DUM — Data da última menstruação

IG — Idade gestacional

USG — Ultrassom gestacional

MS — Ministério da Saúde

OMS — Organização Mundial da Saúde

PN — Pré-Natal

PT — Pré-Termo

RN — Recém-Nascidos

RP — Ribeirão Preto

SL — São Luís

AUC — *Area Under Curve*

ROC — *Receiver Operating Characteristic*

SVM Radial — *Support Vector Machine Radial*

XGBoost — *Extreme Gradient Boosting*

GLM — *Regressão Logística*

nnet — *Rede neural artificial*

rf — *Randon Forest*

VPP — Valor preditivo positivo

VPN — Valor preditivo negativo

BRITO, Emmanuelle Novaes de Vasconcelos, **PREDIÇÃO DE NASCIMENTO PRÉ-TERMO EXTREMO EM DUAS CIDADES BRASILEIRAS**, 2026, Dissertação (Mestrado em Saúde Coletiva) - Programa de Pós-graduação em Saúde Coletiva, Universidade Federal do Maranhão, São Luís, 183p.

RESUMO

Introdução: O nascimento pré-termo (NPT) extremo é um importante problema de saúde pública, associado a elevada morbimortalidade neonatal e infantil. Sua ocorrência resulta de múltiplos fatores: clínicos, sociodemográficos e comportamentais, o que dificulta sua prevenção. Apesar de sua baixa frequência, seus impactos são significativos para a saúde materno-infantil. Diante desse contexto, modelos de aprendizado de máquina podem prever gestantes com maior risco de NPT extremo. **Objetivo:** Determinar modelos preditivos baseados em técnicas de aprendizado de máquina (AM), para identificar preditores de NPT extremo, em nascidos vivos de duas cidades brasileiras. **Métodos:** Trata-se de um estudo analítico, de natureza transversal, aninhado às coortes BRISA de pré-natal e nascimento de 2010 em São Luís (SL), Maranhão, e Ribeirão Preto (RP), São Paulo, Brasil, cuja escolha fundamenta-se na participação de ambas na pesquisa das coortes e em suas distintas características sociodemográficas e econômicas, permitindo a análise de diferentes contextos maternos-infantis brasileiros. O conjunto de dados analisados foi proveniente de mulheres com gestação única e seus bebês, atendidos em serviços públicos e privados. A amostra foi composta por 6.414 mulheres/crianças em SL e 8.273 em RP, totalizando 14.687 observações. Realizou-se análise de sensibilidade com a variável ameaça de parto prematuro na gestação atual, por representar alterações fisiopatológicas em curso, possivelmente relacionadas ao início do desfecho, sendo um importante preditor. Para avaliar sua influência no desempenho dos modelos, foram construídos seis cenários analíticos, correspondentes às coortes combinada, de São Luís e de Ribeirão Preto, com e sem a inclusão dessa variável, permitindo verificar seu impacto na capacidade preditiva e na estabilidade dos modelos. Selecionamos, com base em especialistas e na literatura, 119 variáveis comuns a SL-RP, 121 variáveis em SL e 119 em RP, relacionadas a fatores socioeconômicos, demográficos, comportamentais, reprodutivos e atenção à saúde da mulher, como potenciais preditoras ao NPT extremo. Em seguida, utilizando-se o algoritmo BORUTA, elegeu-se 50 variáveis e elencamos oito, com maior número de indicações e maior importância na predição. Com base em estudos anteriores, aplicaram-se 14 algoritmos de aprendizado de máquina em cada cenário para construção dos modelos: regressão logística penalizada (*plr*), *Support Vector Machines* (*svmLinear* e

svmRadial), *Random Forest* (*rf*, *ranger* e *Rborist*), *boosting* (*glmboost* e *AdaBoost.M1*), métodos baseados em vizinhança (*Knn* e *KKnn*), redes neurais artificiais (*nnet*), árvores e modelos baseados em partição (*cforest* e *ctree2*) e modelos aditivos generalizados (*gam*). O desempenho foi avaliado por métricas como acurácia, sensibilidade, especificidade, precisão, F1-score e área sob a curva ROC (AUC-ROC). **Resultados:** Os cenários resultaram em desbalanceamento entre NPT extremo e não NPT extremo. Nos seis cenários (SL-RP, SL e RP, três incluindo a variável ameaça de parto prematuro na gestação atual e três excluindo a variável ameaça de parto prematuro na gestação atual), o desempenho variou entre os modelos: No cenário incluindo a variável, em SL-RP o *knn* apresentou sensibilidade de 0.774, AUC-ROC de 0.806 e F1-score de 0.049. No cenário SL, o modelo *svmRadial* obteve sensibilidade de 0.653, AUC-ROC de 0.732 e F1-score de 0.063. Já no cenário RP, o *ranger* apresentou os melhores resultados, com AUC-ROC de 0.86, sensibilidade de 0.771 e F1-score de 0.053. Excluído a variável, em SL-RP o *nnet* apresentou sensibilidade de 0.561, AUC-ROC de 0.648 e F1-score de 0.019. No cenário SL, o modelo *gam* obteve sensibilidade de 0.623, AUC-ROC de 0.587 e F1-score de 0.024. Já no cenário RP, o *Rborist* apresentou os melhores resultados, com AUC-ROC de 0.629, sensibilidade de 0.573 e F1-score de 0.016. Em geral, o *knn* ao incluir a variável e o *nnet* ao excluir a variável, se destacaram pelos maiores valores de AUC-ROC e sensibilidade, apresentaram melhores desempenho para prever o NPT extremo. As oito variáveis preditoras foram: Ameaça de parto prematuro na gestação atual; ameaça de aborto na gestação atual; ocupação da mãe; renda per capita; álcool na gestação; cor da mãe autodeclarada; gravidez planejada e diagnóstico de hipertensão antes ou durante a gestação. **Conclusão:** É possível prever, com boa performance, a ocorrência de NPT extremo, utilizando variáveis de saúde materna e características do feto, socioeconômicas, comportamentais e de acesso ao pré-natal. A identificação precoce de bebês em risco pode impactar positivamente a qualidade de vida da mãe e do recém-nascido. As descobertas deste estudo podem estimular investigações semelhantes, testando diferentes técnicas, populações e contextos clínicos, para o enfrentamento de um dos principais desafios de saúde neonatal.

Palavras-chave: Saúde materno-infantil. Coorte de nascimento. Recém-nascido prematuro. Prematuridade extrema. Aprendizado de Máquina.

ABSTRACT

Introduction: Extreme preterm birth (EPTB) is a major public health problem associated with high neonatal and infant morbidity and mortality. Its occurrence is influenced by multiple clinical, sociodemographic, and behavioral factors, making prevention particularly challenging. Although relatively uncommon, EPTB has substantial consequences for maternal and child health. In this context, machine learning (ML) techniques offer a promising approach for identifying pregnant women at increased risk of EPTB. **Objective:** To develop and evaluate machine learning-based predictive models to identify predictors of extreme preterm birth among live births from two Brazilian cities. **Methods:** This analytical cross-sectional study was nested within the 2010 BRISA prenatal and birth cohorts conducted in São Luís, Maranhão, and Ribeirão Preto, São Paulo, Brazil. The inclusion of these cohorts allowed the investigation of maternal and child health in distinct socioeconomic and demographic settings. The study included singleton pregnancies receiving care in public and private healthcare services, comprising 6,414 mother–child pairs from São Luís and 8,273 from Ribeirão Preto, totaling 14,687 observations. A sensitivity analysis was performed using the variable threatened preterm labor during the current pregnancy, which represents ongoing pathophysiological changes potentially related to the onset of the outcome and is considered an important predictor. Six analytical scenarios were developed, corresponding to the combined cohort and the São Luís and Ribeirão Preto cohorts, each analyzed with and without this variable. Based on expert knowledge and previous literature, 119 variables common to both cohorts, 121 variables in São Luís, and 119 variables in Ribeirão Preto were selected as potential predictors, including socioeconomic, demographic, behavioral, reproductive, and maternal healthcare characteristics. The BORUTA feature selection algorithm identified 50 relevant variables, from which the eight most important predictors were retained. Fourteen machine learning algorithms were evaluated in each scenario: penalized logistic regression (PLR), support vector machines (svmLinear and svmRadial), random forest (rf, ranger, and Rborist), boosting methods (glmboost and AdaBoost.M1), nearest neighbor methods (*Knn* e *KKnn*), artificial neural networks (nnet), tree-based partitioning models (cforest and ctree2) and generalized additive

models (GAM). Model performance was assessed using accuracy, sensitivity, specificity, precision, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC).

Results: All analytical scenarios showed class imbalance between EPTB and non-EPTB births. Model performance varied across the six scenarios. When threatened preterm labor was included, kNN showed the best overall performance in the combined cohort (sensitivity = 0.774, AUC-ROC = 0.806, F1-score = 0.049), svmRadial performed best in São Luís (sensitivity = 0.653, AUC-ROC = 0.732, F1-score = 0.063), and ranger achieved the highest performance in Ribeirão Preto (AUC-ROC = 0.860, sensitivity = 0.771, F1-score = 0.053). When this variable was excluded, nnet performed best in the combined cohort (sensitivity = 0.561, AUC-ROC = 0.648, F1-score = 0.019), GAM in São Luís (sensitivity = 0.623, AUC-ROC = 0.587, F1-score = 0.024), and Rborist in Ribeirão Preto (AUC-ROC = 0.629, sensitivity = 0.573, F1-score = 0.016). Overall, kNN and nnet demonstrated the best predictive performance in scenarios with and without the threatened preterm labor variable, respectively. The eight most important predictors were threatened preterm labor during the current pregnancy, threatened miscarriage during the current pregnancy, maternal occupation, per capita income, alcohol consumption during pregnancy, self-reported maternal race/skin color, planned pregnancy, and hypertension diagnosed before or during pregnancy. **Conclusion:** Machine learning models can predict extreme preterm birth with good predictive performance using maternal health, fetal, socioeconomic, behavioral, and prenatal care variables. Early identification of pregnancies at increased risk may support timely interventions and improve maternal and neonatal outcomes. These findings provide evidence for the potential application of machine learning in prenatal risk stratification and encourage further validation in different populations and clinical settings.

Keywords: Maternal and child health. Birth cohort. Preterm newborn. Extreme preterm infant. Machine learning.

SUMÁRIO

1	INTRODUÇÃO.....	14
2	JUSTIFICATIVA.....	16
3	OBJETO DE ESTUDO, HIPÓTESES E OBJETIVOS	17
	3.1. Objeto	17
	3.2. Hipóteses	17
	3.3. Objetivo geral	18
	3.4. Objetivos específicos	18
4	REFERENCIAL TEÓRICO.....	19
	4.1. Definição e Prevalência de Nascimento Pré-Termo (NPT) extremo	19
	4.2. Fatores associados ao NPT extremo	21
	4.3. Preditores de NPT Extremo	26
	4.4. Aprendizado de Máquina e Predição de NPT extremo	27
5	MÉTODOS.....	38
	5.1. Tipo de estudo	38
	5.2. Local do estudo.....	38
	5.3 Fonte de dado.....	39
	5.4 Critérios de inclusão e não inclusão	42
	5.5 Coleta de Dados	42
	5.6 Variáveis do estudo.....	42
	5.7. Análise dos dados e modelos preditivos	44
6	RESULTADOS	53
7	CONSIDERAÇÕES FINAIS	95
	REFERÊNCIAS	97
	ANEXOS.....	110
	APÊNDICE	177

1. INTRODUÇÃO

O nascimento pré-termo (NPT) é um importante problema de saúde pública além de ser um assunto de extrema importância em todo o mundo. Trata-se de uma das principais causas de morte neonatal e ocupa a segunda colocação entre os óbitos de crianças menores de 5 anos, especialmente no período neonatal (Adam *et al.*, 2015; WHO, 2023). O recém-nascido pré-termo (com idade anterior às 37 semanas de gestação) é classificado em diferentes graus de acordo com a idade gestacional da mulher, onde antes de completadas as 28 semanas de gestação, define-se recém-nascido como pré-termo extremo, pré-termo grave a moderado entre 28 a menos de 34 semanas e o tardio entre 34 a menos de 37 semanas gestacionais (Rocha, 2019; Martinelli *et al.*, 2021). Em 2018, de acordo com a Organização Mundial da Saúde (OMS), aproximadamente 15 milhões de bebês nasceram pré-termo, ou seja, com uma média de 1 para cada 10 bebês. Contudo, cerca de 75% dos casos poderiam ter sido evitados, desde que fossem prestados os cuidados adequados durante o pré-natal e o parto (WHO, 2023).

O NPT possui etiologia multifatorial e inter-relacionada, complexa e por vezes desconhecida (Brasil, 2006; Oliveira *et al.*, 2016; WHO, 2023). A literatura nacional e internacional abordam como potenciais fatores associados ao NPT, a raça/cor negra, gravidez na adolescência, idade avançada, baixo poder socioeconômico, doenças preexistentes, estilo de vida, depressão, intervalo interpartal menor que dois anos, complicações na gestação atual, baixa adesão a assistência pré-natal e tipo parto (Baer *et al.*, 2018; Ouattara *et al.*, 2015; Weber *et al.*, 2018; Adhikari *et al.*, 2019; Carter *et al.*, 2020; Dias *et al.*, 2022).

Poucos autores têm focado na etiologia do NPT extremo (Gao *et al.*, 2019). Sabe-se que os fatores relacionados ao NPT extremo envolvem desde as condições socioeconômicas, comportamentais, etiológicas, demográficas, psicossociais e de acesso às ações e serviços de saúde, além das causas não identificadas (WHO, 2012; WHO, 2023). Em pesquisa realizada por Pereira (2016), identificou-se que o NPT extremo está associado como uma das principais causas de morbimortalidade neonatal em países de baixa e média renda, ocasionando um elevado número de déficits neurológicos e óbitos infantis. Nesse sentido, o risco de mortalidade em crianças que nascem prematuras extrema é maior devido ao incompleto desenvolvimento fetal e alta suscetibilidade de adquirir infecções (Suff; Story; Shennan, 2019; WHO, 2023).

Diante disso, apesar da baixa taxa do NPT extremo, tornando um evento raro em relação às outras classificações do NPT, a prevenção é um desafio por se tratar de um aspecto que envolve múltiplos fatores, desde o socioeconômico, fisiológicos e educacionais (Bittar;

Fonseca; Zugaib, 2010; Gao et al., 2019). Dessa forma, determinar quais mulheres necessitam de um maior acompanhamento e/ou algum tipo de intervenção é crucial para reduzir as taxas de mortalidade neonatal e os altos custos que o NPT extremo gera. Por esse motivo, auxílios para tomada de decisão e modelos de previsão com inteligência artificial (IA) se tornam relevantes para os gestores e também profissionais de saúde (Bazo, 2020).

Qi Sun *et al.*, 2022, Lacy *et al.*, 2023, Araújo *et al.*, 2023, Ding *et al.*, 2024 realizaram estudos para prever NPT, enquanto Gao *et al.*, 2019 destacaram-se entre os poucos autores que analisaram a predição do NPT extremo como desfecho. Os autores fizeram uso de algoritmos de aprendizado de máquina, dentre eles a Regressão Logística (GLM), *Random Forest* (RF), *Naive Bayes* (NBM), *Support Vector Machines* (SVM), Redes neurais artificiais (NNET), *Extreme Gradient Boosting* (XGBoost). O RF e as NNET obtiveram melhores métricas de desempenho em contexto de base de dados desbalanceada.

Diante da escassez de estudos que investiguem fatores de risco e desenvolvam modelos de aprendizado de máquina voltados à predição do NPT extremo (Gao et al., 2019), evidencia-se a necessidade de prever este evento raro, mas de extrema significância à saúde materna e infantil. Dessa forma, as perguntas que orientam esta dissertação são: É possível prever, com boa performance, a ocorrência de NPT extremo, a partir de dados relativos ao período pré-gestacional e primeiros meses de gestação? Quais algoritmos de aprendizado de máquina têm melhor performance na predição do NPT extremo? Quais são as principais variáveis preditoras do NPT extremo?

Diante desse cenário, esta dissertação propõe o desenvolvimento de modelos preditivos baseados em técnicas de aprendizado de máquina para identificar gestantes com risco de NPT extremo, utilizando dados das coortes de pré-natal e nascimento BRISA 2010 de São Luís e Ribeirão Preto. Espera-se que os resultados subsidiem a formulação e a implementação de políticas e estratégias voltadas à redução da prematuridade extrema em contextos de vulnerabilidade, além de fornecer evidências aplicáveis a análises em nível nacional e em cenários geográficos comparáveis.

2. JUSTIFICATIVA

Tendo em vista a sobrevida, os bebês prematuros extremo são mais suscetíveis às consequências negativas, quando comparados aos bebês não pré-termo extremo (Brasil, 2020). Entender os mecanismos biológicos que levam ao NPT extremo é fundamental para o desenvolvimento de intervenções mais eficazes. Nesse sentido, a predição pode contribuir para intervenções em momento oportuno, consequentemente a redução de eventos secundários à prematuridade extrema, como problemas respiratórios, distúrbios neurológicos e dificuldades alimentares, portanto, de extrema relevância na obstetrícia e na neonatologia.

Ao identificar o risco de NPT extremo antecipadamente e intervir de forma proativa, é possível aprimorar substancialmente a saúde e evitar os desfechos desfavoráveis para o recém-nascido (Carvalho *et.al*, 2021). Assim, pesquisas que visem à predição de NPT extremo estão de acordo com as preocupações atuais de saúde pública no mundo todo. Diante disso, técnicas de predição com uso de dados de fácil mensuração ao longo da gestação – como informações relacionadas a fatores socioeconômicos, comportamentais e clínicos maternos – amplia a viabilidade de implementação dessa solução na prática obstétrica, permitindo a construção de modelos preditivos capazes de apoiar a tomada de decisão clínica e contribuir para a redução do NPT extremo.

Além disso, entender quais variáveis podem prever o NPT extremo a partir de sinais precoces e realizar a triagem para esses problemas, possibilita intervenções mais eficazes e direcionadas, implementando as ações de promoção à saúde da mulher durante todo o período gestacional, além de prevenir complicações geradas por esse processo (Carvalho *et al.*, 2021). Ressalta-se também que pesquisas desse tipo oferecem uma abordagem tecnológica promissora, utilizando técnicas de aprendizado de máquina para realizar predições relevantes, possibilitando analisar grandes volumes de dados, identificando padrões e correlações. Ao aplicar algoritmos de aprendizado de máquinas em estudos sobre NPT extremo, podemos identificar alterações nos padrões dos fatores biopsicossociais associados a essa problemática, contribuindo para realização de intervenções oportunas e redução dos casos de NPT e das complicações associadas. Essa abordagem pode proporcionar novas compreensões sobre problemas de saúde materno e infantil, oferecendo perspectivas e soluções inovadoras para um desafio crescente e complexo (Bazo, 2020).

Este estudo tem o potencial de gerar conhecimentos valiosos sobre os principais preditores da ocorrência de NPT extremo, especialmente em contextos semelhantes aos de cidades como São Luís, MA e Ribeirão Preto, SP, podendo ser útil para prevenção do problema

quando aplicado para apoiar a tomada de decisão clínica ou a implementação de políticas públicas. Assim, os resultados podem contribuir para a melhoria das práticas clínicas e intervenções comunitárias voltadas para a prevenção desse problema.

Adicionalmente, esta pesquisa integra um projeto de maior abrangência desenvolvido no âmbito da estratégia estadual Plano Maranhão 2050, um planejamento estratégico de longo prazo voltado ao desenvolvimento sustentável e à redução das desigualdades sociais. Entre seus eixos estruturantes, destaca-se o eixo *Sociedade saudável, segura e justa*, que contempla ações direcionadas ao fortalecimento da atenção à saúde, incluindo a redução de mortes evitáveis, a qualificação da assistência materno-infantil e o monitoramento de indicadores de saúde (Maranhão, 2024). Nesse contexto, os resultados deste estudo poderão fornecer evidências científicas para subsidiar o planejamento e a implementação de estratégias voltadas ao aprimoramento da atenção materno-infantil no estado do Maranhão, contribuindo para o alcance das metas estabelecidas pelo Plano Maranhão 2050, sem perder de vista sua potencial aplicabilidade em outros cenários brasileiros.

Dada a grande diversidade da população brasileira, é fundamental realizar estudos populacionais regionalizados que investiguem condições específicas, considerando a inter-relação entre diversos fatores. Nesse sentido, identificar preditores da prematuridade extrema com uso de aprendizado de máquinas, capazes de fazer previsões com alto nível de confiança e de forma rápida é essencial para implementar as políticas públicas de saúde e sociais e para o planejamento mais eficaz das estratégias de cuidado materno-infantil.

3. OBJETO DE ESTUDO, HIPÓTESES E OBJETIVOS

3.1. Objeto

Predição de NPT extremo em duas cidades brasileiras.

3.2. Hipóteses

- **Hipótese**

A aplicação de técnicas de aprendizado de máquina permite determinar características maternas e fetais que predizem casos de NPT extremo em duas cidades brasileiras, com boa performance.

3.3. Objetivo geral

Determinar modelos baseados em técnicas de aprendizado de máquina para identificar preditores do NPT extremo com dados das coortes BRISA de pré-natal e nascimento de 2010, em São Luís – MA e Ribeirão Preto – SP.

3.4. Objetivos específicos

- Caracterizar a amostra do estudo levando em consideração dados sócio-demográficos, clínicos e obstétricos das coortes BRISA;
- Determinar modelos de aprendizado de máquina para prever o NPT extremo;
- Comparar o desempenho de diferentes algoritmos de aprendizado de máquina na predição do desfecho, por meio de métricas como acurácia, sensibilidade, especificidade, valor preditivo positivo (VPP), valor preditivo negativo (VPN) e área sob a curva *Receiver Operating Characteristic* (ROC-AUC);
- Identificar os preditores mais relevantes para o NPT extremo com base na importância atribuída pelo modelo.

4. REFERENCIAL TEÓRICO

4.1. Definição e Prevalência de Nascimento Pré-Termo (NPT) extremo

A prematuridade é definida como o nascimento ocorrido antes de 37 semanas completas de gestação e constitui um dos principais problemas de saúde pública, em razão de sua elevada contribuição para a morbimortalidade neonatal e infantil. De acordo com a Organização Mundial da Saúde (OMS), os recém-nascidos pré-termo podem ser classificados segundo a idade gestacional em pré-termo extremo (22 a <28 semanas), pré-termo grave (28 a <32 semanas), pré-termo moderado (32 a <36 semanas) e pré-termo tardio (36 a <37 semanas) (WHO, 2023). Além da idade gestacional, os recém-nascidos também podem ser classificados conforme o peso ao nascer, considera-se baixo peso ao nascer quando o peso é inferior a 2.500 gramas (g), muito baixo peso ao nascer quando inferior a 1.500 g e extremo baixo peso ao nascer quando inferior a 1.000 g, independentemente da idade gestacional (Brasil, 2011; 2017).

Os recém-nascidos prematuros extremos apresentam características decorrentes da imaturidade dos órgãos e sistemas, incluindo peso muito reduzido, pele fina e translúcida, escassez de tecido adiposo, desenvolvimento incompleto das cartilagens, reflexos neurológicos imaturos, além de dificuldades na sucção, deglutição e regulação térmica (Marcondes et al., 2020). Em consequência dessas condições, apresentam maior susceptibilidade a complicações respiratórias, neurológicas, infecciosas, gastrointestinais e oftalmológicas, como síndrome do desconforto respiratório, hemorragia peri-intraventricular, sepse neonatal, enterocolite necrosante, retinopatia da prematuridade e óbito neonatal (Sousa et al., 2017).

O prognóstico dos prematuros extremos está diretamente relacionado à idade gestacional e ao peso ao nascer. Embora a sobrevivência seja possível a partir de 23 semanas de gestação, os riscos de sequelas neurológicas e de outras complicações permanecem elevados. Acima de 27 semanas, observa-se melhora significativa dos desfechos clínicos e do desenvolvimento neuropsicomotor (Brasil, 2020). Os Recém-nascidos com idade gestacional entre 23 e 24 semanas e peso inferior a 500 g são considerados pertencentes à denominada “zona cinzenta” da viabilidade neonatal, caracterizada por elevada incerteza quanto à sobrevivência e às morbidades futuras, exigindo decisões assistenciais complexas e multidisciplinares (Monteiro et al., 2022). Por outro lado, para recém-nascidos com mais de 25 semanas de gestação e peso superior a 500 g, o início de cuidados intensivos é amplamente

recomendado, uma vez que a maioria apresenta chances consideráveis de sobrevivência e melhores prognósticos (Aziz et al., 2020; Bouzada et al., 2018).

A sobrevivência dos prematuros extremos depende não apenas das condições biológicas do recém-nascido, mas também da estrutura dos serviços de saúde, disponibilidade tecnológica, organização da assistência neonatal e práticas clínicas adequadas (Alleman et al., 2013; Aziz et al., 2020; Barfield, 2018; Barros et al., 2018). Além disso, os óbitos neonatais precoces estão fortemente associados a fatores perinatais, socioeconômicos e assistenciais, muitos dos quais são potencialmente evitáveis por meio de ações qualificadas dos serviços de saúde (Almeida, Couto e Trapani Júnior, 2019).

Apesar de representarem apenas 1% a 2% dos nascidos vivos, os prematuros extremos são responsáveis por aproximadamente 25% a 30% da mortalidade infantil e por 50% a 70% da mortalidade neonatal (Aredes, Santos e Fonseca, 2017; WHO, 2023). Entre os sobreviventes, observa-se maior risco de alterações do crescimento, déficits motores, cognitivos e sensoriais, transtornos comportamentais e psicológicos, além de maior predisposição ao desenvolvimento de doenças crônicas na vida adulta, como diabetes mellitus tipo 2 e doenças cardiovasculares (ONG, 2015).

A premissa fundamental da prematuridade é que os riscos neonatais aumentam progressivamente à medida que diminui a idade gestacional. Dessa forma, independentemente da classificação utilizada, quanto menor a idade gestacional ao nascimento, maiores são os riscos de morbidade e mortalidade neonatal (Bortoletto et al., 2020). Nesse contexto, a prematuridade extrema constitui atualmente a principal causa de morte em crianças menores de cinco anos, especialmente no período neonatal (WHO, 2023).

Globalmente, estima-se que ocorram cerca de 13,4 milhões de nascimentos prematuros por ano, correspondendo a aproximadamente 10% dos nascimentos mundiais (WHO, 2023). A incidência varia entre 5% e 10% nos países desenvolvidos, podendo atingir até 18% nos países em desenvolvimento (Liu et al., 2016). Em 2015, aproximadamente 45% dos óbitos em crianças menores de cinco anos ocorreram no período neonatal, sendo as complicações decorrentes da prematuridade a principal causa dessas mortes (Liu et al., 2016).

A maioria dos nascimentos pré-termo ocorre em países de baixa e média renda, especialmente nas regiões do sul da Ásia e da África Subsaariana. Na América Latina e Caribe,

a prevalência estimada é de 8,6%, enquanto no Brasil aproximadamente 11,7% dos partos ocorrem antes de 37 semanas de gestação, totalizando cerca de 300 mil nascimentos prematuros por ano e posicionando o país entre os dez com maior número absoluto de prematuros no mundo (WHO, 2012).

No cenário nacional, um estudo conduzido por Alberton et al. (2023) identificou 3.503.085 nascimentos pré-termo entre 2011 e 2021, correspondendo a uma prevalência de 11,1%, valor superior à média global de 10%. A distribuição foi relativamente homogênea entre as regiões brasileiras, variando de 10,8% a 11,6%. Em relação à prematuridade extrema, estimou-se que ela representou aproximadamente 16,7% do total de nascimentos prematuros registrados no país em 2019 (Aynalem et al., 2021).

Dados do Sistema de Informações sobre Nascidos Vivos (SINASC) demonstram que, em 2023, o Maranhão apresentou taxa de 0,55% de nascimento pré-termo extremo e 11,8% de nascimentos pré-termo. No município de São Luís, essas prevalências foram de 0,86% e 13,8%, respectivamente. No estado de São Paulo, registraram-se 0,66% de nascimentos pré-termo extremos e 11,1% de prematuridade total, enquanto em Ribeirão Preto as prevalências corresponderam a 0,65% e 12,4%, respectivamente (Brasil, 2026).

Corroborando a esses estudo, de acordo com o ministério da saúde, entre 2011 e 2019, as taxas de prematuridade no Brasil variaram de 6,2 a 9,3 por mil nascidos vivos. No mesmo período, observou-se aumento da mortalidade neonatal e infantil, passando de 10,0 para 14,2 por mil nascidos vivos, evidenciando a relevância da prematuridade como importante determinante dos indicadores de saúde materno-infantil (Brasil, 2019).

4.2. Fatores associados ao NPT extremo

São vários os fatores associados ao nascimento pré-termo extremo. De acordo com a OMS (2023) aspectos como desigualdade de renda, colabora para o aumento da mortalidade devido à falta de cuidados, assistência à saúde, acesso e recursos, agravando ainda mais as taxas de sobrevivência (WHO, 2023).

Segundo a literatura, alguns fatores apresentam maiores associações com a prematuridade, destacando-se: dados sociodemográficos, baixo nível socioeconômico; antecedente de parto pré-termo; idade materna menor que 20 anos ou maior que 40; baixa estatura materna (inferior a 1,52 metros); gestação de gêmeos; presença de sangramentos vaginais no 2º trimestre; o amadurecimento cervical; etilismo, ser mãe solteira; ocupação

materna em atividade profissional remunerada; baixo estado nutricional; alteração de peso ou também o peso inadequado da mãe; a raça/cor; infecções do trato urinário; exposição a substâncias tóxicas; depressão; baixa qualidade ou ausência da assistência de pré-natal; e tipo de parto escolhido (Guimarães *et al.*, 2017; Gao *et al.*, 2019 e Adhikari *et al.*, 2019).

Em relação ao total de consultas pré-natais, conforme o Ministério da Saúde (2013), recomenda-se no mínimo 6 (seis), com acompanhamento intercalado entre médico generalista e enfermeiro, além de consulta odontológica e com outros profissionais que se julguem necessários. Conforme o seguinte cronograma: até 28ª semana – mensalmente; da 28ª até a 36ª semana – quinzenalmente; da 36ª até a 41ª semana – semanalmente (Brasil, 2013).

Alberton *et al.* (2023), em estudo realizado sobre pesquisa da prevalência e tendência temporal da prematuridade no Brasil entre 2011 a 2021, inferem também que fatores ambientais e maternos, como poluição do ar, histórico de complicações gestacionais, infecções bacterianas e/ou virais, estão associados ao maior risco de parto pré-termo. Entre as infecções virais maternas, merece destaque a COVID-19, doença causada pelo coronavírus (SARS-CoV-2), e sua possível influência no nascimento prematuro, pode ocorrer por questões patológicas de deflagração do trabalho de parto antes do termo ou pela interrupção da gestação por conta de complicações maternas da infecção, como a síndrome respiratória.

Destacou-se também que durante a pandemia, o cenário epidemiológico desfavorável às gestantes, evidenciou que áreas nas quais havia maior proporção de gravidez na adolescência, precário acesso aos serviços de saúde e baixos níveis educacionais coincidiram com municípios menos estruturados para responder à emergência da covid-19, carecendo de apoio aos quadros graves provocados pela infecção, podendo apresentar intercorrências maternas e fetais com maior frequência (Alberton *et al.*, 2023).

No tocante às características sociodemográficas maternas, estudos destacam que a faixa etária materna com maiores índices de partos prematuros ocorrem principalmente em mulheres muito jovens (≤ 15 anos) ou em mulheres com idade a partir de 35 anos (Pitilin *et al.*, 2021; Vanin *et al.*, 2020; Carvalho *et al.*, 2018; Freitas; Araújo, 2015; Xu *et al.*, 2014).

Há evidências consistentes de que disparidades étnico raciais aumentam o risco de NPT (Menon *et al.*, 2011). Estudos nos Estados Unidos e no Reino Unido mostram que mulheres negras têm maior probabilidade de parto prematuro em comparação às mulheres brancas (Vianna M., 2006; Jongh BE *et al.*, 2012). No Brasil, mulheres pretas e pardas apresentam menor acesso e início tardio ao pré-natal. Dados indicam que a cor da pele é um fator de risco que influencia a prematuridade, mesmo após controle por renda e escolaridade, sugerindo que

além da desigualdade socioeconômica, a discriminação racial também influencia esses desfechos (Chacham A., 2001; Oliveira KA et al., 2018).

Estudos como o de Guimarães *et al.* (2017), identificaram estatisticamente uma maior prevalência de prematuridade extrema entre as mães com 15 anos ou menos (14,4%) e com 35 anos ou mais de idade (9,3%), quando comparadas às mulheres de 20 a 34 anos (7,6%).

Quanto à escolaridade, evidencia-se que a baixa escolaridade materna é fator associado à prematuridade, o estudo de Ginar (2020), aponta que mulheres com menos de sete anos completos de estudo demonstraram 10% maior prevalência de prematuridade quando comparadas àquelas com maior tempo de escolaridade. Outros estudos também apontaram que a prematuridade foi observada em maior número naqueles bebês cujo as mães possuíam escolaridade até o ensino fundamental (Chermont *et al.*, 2020; Souza *et al.*, 2019).

Em relação à renda, segundo Pitilin *et al.* (2021) e Xu *et al.* (2014) reforçam que a baixa renda é um fator crucial em partos prematuros. Entretanto, Defilipo *et al.* (2022) relatam que quanto maior a renda, maior a probabilidade de ocorrer nascimento pré-termo, associado ao tipo de parto. Porém, em um estudo realizado no município de Piracicaba, observou-se uma tendência de diminuição na porcentagem de prematuridade com o aumento da renda familiar média (Tuon *et al.*, 2016).

Outro estudo, realizado na Nova Zelândia, com o objetivo de avaliar as mudanças na incidência de parto prematuro ao longo de 20 anos, descobriu que as taxas de prematuridade aumentaram em 71,9% na população que vive nas áreas mais ricas e apenas 3,5% nas áreas mais pobres, desafiando o pensamento tradicional sobre a associação de fatores socioeconômicos com o parto prematuro (Craig; Thompson; Mitchell, 2002).

A menor frequência em consultas de pré-natal, infecção do trato urinário durante a gestação, histórico de aborto, cesariana, parto prematuro, gravidez gemelar, multiparidade, primeira gestação são fatores que estão associados ao índice elevado de partos prematuros extremos e aumento de complicações. A assistência pré-natal deficitária, menos de 5 consultas ao completar 28 semanas de gestação, e apresentar intercorrências clínicas gestacionais não diagnosticada, aumentam as chances de prematuridade extrema em cinco e quatro vezes, respectivamente (Carvalho *et al.*, 2021).

Em estudos realizados por Defilipo *et al.* (2020) identificaram também em suas pesquisas outros fatores relacionados à maior predisposição da prematuridade: filho prematuro prévio, primiparidade, vítima de violência contra a mulher gestante, parto cesáreo e maior renda familiar.

A realização de mais de seis consultas pré-natais foi considerada um fator de proteção, com menores chances de nascimento prematuro, porém a violência sofrida pela gestante apresentou uma consistente e forte associação com o NPT (Defilipo *et al.*, 2020).

Uma pesquisa realizada em Divinópolis, Minas Gerais, através da análise do Sistema de Informações sobre Nascidos Vivos, demonstrou que as mães que realizaram seis consultas de pré-natal ou menos, a prematuridade foi três vezes superior (16,0%) em relação às que fizeram sete ou mais consultas (5,3%). No que tange, ao tipo de parto, identificou-se maior frequência de prematuros nos partos cesáreos (8,9%) do que nos partos vaginais (6,5%). Houve, também, maior frequência de prematuridade entre as mães que tiveram quatro ou mais filhos vivos (Guimarães *et al.*, 2017).

Outrossim, é fundamental destacar que a presença de fatores de risco não garante, por si só, que um parto prematuro extremo ocorrerá. Dados do Setor de Baixo Peso Fetal da Clínica Obstétrica do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (HC-FMUSP) mostram que aproximadamente 25% das mulheres que tiveram um parto prematuro extremo em uma gestação anterior (considerado um dos fatores de risco mais significativos) podem enfrentar parto prematuro extremo em uma gravidez futura. No entanto, 75% dessas mulheres conseguem levar a gravidez até o termo (Bittar; Fonseca; Zugaid, 2010; Ahumada-Barrios; Alvarado, 2016; Guimarães *et al.*, 2017; Gao *et al.*, 2019; Adhikari *et al.*, 2019; Almeida; Bartosch; Macedo 2021 e Carvalho *et al.*, 2021).

Em um estudo realizado em um hospital localizado no Peru, foi evidenciado através da análise das variáveis múltiplas, segundo a qual foram estatisticamente significativos para parto prematuro, os seguintes fatores: gravidez atual gemelar, controle pré-natal inadequado (<6 controles), controle pré-natal ausente (nenhum controle), antecedente de parto prematuro e pré-eclâmpsia (Ahumada-Barrios; Alvarado, 2016).

Quanto aos hábitos maternos, o tabagismo, o consumo de bebidas alcoólicas, a dependência por drogas, têm maior chance da gestação resultar em NPT extremo. O tabagismo materno pode causar uma variedade de anormalidades neurológicas e metabólicas devido ao deficiente fornecimento de oxigênio, de nutrientes e estoques de energia para o feto, além disso, drogas como a cocaína, podem causar vasoconstrição uterina, provocando danos placentário e fetal. Esses danos podem ocasionar o descolamento prematuro da placenta e predispor ao nascimento pré-termo extremo, aborto espontâneo e morte fetal (Carvalho *et al.*, 2021).

A principal abordagem para reduzir a incidência de parto prematuro extremo é prevenir os riscos associados, particularmente através de sua identificação e monitoramento, como também ações voltadas à saúde materna e uma assistência à gestação de qualidade, prestada por

profissionais médicos e enfermeiros, contribuem para melhorar a saúde materno-infantil e assim reduzir a mortalidade neonatal. Ademais, identificar fatores associados é fundamental para propor e direcionar medidas de promoção à saúde materna e neonatal adequadas a essa realidade (Gaiva; Modes; Stevanato, 2015).

Nesse sentido, a OMS (2018) identifica que o NPT extremo pode ocorrer por variadas razões e grande parte dos casos quando identificados acontecem espontaneamente. É válido ressaltar que casos comuns podem estar associados a infecções, condições crônicas (hipertensão e diabetes), múltiplas gestações, influência genética e causas indefinidas.

O NPT extremo é uma complicação obstétrica significativa, com implicações graves para a saúde neonatal e, sua alta taxa de morbimortalidade neonatal continua a ser uma preocupação importante e sua prevalência ao longo da última década, destaca a necessidade urgente de estratégias mais eficazes de rastreio e intervenção.

Estudo como o de Almeida; Bartosch; Macedo (2021), identifica que a medição ecográfica transvaginal do comprimento cervical tem sido eficaz como uma ferramenta promissora para identificar grávidas com risco elevado de NPT extremo, sugerindo que esse método pode ser eficaz na detecção precoce de risco, permitindo a implementação de medidas preventivas antes do início do trabalho de parto.

Além disso, o avanço tecnológico na assistência neonatal tem contribuído significativamente para a melhoria das taxas de sobrevivência dos recém-nascidos prematuros de muito baixo peso. A introdução de novas práticas e tecnologias na Unidade de Terapia Intensiva Neonatal (UTIN) tem possibilitado um cuidado mais especializado e eficiente, aumentando assim as chances de sobrevivência desses bebês e melhorando seus desfechos a longo prazo (Mosher, 2017).

Nessa perspectiva, enquanto a medição ecográfica transvaginal do comprimento cervical pode oferecer uma solução para o rastreio do risco de NPT extremo, o avanço contínuo na assistência neonatal também é crucial para garantir o melhor suporte possível aos recém-nascidos prematuros extremo. A combinação de estratégias eficazes de rastreio e cuidados intensivos é essencial para melhorar os resultados para mães e bebês (Carter *et al*, 2020).

Contudo, a prevenção da prematuridade tem sido identificada como um dos maiores desafios obstétricos atuais. Estudos identificam que as ações preventivas se dividem em três níveis: primária (identificação e tratamento dos fatores de risco), secundária (diagnóstico precoce do trabalho de parto prematuro) e terciária (intervenções para reduzir complicações do nascimento prematuro). Destaca-se que, a maior parte dos esforços está na prevenção terciária, com uso de tocolise e corticoides, que diminuem a morbimortalidade, mas não reduzem

significativamente a incidência de partos prematuros. Portanto, avanços na prevenção primária e secundária, seguidos de novas estratégias, são necessários para diminuir a incidência do parto prematuro e prevenir as doenças associadas à prematuridade (Bittar; Fonseca; Zugaid, 2010).

4.3. Preditores de NPT Extremo

A tomada de decisão precoce e oportuna, como intervenções terapêuticas, qualidade na assistência pré-natal, transferência segura e atendimento terciário, são ações que podem ser tomadas de acordo com a predição do NPT extremo (Koullali *et al.*, 2016).

Os atributos de um teste preditivo ou de rastreamento, além de sua capacidade de rastrear com o mínimo de falso-positivos e negativos, estão relacionados com sua funcionalidade, como simplicidade técnica, reprodução adequada de resultados, boa adesão da paciente, inocuidade e baixo custo. No rastreamento do parto prematuro extremo, dentre os vários métodos, os mais estudados e que demonstraram ser mais efetivos e reprodutíveis são a monitoração das contrações uterinas, a avaliação do comprimento do colo uterino e as pesquisas de marcadores bioquímicos específicos (Bittar; Fonseca; Zugaid, 2010; Adhikari *et al.*, 2019; Carter *et al.*, 2020).

Werber *et al.*, 2018, relata que a adição da avaliação de biomarcadores maternos (por exemplo, comprimento cervical, fibronectina fetal e analitos séricos maternos) melhora marginalmente a previsão de NPT extremo e reforça a importância da aplicação de técnicas de aprendizado de máquina que pode revelar algoritmos de triagem mais sensíveis e específicos do que é possível com métodos tradicionais.

Carter *et al.*, 2020, validaram o protótipo do aplicativo QUIPP, estudo baseado em uma coorte prospectiva em 16 hospitais do Reino Unido, ferramenta que identificou a acurácia de 90,4% de risco de NPT extremo em mulheres com menos de 30 semanas de gestação.

Fergus *et al.* (2013), pesquisou em seu estudo, o efeito de utilizar EHG (eletro-histerografia) para medir as contrações uterinas, comparando vários classificadores de aprendizado de máquina com um conjunto de dados aberto, contendo 300 registros (38 prematuros e 262 a termo), usando um filtro de sinal e recursos pré-selecionados, que são adequados para classificar registros de termo e prematuros. Os resultados indicaram 96% de sensibilidade, 90% de especificidade e um valor de área sob a curva de 95% com erro global de 8%.

Nesse contexto, a validação e replicação em outras populações, e incorporação de outras medidas que identificam o risco de NPT extremo, como os biomarcadores maternos, podem ser

um passo para identificar mulheres adicionais que podem se beneficiar de intervenções novas ou atualmente disponíveis.

4.4. Aprendizado de Máquina e Predição de NPT extremo

O aprimoramento tecnológico na área da Inteligência Artificial (IA) tem ocorrido de forma acelerada nos últimos anos, consolidando-se como uma ferramenta de grande impacto social, especialmente no campo da saúde. Definida como qualquer técnica capaz de permitir que máquinas se aproximem da inteligência humana, sendo atualmente um dos pilares da chamada Medicina de Precisão, a qual se fundamenta na interação entre fatores genéticos, ambientais e estilo de vida (Araújo Filho et al., 2023). Nesse cenário, a IA tem como principal objetivo aumentar a eficácia dos diagnósticos e tratamentos, destacando-se como uma estratégia promissora para a análise de bases de dados complexas e com grande volume de informações populacionais, conhecidas como *big data* (Chiavegatto Filho et al., 2023).

Diante disso, o aprendizado de máquina (*machine learning* – AM) configura-se como um importante subconjunto da IA, sendo definido como a ciência dedicada ao desenvolvimento de algoritmos e modelos estatísticos que permitem aos sistemas computacionais realizar tarefas sem instruções explícitas, baseando-se em padrões e inferências extraídos dos dados (Amazon, 2022). O principal diferencial do AM reside na sua capacidade de lidar com conjuntos de dados complexos, volumosos e heterogêneos, combinando múltiplos preditores de forma interativa e não linear, frequentemente apresentando desempenho superior aos métodos estatísticos tradicionais (Batista; Chiavegatto Filho, 2019).

Dentre as estratégias para a construção de sistemas de AM, destaca-se a predição definida como a tarefa cujo objetivo é aprender, a partir de dados disponíveis, um modelo capaz de estimar ou prever o valor associado a uma nova observação. Esse modelo estabelece um mapeamento entre variáveis de entrada e uma variável de saída, permitindo inferir resultados para dados ainda não observados (Faceli et al., 2021).

Portanto o uso de modelos preditivos avançados, baseados em Aprendizado de máquina, estão revolucionando a maneira como fazemos previsões em vários setores, incluindo a medicina, onde está ganhando destaque na predição de nascimento pré-termo extremo. Essa ferramenta permite a análise complexa de dados, a personalização do cuidado, o monitoramento

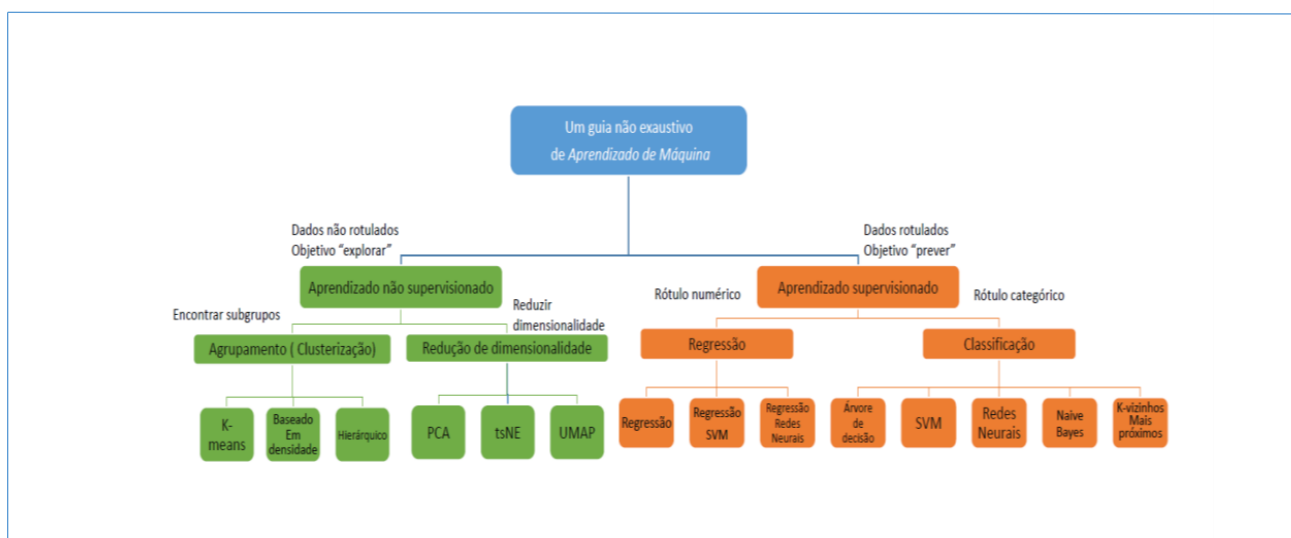
contínuo, a detecção dos fatores de risco e o aprimoramento de protocolos clínicos (SHARIFI-HERIS et al., 2022).

4.4.1 Algoritmos de aprendizado de máquina, métricas de desempenho e predição de NPT extremo

De acordo com Rokach e Maimon (2015), os algoritmos de aprendizado de máquina podem ser classificados em dois grandes grupos: supervisionados e não supervisionados. A distinção entre eles está na necessidade de rotulação prévia dos dados, exigida nos métodos supervisionados, mas dispensável em não supervisionados. Embora a ausência de rotulação torne os métodos não supervisionados mais práticos, ela também implica menor controle e adaptabilidade do modelo, tornando essencial a escolha do algoritmo com base nas características do conjunto de dados e nos objetivos do projeto.

Ressalta-se que os algoritmos não supervisionados são comumente utilizados para fins de clusterização, ou seja, permitindo o agrupamento de pacientes com base em características semelhantes e redução de dimensionalidade. Já os algoritmos supervisionados podem ser segmentados em regressão quando o rótulo ou desfecho é numérico (representa valores quantitativos) e classificação para rótulo ou desfecho categórico (representa categorias ou grupos), resultando nos seguintes subconjuntos: *Regressão logística-modelo linear que estima a probabilidade de um evento binário, cujo resultado possível assume apenas dois valores distintos, normalmente codificados como 0 e 1, representando a ausência ou presença de um determinado evento de interesse.; *Árvores de decisão -modelo em formato de fluxograma, no qual cada nó representa uma decisão sobre um atributo, os ramos indicam os resultados e as folhas mostram a classe prevista. Apesar de ser simples e interpretável, tende ao sobreajuste (ajusta excessivamente), que pode ser reduzido com técnicas de poda; *Redes neurais-representada por uma única ou múltiplas camadas não lineares que aprende representações complexas dos dados (Rokach e Maimon 2015; Badillo *et al.* 2020); dentre outros. A Figura 1 descreve os grupos de aprendizado de máquina e alguns de seus algoritmos, conforme acima abordado.

Figura 1. Exemplo de nomenclatura dos diferentes grupos de aprendizado de máquina



Fonte: Adaptado de Badillo *et al.* (2020)

Vale destacar que a literatura nacional e internacional tem utilizado diferentes tipos de algoritmos de AM para prevê o risco de NPT extremo e não NPT extremo, com destaque entre eles a Regressão logística(*glm*), Random Forest (*rf*), Redes Neurais Artificiais (*nnet*), Support Vector Machine (*svm*) Linear e Radial, Extreme Gradient Boosting (*XGBoost*) e que utilizaremos na presente dissertação. Essas abordagens são brevemente abaixo descritas (Adhikari *et al.*, 2019; Alleman *et al.*, 2015; Fergus *et al.*, 2013; Damaso *et al.*, 2019; Gao *et al.*, 2019; Carter *et al.*, 2020; Sun *et al.*, 2022 e Ding *et al.*, 2024):

- **Regressão logística penalizada (*plr*):** modelo supervisionado, que estabelece uma relação entre a probabilidade de ocorrência da variável categórica de interesse e as variáveis preditoras. Possui representação gráfica em formato de “S” (curva logística). É utilizada para tarefas de classificação, com resultados binários para a variável de interesse, sendo um tipo de regressão logística com regularização (L1/L2) para evitar o *overfitting* e selecionar variáveis (Batista; Chiavegatto Filho, 2019).
- **Random Forests (*rf*, *ranger* e *Rborist*):** é um método ensemble que combina múltiplas árvores de decisão para classificação e regressão, gerando previsões mais precisas, robustas a ruídos e tolerantes a outliers. Funciona a partir de amostras aleatórias dos dados e agregação dos resultados, sendo eficaz também em cenários com dados desbalanceados. Entre as implementações, o **ranger** é mais rápido, completo e

amplamente utilizado, enquanto o **Rborist** é mais leve em memória, porém menos atualizado e com menos recursos (Sun et al., 2007 e Liu; Wang; Zhang, 2012).

- **Redes Neurais Artificiais (*nnet*):** utilizada para problemas de classificação e regressão, inspiradas no funcionamento do cérebro humano, possuem múltiplas camadas de processamento, ajustando pesos das conexões para minimizar erros e aprender padrões complexos nos dados, fazendo previsões precisas (Batista; Chiavegatto Filho, 2019).
- **Support Vector Machines (*SVM*) Radial e Linear:** são técnicas de aprendizado supervisionado usadas em classificação e regressão que buscam a melhor fronteira de decisão para separar classes, sendo eficazes especialmente em problemas binários e com alta dimensionalidade. São resistentes ao overfitting, mas podem ter alto custo computacional em grandes conjuntos de dados. O SVM Linear é mais simples e rápido, indicado para dados linearmente separáveis, enquanto o SVM Radial lida melhor com padrões não lineares ao projetar os dados em dimensões maiores. Ambos podem ser aplicados a dados desbalanceados (Hastie; Tibshirani; Friedman, 2009).
- **Boosting (*glmboost e AdaBoost.M1*):** Métodos de Boosting, como *glmboost* e *AdaBoost*, são técnicas supervisionadas para classificação e regressão que constroem modelos de forma sequencial, onde cada novo modelo corrige os erros do anterior. Utilizam modelos fracos (geralmente árvores) e combinam suas previsões por meio de soma ponderada, resultando em maior precisão. São úteis em dados desbalanceados por focarem em instâncias difíceis, embora possam exigir ajustes adicionais em casos mais extremos. (Hastie; Tibshirani; Friedman, 2009).
- **Métodos baseados em vizinhança (K-Nearest Neighbors (*Knn*) e *KKnn*):** é uma técnica de aprendizado supervisionado usada em classificação e regressão que determina o resultado com base nos *k* vizinhos mais próximos de um dado. Na classificação, atribui a classe mais frequente entre esses vizinhos; na regressão, calcula a média dos valores. A escolha do *k* é um hiperparâmetro crucial — geralmente ímpar para evitar empates — e deve ser equilibrada: valores muito baixos podem ser influenciados por outliers, enquanto valores muito altos podem considerar pontos pouco relevantes. Já o *KKnn* fornece maior peso aos vizinhos mais próximos, aumentando sua influência na decisão. (Silva, 2022).

- **Árvores e modelos baseados em partição (*cforest* e *ctree2*):** os modelos baseados em árvores e partição, como Conditional Inference Tree (*ctree/ctree2*) e Random Forest condicional (*cforest*), são técnicas supervisionadas para classificação e regressão que particionam os dados em regiões mais homogêneas por meio de regras de decisão. O *ctree* constrói uma única árvore com base em testes estatísticos, enquanto o *cforest* combina várias árvores, aumentando o desempenho e reduzindo vieses. São úteis em dados desbalanceados, especialmente o *cforest*, embora possam exigir ajustes para melhor desempenho nas classes minoritárias. (Batista; Chiavegatto Filho, 2019).
- **Modelos aditivos generalizados (*gam*):** Os Modelos Aditivos Generalizados são uma extensão dos modelos lineares generalizados que permitem modelar relações não lineares por meio de funções suaves. Trata-se de um método de aprendizado supervisionado, aplicável a problemas de regressão e classificação. Embora possam ser utilizados em bases de dados desbalanceadas, não tratam esse problema diretamente, sendo necessário o uso de técnicas como ponderação de classes ou reamostragem (Hastie; Tibshirani; Friedman, 2009).

Chiavegatto (2018), refere que o desempenho dos algoritmos de aprendizado supervisionado são avaliados por métricas específicas, de acordo com o tipo de problema abordado. Em problemas de classificação, o qual pertence a presente pesquisa, cuja variável desfecho é categórica, utilizam-se métricas como acurácia, precisão, sensibilidade (*recall*), especificidade e F1-score. Conforme descritos no Quadro 1 abaixo.

Quadro 1. Métricas de desempenho utilizadas em aprendizado de máquina de acordo com o tipo de problema

Tipo de problema	Métrica de desempenho	Descrição
Classificação	Acurácia	Proporção de previsões corretas
	Precisão	Proporção de verdadeiros positivos entre os positivos previstos
	Sensibilidade (<i>Recall</i>)	Capacidade do modelo em identificar corretamente os casos positivos
	Especificidade	Capacidade do modelo em identificar os casos negativos
	AUC (Area Under the Curve) – área sob a curva ROC	Mede a capacidade do modelo de distinguir entre as classes (positivo x negativo)

	F1-score	Média harmônica entre precisão e revocação (<i>Recall</i>)
--	----------	--

Fonte: Adaptado de Faceli *et. al.* (2011) e Chiavegatto (2018)

Estudos realizados em diferentes países reportam desempenhos variados dos modelos preditivos relacionados ao NPT extremo e não NPT extremo. Pesquisas no Canadá e nos Estados Unidos utilizaram regressão logística, bootstrapping e redes neurais, obtendo valores moderados a elevados de AUC, sensibilidade e especificidade. No Reino Unido e no Brasil, modelos tradicionais sem uso de aprendizado de máquina também apresentaram resultados satisfatórios, embora com desempenho inferior em alguns cenários. Apesar dos resultados promissores, identificamos poucos estudos que abordam predição de NPT extremo na literatura nacional e internacional, além da limitação externa, tornando-se uma lacuna do conhecimento. Ademais, ressaltam a necessidade de novas pesquisas para aprimorar a predição na identificação do NPT extremo e não NPT extremo (Adhikari et al., 2019; Alleman et al., 2015; Fergus et al., 2013; Damaso et al., 2019; Gao et al., 2019; Carter et al., 2020). O Quadro 2 abaixo destaca uma matriz de estudos com autores pesquisados para a presente dissertação, algoritmos, variáveis utilizadas e métricas de desempenho.

Quadro 2. Matriz de estudos, relacionando autores e seus principais achados com uso de algoritmos de AM em NPT e NPT extremo

Autor/ Ano	Título	Amostra	Variáveis utilizadas	Algoritmos utilizados	Métricas de desempenho
Adhikari et al., 2019	O status socioeconômico do bairro prediz o risco de parto prematuro? Um estudo de coorte canadense baseado na comunidade com modelos em conjunto (<i>ensemble</i>) de redes neurais recorrentes (RNN) treinados com dados artificialmente balanceados.	2.187 prematuros De 5.297 observações	Renda, educação, status de emprego (são status socioeconômico do bairro), idade materna, etnia, estado civil, tempo de permanência na cidade, índice de massa corporal, paridade, gravidez pretendida, consumo de fumo, álcool e drogas antes da gestação, educação materna, renda familiar, apoio social durante a gestação, depressão a qualquer momento na gestação, índice de privação da vizinhança (carência do bairro), renda pessoal mediana do bairro, NPT anterior, comprimento cervical curto.	Modelos de regressão logística convencional multivariável, <i>Bootstrapping</i>	<i>Bootstrapping</i> apresentou melhor desempenho, com AUC 0,75

Alleman et al., 2013	Proposta de um método para a predição do parto prematuro com base em dados clínicos, triagem sérica materna padrão e níveis de colesterol.	200 prematuros De 2.699 observações	Características maternas (comprimento cervical e histórico de parto prematuro), colesterol total de 1º trimestre e alteração do colesterol total entre os trimestres, alfafetoproteína e inibina A do 2º trimestre.	Regressão logística e Bootstrapping	Bootstrapping apresentou Especificidade de 90%; AUC = 69,5
Fergus et al., 2013	Predição de Partos Prematuros a partir de Sinais Eletrohistógráficos (EHG) Utilizando Aprendizado de Máquina	38 prematuros De 300 observações	NPT anterior, gemelaridade, anormalidades cervical e uterina, hemorragia recorrente anteparto, doenças e infecções, qualquer procedimento ou cirurgia invasivos, baixo peso materno, obesidade, etnia e privação social, longas horas de trabalho/noites, uso de álcool e drogas e deficiência de ácido fólico, contrações uterinas	Redes Neurais Artificiais	96% de sensibilidade, 90% de especificidade e área sob a curva (AUC) de 0,95
Gao et al., 2019	Aprendizado Profundo Prediz Parto Prematuro Extremo a partir de Registros Eletrônicos de Saúde no Vanderbilt University Medical Center.	132 prematuro extremo De 25.689 observações	Idade da mãe, raça, multiparidade, gestação gemelar, comprimento cervical curto, distúrbio hipertensivo, lúpus eritematoso sistêmico, uso de sulfato de hidroxicloroquina, presença de depressão, uso de álcool e drogas	Redes Neurais Recorrentes e Modelos Tradicionais (regressão linear, máquinas de vetores de suporte e reforço de gradiente); modelos em conjunto (ensemble) de redes neurais recorrentes (RNN) treinados com dados artificialmente balanceados	Redes Neurais Recorrentes apresentou 96,5% de sensibilidade e 69,8% de especificidade, AUC 0,82
Khatibi et. al., 2019	Análise de Big Data para Predição de Parto Prematuro Induzido pelo Profissional de Saúde e de Partos Prematuros Espontâneos, e Classificação das Características Preditivas	116.080 prematuro De 1.431.597 observações	Idade da mãe, hipertensão crônica, doença cardiovascular, diabetes gestacional, IMC, poluentes do ar, hábitos nutricionais, estilo de vida, comprimento cervical, nacionalidade paterna, anormalidades fetais, propriedades demográficas.	Random Forest, Árvore de decisão e Suport Vector Machine	Árvore de decisão com Acurácia de 81% AUC de 0,68

Lacey et al., 2020	Análise de Compostos Orgânicos Voláteis, uma Nova Ferramenta na Busca pela Predição do Parto Prematuro — um Estudo de Coorte Observacional, de janeiro de 2017 a agosto de 2018	39 prematuros De 493 observações	Idade da mãe, uso do fumo, vaginose bacteriana, IMC, etnia	Aprendizado de máquina (Análise de compostos orgânicos voláteis (COV))	AUC=0.84, com sensibilidade de 0.73 (95%CI 0.64–0.81) e especificidade de 0.90 (95%CI 0.82–0.95)
Bazo, 2020	Desenvolvimento e análise de modelos de estatística multivariada baseados em Aprendizado de Máquina: predição de partos pré-termo na coorte de nascimento BRISA em Ribeirão Preto	77 prematuros e 7.699 observações	Parto prematuro anterior, ameaça de parto PT (antes do tempo) na gestação atual, quantidade de consultas do pré-natal no segundo e terceiro trimestre da gestação, realização do pré-natal e tipo do serviço onde o mesmo foi realizado, número de fetos, hospitalização durante a gestação, ocorrência de doenças durante a gestação atual, hipertensão gestacional, anemia e quantidade de filhos vivos, hábito de fumo antes e durante a gestação, consumo de álcool na gestação, idade da mãe, trabalho no início da gestação, mês de início do pré-natal e a quantidade de ultrassons realizada em todo pré-natal	Random Forest (RF), Support vector machines (SVM), Naive Bayes (NB), Decision Tree (DT), Linear discriminant analysis (LDA) e K-nearest neighbors (KNN)	RF apresentou a maior acurácia (0,913, sensibilidade 0,81 e área sob a curva AUC=0,842. KNN apresentou a maior especificidade (e=0,972).
Koivu e Sairanen, 2020	Previsão do risco de natimortos e gestações prematuras com aprendizado de máquina	1.532.538 prematuros, De 15.883.784, contendo gestações ocorridas entre 2013 e 2016 nos Estados Unidos, foram fornecidos pelo CDC, Centro Nacional de Estatísticas	Idade, raça, estado civil, escolaridade, programa especial de suplementação alimentar, fumo, índice de massa corpórea(IMC), altura peso, paridade, diabetes pré-gestacional e gestacional, hipertensão pré-gestacional e gestacional, parto prematuro anterior, infertilidade, cesarianas anteriores, gonorréia, sífilis, clamídia, hepatite B e hepatite C.	Regressão logística, redes neurais artificiais e árvores de decisão com reforço de gradiente.	A regressão logística alcançou melhor AUC de 0,73 e 0,74

		de Saúde, por meio do seu Sistema Nacional de Estatísticas Vitais			
Araújo et al, 2023	Métodos de aprendizado de máquina aplicados para a previsão de prematuridade e do peso ao nascer no Brasil com base em dados públicos	Base do SINASC/DATASUS 2023, com 35 variáveis selecionadas (maternas, socioeconômicas e obstétricas)	Escolaridade da mãe, raça/cor mãe, número de consultas no pré-natal, IDH do município, mês de início do pré-natal, idade da mãe, tipo de gestação	Random Forest e Regressão Logística	Regressão Logística apresentou melhor desempenho com Acurácia 0,91 e F1score 0,60
Sun Q, Zou X, et al 2022	Modelo de Predição de Parto Prematuro Baseado em Aprendizado de Máquina Utilizando Registros Eletrônicos de Saúde em um hospital na China, entre 2008 à 2018	4.775 prematuros De 9.550 observações (gestantes)	Idade da mãe, exame físico, exame de sangue (glóbulos vermelhos), contagem de glóbulos brancos e plaquetas, tira de teste de urina (pH da urina e glicosúria) e exame ginecológico (vaginose bacteriana, grau de limpeza da vagina e infecção vaginal por fungos)	Naive Bayesian, Support Vector Machine, Random Forest, Redes Neurais Artificiais, K-means e regressão logística	Random Forest Precisão: 0,816; AUC 0,885
Ding, L., Yin, X., Wen, G. et al., 2024	Predição de Parto Prematuro Utilizando Aprendizado de Máquina: uma Análise Abrangente Baseada em Dados em Larga Escala de Pesquisa com Pré-Escolares em Shenzhen, China de 2021 e 2022	6.556 prematuros De 84.050 Pares de mães e filhos	Idade da materna na concepção, modo de gravidez, paridade, complicações na gravidez, ambiente nutricional e metabólico, determinantes sociais (tabagismo) e condições fetais (sexo, retardo de crescimento intrauterino)	L1-Regularised Logistic Regression, Light Gradient Boosting Machine (LightGBM), Naive Bayes, Random Forest, Support Vector Machine, and Extreme Gradient Boosting (XGBoost).	XGBoost com melhor desempenho (AUC-ROC (Receiver Operating Characteristic) de 0,752 e 0,757.

Fonte: Elaborado pela autora, com base nos estudos selecionados.

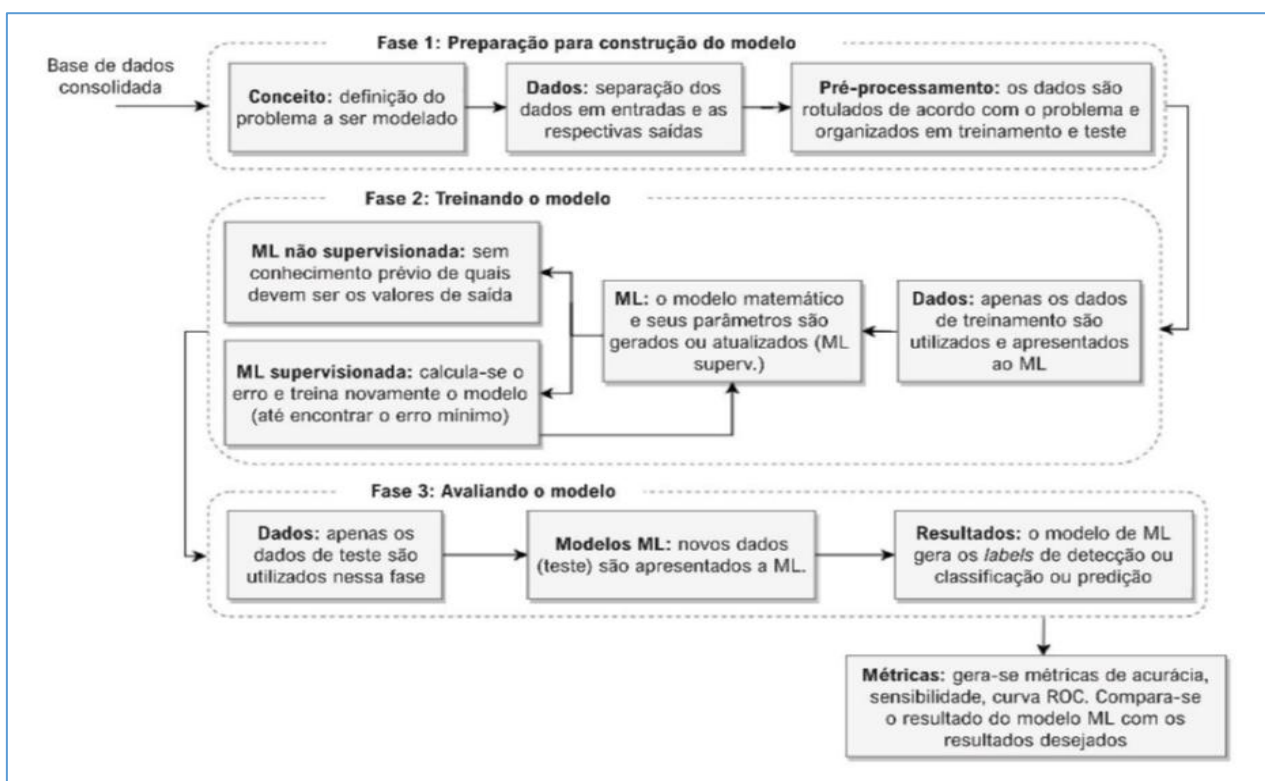
4.4.2 Etapas do desenvolvimento dos modelos preditivos

Na presente pesquisa utilizaremos modelo supervisionado de classificação por ser um evento binário, categorizado em NPT extremo e não NPT extremo, o qual durante o desenvolvimento para a predição do NPT extremo envolve etapas fundamentais, como o pré-processamento dos dados, a divisão em conjuntos de treinamento e teste, a otimização de hiperparâmetros, o treinamento e a avaliação do desempenho por métricas adequadas. Diante desse contexto, a qualidade e a confiabilidade dos dados são determinantes para a eficácia do modelo, tornando a consolidação da base de dados e o pré-processamento uma etapa crítica do processo. Abaixo segue a figura 2 representando as etapas (Rowe, 2019; Paixão et al., 2022).

Diante das pesquisas abordadas, vale ressaltar que na área da saúde, modelos preditivos podem estimar a probabilidade de certos desfechos com base em características socioeconômicas, demográficas, hábitos de vida e condições de saúde, entre outros fatores. A combinação desses modelos com medidas de saúde pública em larga escala pode ajudar a reduzir custos e melhorar a eficácia das intervenções, como tratamentos e ações preventivas (Steyerberg, 2009).

Além disso, entender o risco de desfechos específicos pode auxiliar os gestores na formulação e avaliação de políticas públicas, permitindo uma abordagem mais precisa e eficaz para intervenções preventivas, equilibrando riscos e benefícios (Santos *et al.*, 2019).

Figura 2. Exemplo das etapas do desenvolvimento de aprendizado de máquina com modelos supervisionados em problemas de classificação



Fonte: Paixão *et al.* (2022)

Portanto, o presente estudo avaliou a utilização de algoritmos de AM para a predição do NPT extremo, a partir de dados transversais envolvendo o pré-natal e nascimento de um estudo de coorte situado em 2 cidades brasileiras. O trabalho almejou ainda identificar os principais preditores relacionados ao NPT extremo, bem como algoritmos com melhores desempenhos para o desfecho proposto.

5. MÉTODOS

5.1. Tipo de estudo

Trata-se de um estudo transversal com dados oriundos das coortes BRISA de mulheres com gestações únicas, atendidas em serviços públicos e privados de São Luís e Ribeirão Preto, Brasil. Os dados utilizados provêm de um estudo de coorte denominado “Fatores etiológicos do nascimento pré-termo e consequências dos fatores perinatais na saúde da criança: coortes de nascimentos e pré-natais em duas cidades brasileiras-BRISA” (acrônimo de “Brazilian Birth Cohort Studies, Ribeirão Preto and São Luís”), as quais possuem diferentes contextos socioeconômicos. O objetivo principal da coorte original foi investigar os fatores intergeracionais e suas influências nos indicadores de saúde ao longo da vida, com uma abordagem integrada e colaborativa nas duas cidades, São Luís (MA) e Ribeirão Preto (SP). Ocorreu durante o ano de 2010 nas duas cidades simultaneamente, as análises focaram em eventos perinatais e, posteriormente, examinaram o capital humano, saúde mental, nutrição e sinais de doenças não transmissíveis e todas as quatro coortes coletaram dados sobre pais, mães e seus filhos. Os procedimentos e instrumentos utilizados foram padronizados nas duas localidades para possibilitar comparações futuras (Pereira *et al.*, 2016; Lima *et al.*, 2018; Barbosa *et al.* 2020; Abreu *et al.*, 2023). Nesse estudo para compor o modelo computacional de predição, dados referentes às coortes de pré-natal e nascimento de São Luís (SL) e Ribeirão Preto (SP), foram utilizados.

5.2. Local do estudo

A cidade de São Luís (MA) está localizada no meio-norte do litoral nordestino brasileiro, pertencente a mesorregião norte maranhense, é banhada e limitada pelas Baías de São Marcos e São José, faz divisa com os municípios de Alcântara, São José de Ribamar e Raposa. Em 2010, a população da cidade era de 1.011.943 milhão de habitantes. Destes 538.181 (53,2%) mil eram mulheres, com Índice de Desenvolvimento Humano (IDH) de 0,768, ocupando na ocasião a posição de 249º cidade na classificação do IDH entre as 5.565 cidades brasileiras e 1º entre as 217 cidades do estado do Maranhão, com taxa de escolarização (6 a 14 anos de estudo) de 96,8%. Sua atividade econômica é baseada na metalúrgica do alumínio, na exportação do minério de Carajás e na produção estadual de soja, além do comércio de serviços (IBGE, 2010).

Ribeirão Preto (SP) está localizado na região nordeste do estado de São Paulo e sudeste do Brasil. Em 2010, a população da cidade era de 604 mil habitantes. Destes 314.828 (50,7%)

mil eram mulheres, com Índice de Desenvolvimento Humano (IDH) de 0,800, ocupando na ocasião a 44ª posição entre os 5.565 municípios brasileiros e 19ª entre os 645 municípios de seu estado, com taxa de escolarização (6 a 14 anos de estudo) de 96,9%. É um dos principais centros econômicos do interior paulista, com destaque na indústria, agronegócio, comércio e setor de serviços, possui uma infraestrutura desenvolvida e uma crescente urbanização, é um importante pólo de ensino superior marcado pela presença da renomada instituição de ensino superior, Universidade de Ribeirão Preto, contribuindo para a formação de mão de obra qualificada (IBGE, 2010).

Portanto a escolha das duas cidades, fundamenta-se na participação de ambas na pesquisa das coortes e em suas distintas características sócio-demográficas e econômicas, permitindo a análise de diferentes contextos materno-infantis brasileiros.

5.3. Fonte de dados

Neste estudo foram utilizados dados provenientes das coortes de Pré-Natal e Nascimento Brasileira de São Luís e Ribeirão Preto (BRISA), desenvolvido pela Universidade Federal do Maranhão (UFMA) em parceria com a Faculdade de Medicina de Ribeirão Preto (FMRP). Utilizou-se dados das duas coortes: a iniciada no pré-natal (2010) e a iniciada no nascimento (2010).

As duas primeiras coortes (iniciadas no pré-natal) incluiu mulheres com gestação única, de 22 a 25 semanas de idade gestacional, sendo 1.447 gestantes em São Luís e 1.417 em Ribeirão Preto. Destas, houveram 65 perdas em São Luís, resultando em 1.382 gestantes entrevistadas e 17 em Ribeirão Preto, resultando em 1.400 por ocasião do pré-natal. A amostra da coorte de pré-natal em São Luís foi de conveniência, as mulheres foram recrutadas em três maternidades públicas e um Centro de Saúde. Em Ribeirão Preto, as mulheres foram recrutadas a partir de um cadastro de gestantes existente no município (Rocha, 2019).

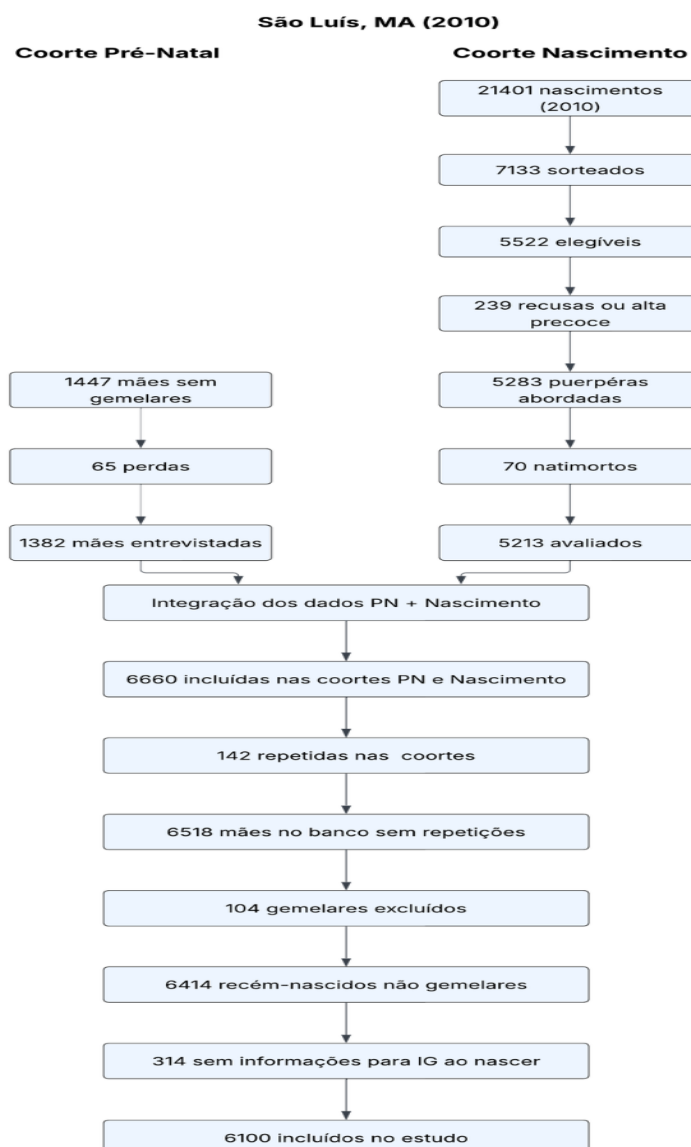
As outras duas coortes iniciadas por ocasião do nascimento da criança, foram representativas do nascimento. Em São Luís a amostra foi probabilística e estratificada por maternidade, com sorteio sistemático, e selecionado um a cada três nascimentos (Rocha, 2019).

Para seleção dos hospitais, realizou-se um levantamento dos dados registrados no Sistema de Informações de Nascidos Vivos (SINASC) e Cadastro Nacional de Estabelecimentos de Saúde (CNES) do Ministério da Saúde, no ano de 2008. No período entre 1º de janeiro a 31 de dezembro de 2010, todos os nascidos vivos ou mortos, de partos únicos ou múltiplos, cujas mães fossem residentes no município de São Luís, foram catalogados

através de uma ficha de controle de nascimento. O estudo pretendia analisar 1/3 de todos os nascimentos ocorridos em 2010, o que correspondeu a aproximadamente 6.000 nascimentos, segundo os dados do levantamento realizado no SINASC. Perdas por recusa ou alta precoce representaram 4.6%.

Nas unidades selecionadas ocorreram 21.401 nascimentos, dos quais foram sorteados 1/3 (7.133). Destes, 5.522 eram residentes no município há pelo menos três meses e, portanto, elegíveis, 239 recusaram ou tiveram alta precoce, resultando em 5.283 puérperas abordadas. Foram entrevistadas, por ocasião do nascimento do bebê, após a exclusão de 70 natimortos, 5.213 puérperas. A amostra resultou em 6.660 dados de pré-natal e nascimento, após a exclusão de 142 gestantes repetidas nas duas coortes, 104 gemelar e 314 com informações ausentes para a idade gestacional, resultando em 6.100 amostras (Figura 3).

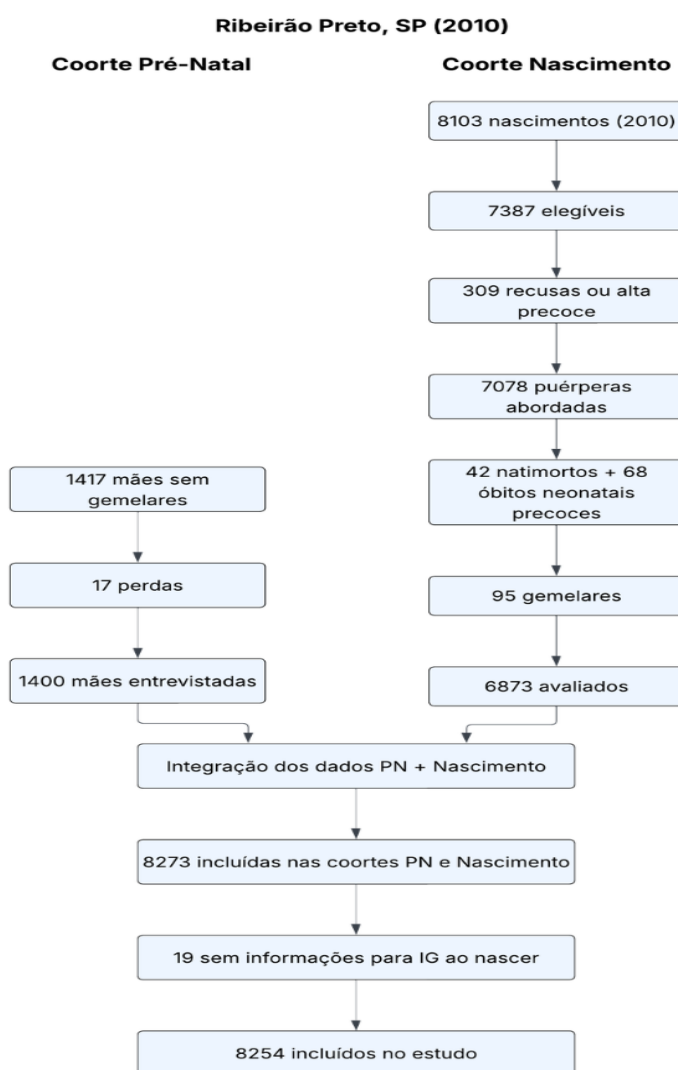
Figura 3. Fluxograma das coortes de pré-natal e nascimento, São Luís – MA, 2010



Fonte: Coortes BRISA

Em Ribeirão Preto, por ocasião do nascimento foram incluídos todos os nascimentos ocorridos em oito instituições, hospitais e maternidades, de caráter público, conveniado e/ou privado, que prestaram assistência ao parto e ao recém-nascido de residentes em Ribeirão Preto no ano de 2010. Do total de 8.103 nascimentos, 7.387 puérperas eram elegíveis, destas 309 não foram encontradas nos hospitais ou recusaram-se a participar do estudo, resultando em 7.078 puérperas entrevistadas. Após exclusão de gemelares (95), natimortos (42) e óbitos infantis (68), o percentual de perda foi de 3,8%, resultando em 6.873 mães de nativos e com partos únicos. Dessa forma, a amostra total do estudo de Ribeirão Preto, incluindo dados do pré-natal e nascimento compreendeu 8.273 e após a exclusão de 19 puérperas sem informação da idade gestacional no nascimento, resultou em 8254 nativos (Figura 4).

Figura 4. Fluxograma das coortes de pré-natal e nascimento, Ribeirão Preto – SP, 2010



5.4 Critérios de inclusão e não inclusão

Foram incluídas mulheres com gestação única que realizaram a 1ª ultrassonografia antes de 20 semanas de gestação e que estavam com 22 a 25 semanas gestacionais e, nascidos após a 20ª semana de idade gestacional (IG) e peso maior ou igual a 500g. Também participaram mulheres com informações completas ou suficientemente detalhadas sobre pré-natal e nascimento que responderam ao questionário BRISA de São Luís e Ribeirão Preto, em 2010.

Foram excluídos dados das coortes de pré-natal e nascimento (BRISA) de mulheres e nascidos vivos de mães não residentes no município, maternidades as quais ocorreram menos de 100 partos ao ano e informações de mulheres graves que estavam internadas em UTI, impossibilitadas em responder aos questionários..

5.5 Coleta de Dados

A coleta de dados foi realizada através do banco de dados e dos questionários de pré-natal e nascimento do banco de dados BRISA 2010 de São Luís e Ribeirão Preto, solicitados aos e-mails pesquisacoortes.slz@ufma.br e rpbrisa@gmail.com. Os dados foram coletados por uma equipe de graduados da área da saúde, inicialmente treinados e identificados para a aplicação dos questionários e procedimentos. Foi realizado estudo piloto com simulação de todas as etapas da pesquisa em todas as unidades hospitalares por 24 horas para checagem e ajustes técnicos (Ribeiro *et. al.*, 2015).

Em relação aos instrumentos utilizados e coleta de dados, todas as gestantes e puérperas foram avaliadas por meio de entrevistas, com a utilização de questionários padronizados especificamente para o estudo, sendo um questionário para os dados da mãe e outro para o filho (Ribeiro *et. al.*, 2015).

5.6 Variáveis do estudo

A variável dependente ou desfecho “NPT extremo”, foi criada a partir da categorização da idade gestacional (IG), sendo binária, com 0 para parto não pré-termo extremo e 1 para parto pré-termo extremo (menos de 28 semanas de gestação ou menos de 196 dias).

A idade gestacional foi estimada com base na data da última menstruação (DUM) e na idade gestacional obtida por ultrassonografia (USG), informadas pela mãe durante a entrevista. Quando houve concordância entre as medidas — considerando uma margem de erro de $\pm 7\%$ para a USG — utilizou-se o tempo de amenorreia para o cálculo. Em casos de discordância, priorizou-se a estimativa por USG. Posteriormente, a idade gestacional foi ajustada para a data

do parto, resultando na medida final, conforme os métodos descritos por Hutcheon et al. (2016) e Medeiros et al. (2015).

Tendo em vista a diversidade e quantidade de fatores de risco relevantes que predis põe ao NPT (nascimento pré-termo) descritos na literatura (Bittar; Fonseca; Zugaid, 2010; Fergus *et al.*, 2013; Ahumada-Barrios, Alvarado, 2016; Guimarães *et al.*, 2017; Adhikari *et al.*; Damaso *et al.* e Gao *et al.*, 2019; Carter *et al.*, 2020; Carvalho *et al.*; Almeida; Bartosch; Macedo 2021). Inicialmente foram selecionadas manualmente pelo presente autor da dissertação, especialista na área de saúde materno-infantil e conforme Simhan; Caritis (2007) que classifica os fatores de risco de acordo com estudos epidemiológicos materno-infantis, 119 variáveis independentes ou preditoras comuns a base de dados de SL e RP, 121 em SL e 119 apenas em RP; com base na plausibilidade teórica do problema e nas variáveis disponíveis no banco de dados dos Questionários Brisa de Entrevista do Pré-natal e do Nascimento 2010.

Em seguida, foi aplicado uma estratégia para facilitar os estudos de investigação conforme Simhan; Caritis (2007), classificando de acordo com estudos epidemiológicos materno-infantis, pelos seguintes fatores: Socioeconômicos, Demográficos e Ambientais; Reprodutivos; Comportamentais e Atenção à Saúde Materna / Pré-natal e Parto. Essa categorização visa estruturar a análise e facilitar a identificação de padrões relevantes ao desfecho estudado.

Posteriormente, aplicou-se o algoritmo BORUTA para selecionar as variáveis mais importantes como preditoras do desfecho (Kursa; Rudnicki, 2010), resultando em 50 variáveis, e elencamos oito com maior número de indicações e maior importância na predição sendo que, diante de sua relevância clínica, realizamos análise de sensibilidade com a variável “ameaça de parto prematuro na gestação atual”, destacada como forte preditor de parto pré-termo extremo, por refletir mecanismos fisiopatológicos já em curso, possivelmente relacionadas ao início do desfecho e sua divisão em dois grupos (com e sem a condição) permitiu avaliar diretamente seu impacto no desempenho dos modelos (Yu QY et al., 2024).

Convém destacar que utilizamos também métodos estatísticos convencionais, como média, mediana e p-valor, para compreender o comportamento e a relevância das variáveis candidatas na predição do NPT extremo.

Para identificar a origem das variáveis utilizadas, dispomos no Anexo A o questionário do pré-natal/entrevista mãe, no Anexo B está o questionário pré-natal auto aplicado e os Anexos C e D, respectivamente, estão os questionários referentes ao nascimento mãe e do recém-nascido (RN).

Portanto, com base nos questionários de pré-natal e nascimento BRISA de São Luís e Ribeirão Preto, as variáveis preditoras foram organizadas em blocos de estudo, significado, tipo da variável, valores possíveis e autor referência conforme ilustrado no Quadro 1, que encontram-se no apêndice.

5.7. Análise dos dados e modelos preditivos

5.7.1. Pré-processamento dos dados

O pré-processamento foi conduzido de forma sistemática para garantir qualidade, consistência e adequação dos dados aos algoritmos de aprendizado de máquina, envolvendo etapas de integração, limpeza, transformação e preparação das variáveis (Harrison; Sidey-Gibbons, 2021).

5.7.1.1. *Consolidação da base de dados*

As bases das coortes de pré-natal (amostragem por conveniência) e nascimento (amostragem probabilística) de São Luís e Ribeirão Preto foram integradas por *linkage* determinístico, utilizando identificador único no software Stata®. Definimos seis cenários analíticos: três com a variável ameaça de parto prematuro na gestação atual e três excluindo essa variável.

- SL-RP combinado: 119 variáveis e 14.687 observações (107 NPT extremo);
- São Luís (SL): 121 variáveis e 6.414 observações (55 NPT extremo);
- Ribeirão Preto (RP): 119 variáveis e 8.273 observações (45 NPT extremo).

5.7.1.2. *Exclusão de observações e controle de duplicidade*

Inicialmente, foram excluídas observações com valores ausentes no desfecho, resultando em:

- 14.354 observações (SL-RP),
- 6.100 (SL)
- 8.254 (RP).

5.7.1.3. *Tratamento de dados faltantes*

O tratamento de *missing data* seguiu critérios de completude:

- Variáveis com $\geq 50\%$ de valores ausentes foram excluídas;
- Variáveis com até 20% de dados faltantes foram imputadas por meio do algoritmo *k*-nearest neighbors (*k*-NN).

O k-NN estima valores ausentes com base nos vizinhos mais próximos no espaço das variáveis, utilizando, em geral, a distância euclidiana (Faceli et al., 2011; Batista; Chiavegatto Filho, 2019).

5.7.1.4. Tratamento das variáveis categóricas

As variáveis categóricas foram transformadas por meio de codificação one-hot (variáveis dummy), convertendo categorias em representações binárias independentes, assumindo o valor 1 (um) para indicar a presença daquela categoria na observação e 0 (zero) para indicar sua ausência, onde cada categoria foi considerada separadamente para este procedimento, permitindo que os algoritmos interpretem corretamente variáveis nominais sem impor ordens artificiais (Kuhn; Johnson, 2013).

5.7.1.5. Tratamento das variáveis quantitativas

As variáveis numéricas foram padronizadas utilizando o escore z , aplicado para ajustar os valores a uma média de 0 e desvio-padrão de 1, garantindo comparabilidade de escala entre preditores. A padronização foi realizada separadamente nos conjuntos de treinamento e teste, evitando vazamento de informação (*data leakage*).

5.7.1.6. Redução de dimensionalidade e refinamento teórico

Foram removidas variáveis com baixa variabilidade (near-zero variance), além de um refinamento baseado no modelo conceitual do estudo e na expertise de pesquisadores da área de saúde materno-infantil. Nesse processo, foram excluídas:

- variáveis não relacionadas ao desfecho,
- variáveis redundantes,
- variáveis conceitualmente sobrepostas ao desfecho.

Posteriormente, aplicou-se o algoritmo Boruta para seleção das variáveis mais relevantes como preditoras (Kursa; Rudnicki, 2010).

Ao final, o conjunto analítico foi reduzido a oito variáveis preditoras, mantendo-se:

- 14.354 observações (SL-RP),
- 6.100 (SL) e
- 8.254 (RP).

5.7.1.7. Balanceamento das classes

Considerando o desbalanceamento do desfecho (baixa frequência de NPT extremo), aplicou-se técnica de subamostragem da classe majoritária (*undersampling*), conforme implementado no pacote UBL. Essa abordagem visa equilibrar a distribuição entre classes e reduzir viés de classificação, melhorando o desempenho dos modelos (Pfob; Lu; Sidey-Gibbons, 2022; Rowe, 2019).

5.7.1.8. Considerações sobre potenciais vieses

Possíveis vieses incluem:

- **viés de informação**, especialmente em variáveis autorreferidas (ex.: hábitos de vida);
- **viés de seleção**, decorrente de perdas de seguimento e da exclusão de maternidades com menor volume de partos.

Para mitigar tais limitações, foram adotadas estratégias como:

- tratamento sistemático de dados faltantes,
- padronização das variáveis,
- balanceamento das classes,
- uso de conjuntos independentes de teste,
- avaliação por métricas de desempenho (ex.: AUC, sensibilidade, F1score).

Essas abordagens contribuem para maior validade interna e estabilidade dos modelos (Hastie; Tibshirani; Friedman, 2009).

5.7.2 Treinamento e ajuste dos modelos preditivos

A etapa de treinamento dos modelos de Aprendizado de Máquina envolve o ajuste dos parâmetros internos dos algoritmos para capturar padrões nos dados e realizar previsões precisas. Para isso, a base de dados foi dividida em conjuntos de treinamento e teste, onde utilizou a validação cruzada k-fold para melhorar a generalização do modelo, sendo utilizado de 5 conjuntos (padrão). Essa técnica divide os dados em k subconjuntos mutuamente exclusivos. Em cada iteração, k-1 subconjuntos foram usados para treinamento, enquanto o subconjunto restante foi usado para teste. Esse processo se repetiu k=5 vezes, alternando o conjunto de teste, evitando que seu desempenho dependa excessivamente de uma única divisão dos dados, garantindo uma avaliação do modelo com melhor performance (Sidey-Gibbons; Sidey-Gibbons, 2019).

Além dos parâmetros internos, os hiperparâmetros também foram ajustados, pois influenciam o desempenho do modelo. Para otimizá-los, utilizamos técnicas como validação

cruzada interna e Grid Search, que testam todas as combinações possíveis de hiperparâmetros. Essas abordagens garantem um melhor ajuste do modelo e aumentam sua capacidade de generalização para novos dados (Batista; Chiavegatto Filho, 2019; Pfob; Lu; Sidey-Gibbons, 2022).

Diante do contexto acima, na presente pesquisa, algoritmos de aprendizado de máquina foram ajustados para identificar padrões nos dados e realizar previsões do desfecho com precisão. A base foi dividida em conjunto de treinamento (80%) e teste (20%), utilizando validação cruzada do tipo k-fold, a fim de avaliar a capacidade de generalização dos modelos e mitigar o risco de overfitting.

Figura 5. Exemplo de Técnica de validação cruzada (5-folds cross-validation)



Fonte: Adaptado de Oschina (2019)

O processamento foi realizado no software R (versão 4.3.3), via RStudio Desktop e pacotes (*haven*, *skimr*, *dplyr*, *caret*, *performanceEstimation*, *Metrics*, *smotefamily*, *doParallel*, *pROC*, *umap*, *Boruta*, *ROSE*, *ebmc*, *UBL*, *gtsummary*, *fastshap* e *shapviz*), seguindo as diretrizes TRIPOD (*Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis*), que orientam o relato de estudos de modelos de predição, promovendo clareza e padronização na apresentação dos métodos e resultados (Moons et al., 2015).

5.7.3 Escolha dos Algoritmos

Os algoritmos de aprendizado de Máquina abrangem diversas famílias, cada uma adequada a diferentes tipos de problemas e conjuntos de dados, com diferentes vantagens e desafios, sendo selecionados conforme a natureza dos dados e a necessidade do estudo (Batista; Chiavegatto Filho, 2019). O tipo de problema do presente estudo, é considerado como de classificação, a variável dependente é categórica e as independentes são categóricas e

numéricas, sendo um evento binário.

Diante disso, foram desenvolvidos e comparados diferentes algoritmos de aprendizado de máquina, escolhidos com base em estudos similares da literatura e por ser desfecho raro com desbalanceamento de dados entre as classes, como os de Fergus et al. (2013), Damaso et al. (2019), Gao et al. (2019), Carter et al. (2020) e Sun et al. (2022). Os principais algoritmos considerados nesta pesquisa, por ser aprendizado supervisionado, incluíram: Regressão logística penalizada (*plr*), máquinas de vetor de suporte (*svmLinear* e *svmRadial*), florestas aleatórias (*rf*, *ranger* e *Rborist*), Boosting (*glmboost* e *AdaBoost.M1*), métodos baseados em vizinhança (*Knn* e *KKnn*), redes neurais artificiais (*nnet*), árvores e modelos baseados em partição (*cforest* e *ctree2*) e modelos aditivos generalizados (*gam*). Ademais, a escolha e a comparação entre esses métodos foram realizadas com base em critérios de métricas de desempenho e interpretabilidade.

5.7.4 Teste e Avaliação de desempenho dos algoritmos

A avaliação de algoritmos de Aprendizado de Máquinas consiste em aplicar o modelo treinado a um conjunto de teste, composto por dados que não foram utilizados durante o treinamento. O objetivo dessa etapa é medir a capacidade de generalização do modelo, ou seja, sua eficácia em prever o desfecho (NPT extremo) em dados novos, ainda não vistos pelo modelo.

Posteriormente, métricas de desempenho foram calculadas para comparar a eficácia geral dos modelos (Badillo *et al.*, 2020; Batista; Chiavegatto Filho, 2019). As principais métricas utilizadas, para mitigar problemas de classificação de duas classes, são descritas a seguir e podem ser extraídas e organizadas conforme apresentadas na matriz de confusão e curva ROC, exemplificadas nas Figuras 6 e 7 abaixo descritas.

Figura 6. Exemplo de uma matriz de confusão para um problema de classificação em duas categorias (1 e 0).

		Valor previsto		
		1	0	
Valor real	1	Verdadeiro Positivo (VP)	Falso Negativo (FN)	Sensibilidade (Taxa de verdadeiros positivos) $\frac{VP}{VP + FN}$
	0	Falso Positivo (FP)	Verdadeiro Negativo (VN)	Especificidade $\frac{VN}{VN + FP}$ Taxa de falso positivo $\frac{FP}{FP + VN}$
		Precisão $\frac{VP}{VP + FP}$	Taxa de falso negativo $\frac{FN}{VP + FN}$	Acurácia $\frac{VP + VN}{VP + FN + VN + FP}$

Fonte: Adaptado de Badillo *et al.* (2020)

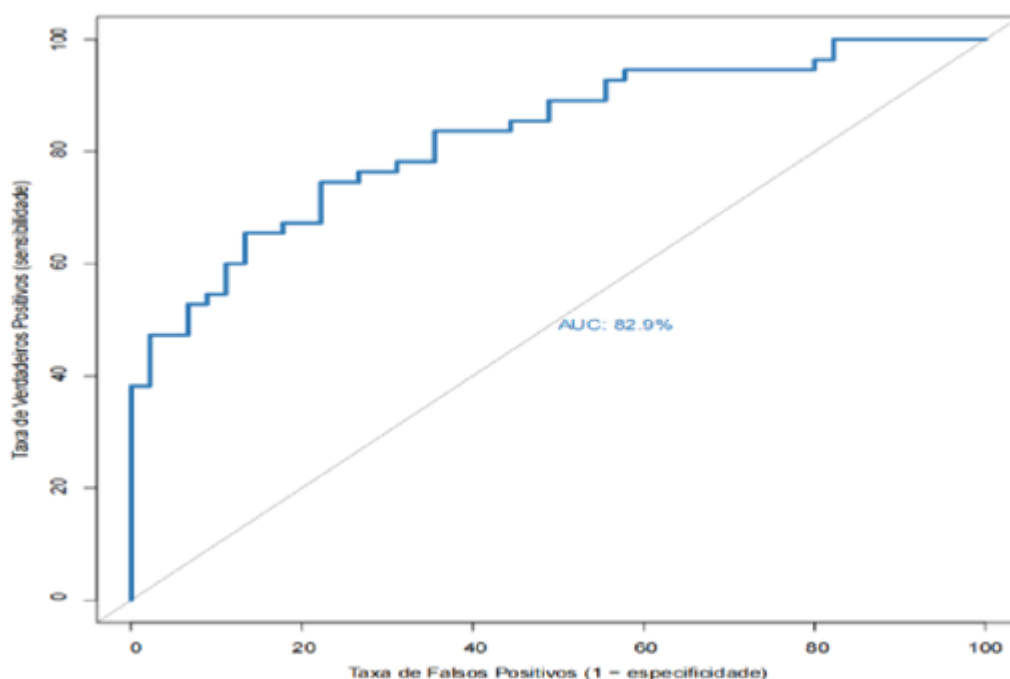
- **Acurácia:** representa a proporção de previsões corretas, considerando tanto os verdadeiros positivos quanto os verdadeiros negativos, em relação ao total de classificações (Badillo *et al.*, 2020).
- **Sensibilidade (Recall):** mede a capacidade do modelo de reconhecer corretamente o NPT extremo, ou seja, a proporção de verdadeiros positivos (Badillo *et al.*, 2020).
- **Especificidade:** indica a habilidade do modelo em identificar corretamente a não ocorrência de NPT extremo, correspondendo à proporção de verdadeiros negativos (Badillo *et al.*, 2020).
- **Valor preditivo positivo (VPP):** mede a probabilidade do teste positivo ser real (Badillo *et al.*, 2020).
- **Valor preditivo negativo (VPN):** mede a probabilidade do teste negativo ser real (Badillo *et al.*, 2020).
- **Precisão:** representa a proporção de instâncias corretamente classificadas como positivas em relação ao total de instâncias classificadas como positivas (Faceli *et al.* 2011).
- **F1-score:** calculado a partir da combinação de duas métricas principais: precisão e recall, representado pela seguinte fórmula (Faceli *et al.* 2011).

$$F1\ Score = 2x \frac{Recall \times Precisão}{Recall + Precisão}$$

Ademais, utilizamos a métrica da curva ROC (*Receiver Operating Characteristic*) a qual permite avaliar o desempenho de modelos de classificação, representando a taxa de

verdadeiros positivos (sensibilidade) em relação à taxa de falsos positivos. O gráfico resultante mostra como o modelo se comporta em diferentes pontos de decisão. No gráfico, o eixo vertical corresponde à sensibilidade, enquanto o eixo horizontal representa a taxa de falsos positivos. A linha diagonal ao centro indica decisões aleatórias, e um modelo eficaz deve ter sua curva posicionada acima dessa linha (Batista; Chiavegatto Filho 2019), como ilustrado na Figura 7.

Figura 7. Exemplo de Curva ROC



Fonte: Batista; Chiavegatto Filho (2019)

Vale ressaltar que a área sob a curva (AUC) é uma métrica que avalia a capacidade do modelo de distinguir corretamente entre classes. Quanto mais próximo de 1 for o valor da AUC, melhor o desempenho do classificador (Batista; Chiavegatto Filho, 2019).

Em relação a análise descritiva dos dados, realizou-se a estimativa de médias e medianas para variáveis numéricas e frequências absolutas e relativas para variáveis categóricas, na população total e segundo os grupos de NPT extremo e não NPT extremo. A normalidade das variáveis quantitativas foi avaliada pelo teste de Shapiro-Wilk, considerando a distribuição dentro dos grupos definidos pelas variáveis independentes. Devido à ausência de normalidade, foram utilizados testes não paramétricos nas análises inferenciais, aplicando-se o teste exato de Fisher para variáveis qualitativas e o teste de soma de postos de Wilcoxon para variáveis quantitativas ou ordinais. Em todas as análises, a inferência estatística foi baseada no p-valor, adotando-se nível de significância de 5% ($\alpha = 0,05$).

Diante desse contexto, ressalta-se que a escolha do modelo envolve também uma análise comparativa de fatores como precisão, interpretabilidade, complexidade computacional e risco de overfitting (Sidey-Gibbons; Sidey-Gibbons, 2019). A precisão reflete a capacidade do modelo de identificar padrões relevantes nos dados; interpretabilidade ou transparência, referindo à capacidade do algoritmo de explicar como chegou a determinada classificação a partir das variáveis preditoras, sendo inversamente proporcional à sua complexidade computacional. Já o risco de overfitting ocorre quando o modelo se ajusta excessivamente aos dados de treinamento, prejudicando sua capacidade de generalização para novos conjuntos de dados (Batista; Chiavegatto Filho, 2019).

Portanto, durante o treinamento e teste, os modelos foram avaliados com base em métricas extraídas da matriz de confusão e da curva ROC. As principais métricas utilizadas foram acurácia, sensibilidade, especificidade, valor preditivo positivo (VPP), valor preditivo negativo (VPN) e área sob a curva ROC (AUC) e F1-score (Badillo et. al, 2020; Batista e Chiavegatto Filho 2019).

5.7.5 Pós-processamento dos dados

O pós-processamento configura-se como uma etapa importante após o treinamento e validação de modelos de aprendizado de máquina, sendo responsável por aplicar métodos voltados à análise, interpretação e validação dos resultados gerados. Essa fase tem como objetivo principal aumentar a transparência e a confiabilidade das predições do modelo, favorecendo sua aplicabilidade prática. Nesta dissertação, uma das abordagens adotadas no pós-processamento é a técnica SHAP (SHapley Additive exPlanations).

O SHAP é uma ferramenta de interpretabilidade de modelos preditivos que visa atribuir a cada variável independente (preditora), uma medida de importância baseada na sua contribuição para o resultado previsto. Por meio de uma abordagem fundamentada na teoria dos jogos cooperativos, o método calcula os valores de Shapley, que representam o impacto de cada variável na predição final, considerando todas as possíveis combinações de atributos (Batista; Chiavegatto Filho, 2019).

No escopo deste estudo, que se propõe a prever o NPT extremo, o uso do SHAP permitiu identificar os principais fatores associados ao desfecho, além de fornecer insights sobre a direção e magnitude dessas associações. Ademais, essa abordagem contribui para a explicabilidade do modelo, facilitando sua interpretação por profissionais da saúde e gestores, ao fornecer justificativas transparentes para a predição realizada.

5.8 Aspectos Éticos e Legais

Em relação aos aspectos éticos, este estudo foi aprovado pelos Comitês de Ética em Pesquisa dos Hospitais Universitário Presidente Dutra (HUUPD), vinculado à UFMA (parecer nº 223/09, registro nº 350/08 e Processo 4771/2008-30) e do Hospital das Clínicas de Ribeirão Preto (HCRP), vinculado à Faculdade de medicina de Ribeirão Preto da Universidade de São Paulo (ofício nº 4116/08 em reunião ordinária 276^a/10.11.08), ambos atendendo às Resolução n. 196/1996 e diretrizes complementares estabelecidas pela Agência Nacional de Saúde Conselho/Ministério da Saúde. Após receberem informações sobre o estudo, as mulheres que concordaram em participar do estudo assinaram o termo de consentimento livre e esclarecido. Documentos encontram-se no anexo F.

6 RESULTADOS

Neste capítulo, os resultados são apresentados em formato de artigo, escrito de acordo com as normas da revista para a qual será submetido, acrescido das outras seções necessárias para a publicação do manuscrito.

Artigo

PREDIÇÃO DE NASCIMENTO PRÉ-TERMO EXTREMO EM DUAS CIDADES BRASILEIRAS

(a ser submetido à revista Ciência & Saúde Coletiva. Fator de impacto 1.2 e Qualis A1)

PREDIÇÃO DE NASCIMENTO PRÉ-TERMO EXTREMO EM DUAS CIDADES BRASILEIRAS

PREDICTION OF EXTREME PRETERM BIRTH IN TWO BRAZILIAN CITIES

Emmanuelle Novaes de Vasconcelos Brito

E-mail: emmanuelle.brito@discente.ufma.br

ORCID: <https://orcid.org/0009-0008-1349-8780>

Bruno Feres de Souza

E-mail: bruno.feres@ufma.br

ORCID: <https://orcid.org/0000-0003-1997-4983>

Erika Bárbara Abreu Fonseca Thomaz

E-mail: erika.barbara@ufma.br

ORCID: <https://orcid.org/0000-0003-4156-4067>

RESUMO

Introdução: O nascimento pré-termo (NPT) extremo é um importante problema de saúde pública, associado a elevada morbimortalidade neonatal e infantil. Sua ocorrência resulta de múltiplos fatores: clínicos, sociodemográficos e comportamentais, o que dificulta sua prevenção. Apesar de sua baixa frequência, seus impactos são significativos para a saúde materno-infantil. Diante desse contexto, modelos de aprendizado de máquina podem predizer gestantes com maior risco de NPT extremo. **Objetivo:** Determinar modelos preditivos baseados em técnicas de aprendizado de máquina (AM), para identificar preditores de NPT extremo, em nascidos vivos de duas cidades brasileiras. **Métodos:** Trata-se de um estudo analítico, de natureza transversal, aninhado às coortes BRISA de pré-natal e nascimento de 2010 em São Luís (SL), Maranhão, e Ribeirão Preto (RP), São Paulo, Brasil, cuja escolha fundamenta-se na participação de ambas na pesquisa das coortes e em suas distintas características sociodemográficas e econômicas, permitindo a análise de diferentes contextos maternos-infantis brasileiros. O conjunto de dados analisados foi proveniente de mulheres com gestação única e seus bebês, atendidos em serviços públicos e privados. A amostra foi composta por 6.414 mulheres/crianças em SL e 8.273 em RP, totalizando 14.687 observações. Realizou-se

análise de sensibilidade com a variável ameaça de parto prematuro na gestação atual, por representar alterações fisiopatológicas em curso, possivelmente relacionadas ao início do desfecho, sendo um importante preditor. Para avaliar sua influência no desempenho dos modelos, foram construídos seis cenários analíticos, correspondentes às coortes combinada, de São Luís e de Ribeirão Preto, com e sem a inclusão dessa variável, permitindo verificar seu impacto na capacidade preditiva e na estabilidade dos modelos. Seleccionamos, com base em especialistas e na literatura, 119 variáveis comuns a SL-RP, 121 variáveis em SL e 119 em RP, relacionadas a fatores socioeconômicos, demográficos, comportamentais, reprodutivos e atenção à saúde da mulher, como potenciais preditoras ao NPT extremo. Em seguida, utilizando-se o algoritmo BORUTA elegeu-se 50 variáveis e elencamos oito, com maior número de indicações e maior importância na predição. Com base em estudos anteriores, aplicaram-se 14 algoritmos de aprendizado de máquina em cada cenário para construção dos modelos: regressão logística penalizada (*plr*), *Support Vector Machines* (*svmLinear* e *svmRadial*), *Random Forest* (*rf*, *ranger* e *Rborist*), *boosting* (*glmboost* e *AdaBoost.M1*), métodos baseados em vizinhança (*Knn* e *KKnn*), redes neurais artificiais (*nnet*), árvores e modelos baseados em partição (*cforest* e *ctree2*) e modelos aditivos generalizados (*gam*). O desempenho foi avaliado por métricas como acurácia, sensibilidade, especificidade, precisão, F1-score e área sob a curva ROC (AUC-ROC). **Resultados:** Os cenários resultaram em desbalanceamento entre NPT extremo e não NPT extremo. Nos seis cenários (SL-RP, SL e RP, três incluindo a variável ameaça de parto prematuro na gestação atual e três excluindo a variável ameaça de parto prematuro na gestação atual), o desempenho variou entre os modelos: No cenário incluindo a variável, em SL-RP o *knn* apresentou sensibilidade de 0.774, AUC-ROC de 0.806 e F1-score de 0.049. No cenário SL, o modelo *svmRadial* obteve sensibilidade de 0.653, AUC-ROC de 0.732 e F1-score de 0.063. Já no cenário RP, o *ranger* apresentou os melhores resultados, com AUC-ROC de 0.86, sensibilidade de 0.771 e F1-score de 0.053. Excluído a variável, em SL-RP o *nnet* apresentou sensibilidade de 0.561, AUC-ROC de 0.648 e F1-score de 0.019. No cenário SL, o modelo *gam* obteve sensibilidade de 0.623, AUC-ROC de 0.587 e F1-score de 0.024. Já no cenário RP, o *Rborist* apresentou os melhores resultados, com AUC-ROC de 0.629, sensibilidade de 0.573 e F1-score de 0.016. Em geral, o *knn* ao incluir a variável e o *nnet* ao excluir a variável, se destacaram pelos maiores valores de AUC-ROC e sensibilidade, apresentaram melhor desempenho para prever o NPT extremo. As oito variáveis preditoras foram: Ameaça de parto prematuro na gestação atual; ameaça de aborto na gestação atual; ocupação da mãe; renda per capita; álcool na gestação; cor da mãe autodeclarada; gravidez planejada e diagnóstico de hipertensão antes ou durante a gestação. **Conclusão:** É possível

predizer, com boa performance, a ocorrência de NPT extremo, utilizando variáveis de saúde materna e características do feto, socioeconômicas, comportamentais e de acesso ao pré-natal. A identificação precoce de bebês em risco pode impactar positivamente a qualidade de vida da mãe e do recém-nascido. As descobertas deste estudo podem estimular investigações semelhantes, testando diferentes técnicas, populações e contextos clínicos, para o enfrentamento de um dos principais desafios de saúde neonatal.

Palavras-chave: Saúde materno-infantil. Coorte de nascimento. Recém-nascido prematuro. Prematuro extremo. Aprendizado de Máquina.

ABSTRACT

Introduction: Extreme preterm birth (EPTB) is a major public health problem associated with high neonatal and infant morbidity and mortality. Its occurrence is influenced by multiple clinical, sociodemographic, and behavioral factors, making prevention particularly challenging. Although relatively uncommon, EPTB has substantial consequences for maternal and child health. In this context, machine learning (ML) techniques offer a promising approach for identifying pregnant women at increased risk of EPTB. **Objective:** To develop and evaluate machine learning-based predictive models to identify predictors of extreme preterm birth among live births from two Brazilian cities. **Methods:** This analytical cross-sectional study was nested within the 2010 BRISA prenatal and birth cohorts conducted in São Luís, Maranhão, and Ribeirão Preto, São Paulo, Brazil. The inclusion of these cohorts allowed the investigation of maternal and child health in distinct socioeconomic and demographic settings. The study included singleton pregnancies receiving care in public and private healthcare services, comprising 6,414 mother–child pairs from São Luís and 8,273 from Ribeirão Preto, totaling 14,687 observations. A sensitivity analysis was performed using the variable threatened preterm labor during the current pregnancy, which represents ongoing pathophysiological changes potentially related to the onset of the outcome and is considered an important predictor. Six analytical scenarios were developed, corresponding to the combined cohort and the São Luís and Ribeirão Preto cohorts, each analyzed with and without this variable. Based on expert knowledge and previous literature, 119 variables common to both cohorts, 121 variables in São Luís, and 119 variables in Ribeirão Preto were selected as potential predictors, including socioeconomic, demographic, behavioral, reproductive, and maternal healthcare

characteristics. The BORUTA feature selection algorithm identified 50 relevant variables, from which the eight most important predictors were retained. Fourteen machine learning algorithms were evaluated in each scenario: penalized logistic regression (PLR), Support Vector Machines (svmLinear and svmRadial), Random Forest (rf, ranger, and Rborist), boosting methods (glmboost and AdaBoost.M1), nearest neighbor methods (*Knn* e *KKnn*), artificial neural networks (nnet), tree-based partitioning models (cforest and ctree2) and generalized additive models (GAM). Model performance was assessed using accuracy, sensitivity, specificity, precision, F1-score, and the area under the receiver operating characteristic curve (AUC-ROC).

Results: All analytical scenarios showed class imbalance between EPTB and non-EPTB births. Model performance varied across the six scenarios. When threatened preterm labor was included, kNN showed the best overall performance in the combined cohort (sensitivity = 0.774, AUC-ROC = 0.806, F1-score = 0.049), svmRadial performed best in São Luís (sensitivity = 0.653, AUC-ROC = 0.732, F1-score = 0.063), and ranger achieved the highest performance in Ribeirão Preto (AUC-ROC = 0.860, sensitivity = 0.771, F1-score = 0.053). When this variable was excluded, nnet performed best in the combined cohort (sensitivity = 0.561, AUC-ROC = 0.648, F1-score = 0.019), GAM in São Luís (sensitivity = 0.623, AUC-ROC = 0.587, F1-score = 0.024), and Rborist in Ribeirão Preto (AUC-ROC = 0.629, sensitivity = 0.573, F1-score = 0.016). Overall, kNN and nnet demonstrated the best predictive performance in scenarios with and without the threatened preterm labor variable, respectively. The eight most important predictors were threatened preterm labor during the current pregnancy, threatened miscarriage during the current pregnancy, maternal occupation, per capita income, alcohol consumption during pregnancy, self-reported maternal race/skin color, planned pregnancy, and hypertension diagnosed before or during pregnancy. **Conclusion:** Machine learning models can predict extreme preterm birth with good predictive performance using maternal health, fetal, socioeconomic, behavioral, and prenatal care variables. Early identification of pregnancies at increased risk may support timely interventions and improve maternal and neonatal outcomes. These findings provide evidence for the potential application of machine learning in prenatal risk stratification and encourage further validation in different populations and clinical settings.

Keywords: Maternal and child health. Birth cohort. Preterm newborn. Extreme preterm infant. Machine learning.

INTRODUÇÃO

O nascimento pré-termo (NPT) ocorre antes de completar as 37 semanas de gestação (Brasil, 2006; Oliveira *et al.*, 2016; Martinelli *et al.*, 2021; WHO, 2023) e representa um importante problema de saúde pública em todo o mundo. Trata-se de uma das principais causas de morte neonatal e ocupa a segunda colocação entre os óbitos de crianças menores de 5 anos, especialmente no período neonatal (Adam *et al.*, 2015; WHO, 2023), sendo relacionado a aproximadamente 1/3 dos óbitos neonatais (WHO, 2023). É ainda relacionado a uma elevada morbidade devido ao incompleto desenvolvimento fetal e alta suscetibilidade a infecções (Suff; Story; Shennan, 2019; WHO, 2023).

Em 2018, de acordo com a Organização Mundial da Saúde (OMS), aproximadamente 15 milhões de bebês nasceram pré-termo, ou seja, com uma média de 1 para cada 10 bebês. Contudo, cerca de 75% dos casos poderiam ter sido evitados, desde que fossem prestados os cuidados adequados durante o pré-natal e parto (WHO, 2019). A maioria dos NPT acontece em países de baixa ou média renda; as maiores taxas ocorrem no sudeste e sul da Ásia e no sul da África subsaariana. Já na América Latina e Caribe, menos de 10% dos partos acontecem antes das 37 semanas de gestação, representando taxa de prematuridade de 8,6%. Em 2019, o Brasil apresentou uma taxa de 13,6% de prematuridade (Aynalem *et al.*, 2021). Em um estudo realizado por Alberton *et al.* (2023) foram notificados 3.503.085 NPT no Brasil, entre 2011 a 2021, correspondendo a uma prevalência de 11,1%, com 300 mil nascidos prematuros anualmente, estando acima da média global de 10%. Entre as regiões do Brasil, a prevalência foi semelhante: Norte (11,6%), Nordeste (11,2%), Sudeste (10,9%), Sul (11,1%), Centro-Oeste (10,8%). Nesse sentido, embora o Brasil apresente cenário semelhante ao dos demais países da América Latina, o país ocupa a décima posição entre os países onde mais se tem nascimentos prematuros no mundo (WHO, 2012).

O NPT pode ser classificado em extremo, grave, moderado e tardio, sendo o NPT extremo (ocorrido antes de completar 28 semanas de idade gestacional) responsável pela maioria dos óbitos entre os prematuros (WHO, 2023). A prevalência de NPT extremo é variada: 0,31% a 0,39 na Noruega (Tingleff *et al.*, 2021; Deb-Rinker *et al.*, 2015; Swamy *et al.*, 2008), 0,02% a 0,50% na Suécia (Crump *et al.*, 2011; Hoglund *et al.*, 2012), 0,15% a 0,38% na Finlândia (Deb-Rinker *et al.*, 2015; Tingleff *et al.*, 2021), 0,05% a 0,39% na Dinamarca (Bilsteen *et al.*, 2018; Morgen *et al.*, 2008), 0,27% na Islândia (Gretarsdottir *et al.*, 2020). Porém, existem poucas estimativas do NPT extremo em países em desenvolvimento (Behboudi-Gandevani *et al.*, 2023). No Brasil, entre 2012 e 2019, as proporções de NPT extremo foi de 0,09%. Todavia, estimativas de 2024 apontam 13.119 nascimentos pré-termo extremo,

representando 0,55%, representando um aumento de 511% (BRASIL, 2025). Portanto, embora a prevalência de NPT extremo seja baixa, representa um grave problema, pela sua alta morbimortalidade (Gette et al., 2025) e pela tendência crescente em diferentes países, especialmente naqueles em desenvolvimento.

O NPT extremo possui etiologia multifatorial, complexa e por vezes desconhecida (Brasil, 2006; Oliveira *et al.*, 2016; WHO, 2023). Alguns fatores relacionados ao NPT extremo envolvem condições socioeconômicas, comportamentais, demográficas, psicossociais e de acesso às ações e serviços de saúde, além das causas não identificadas (WHO, 2012; WHO, 2023). A literatura nacional e internacional aborda como potenciais fatores associados ao NPT, gravidez na adolescência, idade avançada, baixo poder socioeconômico, doenças preexistentes, estilo de vida não saudável da mãe, depressão, intervalo interpartal menor que 2 anos, complicações na gestação atual, baixa adesão a assistência pré-natal e tipo parto (Ouattara *et al.*, 2015; Weber *et al.*, 2018; Adhikari *et al.*, 2019; Gao *et al.*, 2019; Carter *et al.*, 2020; Dias *et al.*, 2022).

Diante disso, nos últimos anos, estudos têm apontado o potencial do uso de inteligência artificial (IA) e aprendizado de máquina (AM) para aprimorar a previsão do NPT extremo, possibilitando a análise integrada de grandes volumes de dados clínicos, sociais e comportamentais. No entanto, a literatura ainda apresenta lacunas importantes, como o número reduzido de pesquisas voltadas especificamente ao NPT extremo, a falta de validação externa dos modelos e a escassez de estudos em populações de países de média e baixa renda (Gao *et al.*, 2019; Bazo, 2020; Lacy *et al.*, 2023). Dessa forma, há necessidade de ampliar as investigações com técnicas de AM aplicadas a bases de dados amplas e heterogêneas, visando desenvolver modelos mais precisos e aplicáveis à prática clínica. Essas abordagens podem contribuir significativamente para a prevenção do NPT extremo e para a melhoria dos desfechos maternos e neonatais (Bazo, 2020; Sun, *et al.*, 2022; Ding *et al.*, 2024; Gul *et al.*, 2024).

Nesse contexto, o presente estudo propõe o desenvolvimento de um modelo preditivo baseado em técnicas de aprendizado de máquina para identificar gestantes com risco de NPT extremo, a partir de dados coletados de coortes de pré-natal e nascimento, a fim de contribuir com conhecimentos e aplicações para o problema mundial de saúde contextualizado.

MÉTODO

Tipo de estudo

Trata-se de um estudo com dados oriundos das coortes prospectivas de pré-natal e nascimento BRISA (acrônimo de “*Brazilian Birth Cohort Studies, Ribeirão Preto, and São*

Luís”). Estas coortes acompanham mulheres e seus filhos que nasceram em instituições públicas ou privadas de São Luís (MA) e Ribeirão Preto (SP), em diferentes contextos socioeconômicos do Brasil (Abreu *et al.*, 2023; Barbosa *et al.* 2020; ima *et al.*, 2018; Pereira *et al.*, 2016; Silva *et al.*, 2014). Para este estudo foram utilizados dados das coortes de São Luís e Ribeirão Preto. A coorte de pré-natal teve início em 2010, na fase de gestação (T0), e as crianças foram seguidas em 2010-2011, por ocasião do parto (T1), em 2012-2013, aos 12-24 meses de idade (T2), e em 2022-2023, quando tinham cerca de 12-13 anos de idade (T3). A coorte de nascimento foi iniciada em 2010 (T1), no parto, e as crianças foram seguidas aos 12-24 meses (T2) e aos 12-13 anos de idade. Nesse estudo serão utilizados dados do T0 e do T1 de ambas as coortes.

Aspectos éticos

O estudo foi aprovado pelos Comitês de Ética em Pesquisa dos Hospitais Universitário Presidente Dutra (HUUPD), vinculado à UFMA (parecer nº 223/09, registro nº 350/08 e Processo 4771/2008-30) e do Hospital das Clínicas de Ribeirão Preto (HCRP), vinculado à Faculdade de medicina de Ribeirão Preto da Universidade de São Paulo (ofício nº 4116/08 em reunião ordinária 276ª/10.11.08), ambos atendendo à Resolução n. 196/1996 e diretrizes complementares estabelecidas pela Agência Nacional de Saúde Conselho/Ministério da Saúde., Todas as gestantes assinaram o termo de consentimento em participação na pesquisa.

Local do estudo

A cidade de São Luís (MA) está localizada no meio-norte do litoral nordestino brasileiro, pertencente a mesorregião norte maranhense, é banhada e limitada pelas Baías de São Marcos e São José, faz divisa com os municípios de Alcântara, São José de Ribamar e Raposa. Em 2010, a população da cidade era de 1.011.943 milhão de habitantes. Destes 538.181 (53,2%) mil eram mulheres, com Índice de Desenvolvimento Humano (IDH) de 0,768, ocupando na ocasião a posição de 249º cidade na classificação do IDH entre as 5.565 cidades brasileiras e 1º entre as 217 cidades do estado do Maranhão, com taxa de escolarização (6 a 14 anos de estudo) de 96,8%. Sua atividade econômica é baseada na metalúrgica do alumínio, na exportação do minério de Carajás e na produção estadual de soja, além do comércio de serviços (IBGE, 2010).

Ribeirão Preto (SP) está localizado na região nordeste do estado de São Paulo e sudeste do Brasil. Em 2010, a população da cidade era de 604 mil habitantes. Destes 314.828 (50,7%) mil eram mulheres, com Índice de Desenvolvimento Humano (IDH) de 0,800, ocupando na ocasião a 44º posição entre os 5.565 municípios brasileiros e 19º entre os 645 municípios de

seu estado, com taxa de escolarização (6 a 14 anos de estudo) de 96,9%. É um dos principais centros econômicos do interior paulista, com destaque na indústria, agronegócio, comércio e setor de serviços, possui uma infraestrutura desenvolvida e uma crescente urbanização, é um importante pólo de ensino superior marcado pela presença da renomada instituição de ensino superior, Universidade de Ribeirão Preto, contribuindo para a formação de mão de obra qualificada (IBGE, 2010).

Portanto a escolha das duas cidades, fundamenta-se na participação de ambas na pesquisa das coortes e em suas distintas características sociodemográficas e econômicas, permitindo a análise de diferentes contextos materno-infantis brasileiros.

Amostra do estudo

A primeira coorte, iniciada no pré-natal, incluiu mulheres com gestação única, de 22 a 25 semanas de idade gestacional, sendo 1.447 gestantes em São Luís e 1.417 em Ribeirão Preto (baseline ou T0). Destas, 1.382 (95,4%) em São Luís e 1.370 (96,7%) em Ribeirão foram reavaliadas por ocasião do parto (T1). A amostra de pré-natal em São Luís foi por conveniência, as mulheres foram recrutadas em três maternidades públicas e um Centro de Saúde. Em Ribeirão Preto, as mulheres foram recrutadas a partir de um cadastro de gestantes existente no município (Rocha, 2019).

As outras duas coortes, iniciada por ocasião do nascimento da criança (T1), foi representativa dos nascimentos dos municípios. Em São Luís, a amostra foi probabilística e estratificada por maternidade, com sorteio sistemático, e selecionado um a cada três nascimentos (Rocha, 2019).

Para seleção dos hospitais, realizou-se um levantamento dos dados registrados no Sistema de Informações de Nascidos Vivos (SINASC) e Cadastro Nacional de Estabelecimentos de Saúde (CNES) do Ministério da Saúde, no ano de 2008. No período entre 1º de janeiro a 31 de dezembro de 2010, todos os nascidos vivos ou mortos, de partos únicos ou múltiplos, cujas mães fossem residentes no município de São Luís, foram catalogados através de uma ficha de controle de nascimento. O estudo pretendia analisar 1/3 de todos os nascimentos ocorridos em 2010, o que correspondeu a aproximadamente 6.000 nascimentos, segundo os dados do levantamento realizado no SINASC. Perdas por recusa ou alta precoce representaram 4,6%.

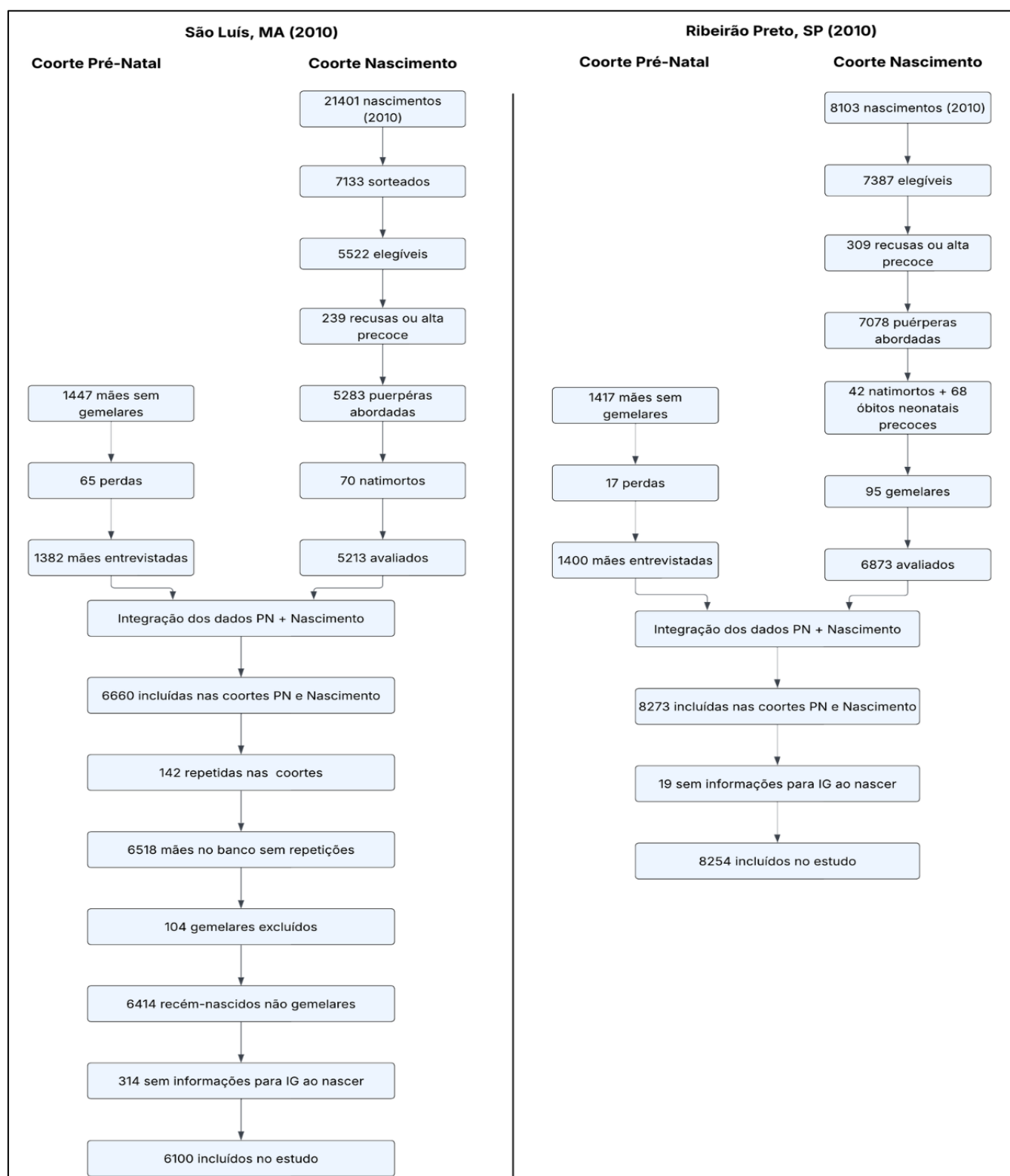
Nas unidades selecionadas ocorreram 21.401 nascimentos, dos quais foram sorteados 1/3 (7.133). Destes, 5.522 eram residentes no município há pelo menos três meses e, portanto, elegíveis. Foram entrevistadas, por ocasião do nascimento do bebê, 5.283 puérperas. Após a

exclusão de 70 natimortos a amostra incluída foi de 5.213 nascidos vivos (**Figura 1A**). Em Ribeirão Preto, por ocasião do nascimento foram incluídos todos os nascimentos ocorridos em oito instituições, hospitais e maternidades, de caráter público, conveniado e/ou privado, que prestaram assistência ao parto e ao recém-nascido de residentes em Ribeirão Preto no ano de 2010. Do total de 8.103 nascimentos, 7387 eram elegíveis, 309 não foram encontradas nos hospitais ou recusaram-se a participar do estudo, resultando em 7.078 puérperas entrevistadas. Após exclusão de gemelares (95), natimortos (42) e óbitos infantis (68), o percentual de perda foi de 3,8%, resultando em 6.873 mães de nativos e com partos únicos. Dessa forma, a amostra total do estudo de Ribeirão Preto incluindo dados do pré-natal e nascimento, compreendeu 8.254 nativos (**Figura 1B**).

Figura 1. Fluxograma da amostra das coortes BRISA São Luís, MA e Ribeirão Preto, SP (2010).

1A. São Luís- MA, 2010

1B. Ribeirão Preto- SP, 2010



Fonte: Coortes BRISA

Critérios de inclusão e não inclusão

Foram incluídas mulheres com gestação única que realizaram a 1ª ultrassonografia antes de 20 semanas de gestação, por ser uma estimativa mais confiável da IG e que estavam com 22 a 25 semanas gestacionais e, nascidos após a 20ª semana de idade gestacional (IG) e peso maior ou igual a 500g. Também participaram mulheres com informações completas ou suficientemente detalhadas sobre pré-natal e nascimento que responderam ao questionário BRISA de São Luís e Ribeirão Preto, em 2010.

Foram excluídos dados das coortes de pré-natal e nascimento (BRISA) de mulheres e nascidos vivos de mães não residentes no município e maternidades as quais ocorreram menos de 100 partos ao ano; gravidez múltipla, considerada um dos principais determinantes do parto prematuro, mesmo na ausência de complicações e mulheres graves que estavam internadas em UTI, impossibilitadas em responder aos questionários.

Coleta de Dados

Foi solicitado o banco de dados BRISA São Luís e Ribeirão Preto 2010 aos e-mails pesquisacoortes.slz@ufma.br e rpbrisa@gmail.com respectivamente, onde os dados foram provenientes da pesquisa realizada no ano de 2010, através dos questionários de pré-natal e nascimento do banco de dados BRISA 2010 de São Luís e Ribeirão Preto. Naquela data fora realizado um estudo piloto com simulação de todas as etapas da pesquisa em todas as unidades hospitalares, identificadas no anexo C, por 24 horas para checagem e ajustes técnicos (Ribeiro *et. al.*, 2015). Todas as gestantes e puérperas foram avaliadas por meio de entrevistas face a face ou, para coleta de dados sensíveis, por entrevistas autoaplicadas, com a utilização de questionários padronizados especificamente para o estudo, sendo um questionário para os dados da mãe e outro para o filho e todas as participantes assinaram o termo de consentimento livre e esclarecido. Também foram coletados dados de prontuários a partir de um roteiro de coleta. Os dados foram coletados por uma equipe de graduados da área da saúde, inicialmente treinados e identificados para a aplicação dos questionários e procedimentos (Silva *et. al.*, 2014).

Variáveis do estudo

A variável dependente ou desfecho “NPT extremo”, foi criada a partir da categorização da idade gestacional (IG), sendo binária, com 0 para parto não pré-termo extremo e 1 para parto pré-termo extremo (menos de 28 semanas de gestação ou menos de 196 dias). A IG foi calculada utilizando informações sobre a data da última menstruação (DUM) e da idade gestacional estimada pela ultrassonografia (USG), fornecida pela mãe durante a entrevista. Quando as

informações eram compatíveis, considerando um erro de +7% para a USG, o tempo de amenorreia foi utilizado para calcular a IG. Caso contrário, a informação da USG foi priorizada. A idade gestacional foi ajustada para a data do parto, gerando a IG definitiva, conforme os métodos descritos por Hutcheon *et al.* (2016) e Medeiros *et al.* (2015).

Tendo em vista a diversidade e quantidade de fatores de risco relevantes que predispõe ao NPT descritos na literatura (Bittar; Fonseca; Zugaid, 2010; Fergus *et al.*, 2013; Ahumada-Barrios, Alvarado, 2016; Guimarães *et al.*, 2017; Adhikari *et al.* 2019; Damaso *et al.* e Gao *et al.*, 2019; Carter *et al.*, 2020; Carvalho *et al.*; Almeida; Bartosch; Macedo 2021), com base nos questionários de pré-natal e nascimento BRISA de São Luís e Ribeirão Preto, as variáveis foram selecionadas manualmente com base na experiência do autor, especialista em saúde materno e infantil e classificadas de acordo com a categoria dos fatores de risco para NPT relatados na literatura, como o proposto por Simhan & Caritis (2007), categorizado em fatores socioeconômicos, demográficos, reprodutivos, comportamentais e atenção à saúde materna. Portanto, as principais variáveis preditoras foram organizadas em blocos, nome da variável e valores possíveis conforme ilustrado no Quadro 1 e os códigos encontram-se no anexo E.

Quadro 1. Exemplo da configuração das variáveis do estudo

Bloco de estudos	Nome da variável	Valores possíveis
Determinantes sociodemográficos	Classificação socioeconômica	Classe A/B; Classe C; Classe D/E
	Chave	
	Cidade	Ribeirão Preto= 1; São Luís=2
	Indivíduo com maior renda na família	A entrevistada; Companheiro; Mãe; Pai; Avó; Avô; Madrasta; Padrasto; Tia; Tio; Irmã; Irmão; Outro; Não sabe
	Cor da mãe	Branca; Não Branca
	Idade da mãe	Número natural
	Ocupação da mãe	Qualificação não-manual; Manual semi-especializada; Manual não-qualificada; Desempregado(a)
	Ocupação do Chefe da Casa	Qualificação não-manual; Manual semi-especializada; Manual não-qualificada; Desempregado(a)
	Frequentou escola	Não; Sim
	Situação conjugal da puérpera	Casada; União consensual (Mora junto); Solteira; Separada/desquitada/divorciada; Viúva
	Último curso que fez	Alfabetização de jovens e adultos; Ensino fundamental ou 1º grau; Ensino médio ou 2º grau; Superior graduação incompleto; Superior graduação completo; Não se aplica; Não sabe
	Renda per capita	Qualquer número real
	Pessoa que realiza atividades domésticas	A sra. faz todo o trabalho; A sra. faz parte do trabalho; Não sabe

Determinantes comportamentais	Tomou bebida alcoólica na gestação	Não; Sim
	Usou drogas ilícitas na gestação	Não; Sim
	Fumou durante a gravidez	Não; Sim
	Duração de caminhada (mínimo 10 min/dia)	Número natural
	Dias da semana de prática de atividades moderadas (mínimo 10min/dia)	Número natural
Saúde materna, pré-natal e parto	Idade Gestacional em semanas	Qualquer número real
	Altura do pai	Qualquer número real
	Peso antes de engravidar	Qualquer número real
	Peso ao final da gravidez	Qualquer número real
	Ganho de peso durante a gestação	Qualquer número real
	Peso pai	Qualquer número real
	Número de comorbidades	Número natural entre 0 e 10
	Comorbidades que acometeram a mãe durante a gestação (hipertensão, diabetes, toxoplasmose sífilis, herpes, sarampo, catapora, rubéola, anemia, ou infecção urinária)	Nenhuma; Uma ou mais
	Mês de início da consulta pré-natal	Número natural entre 0 e 9
	Quantidade consulta pré-natal no 1º trimestre	Número natural
	Quantidade consulta pré-natal no 2º trimestre	Número natural
	Consultas pré-natal	Número natural
	Quantidade consulta com enfermeiro	Número natural
	Quantidade consulta com médico	Número natural
	Médico ou enfermeira encaminhou para tratamento dentário durante a gravidez	Sim; Não; Não se aplica; Não sabe
	Diabetes antes ou durante a gestação	Não; Sim
	Hipertensão na gestação atual	Sim; Não; Não se aplica; Não sabe
	Herpes durante a gestação	Sim; Não; Não se aplica; Não sabe
	Infecção urinária durante a gestação	Sim; Não; Não se aplica; Não sabe
	Rubéola durante a gestação	Sim; Não; Não se aplica; Não sabe
	Sarampo durante a gestação	Sim; Não; Não se aplica; Não sabe
	Sífilis antes da gestação	Sim; Não; Não se aplica; Não sabe
	Sífilis durante a gestação	Sim; Não; Não se aplica; Não sabe
	Toxoplasmose antes ou durante a gestação	Sim; Não; Não se aplica; Não sabe
	Catapora durante a gestação	Sim; Não; Não se aplica; Não sabe
	Episódio de febre alta durante a gestação	Sim; Não; Não se aplica; Não sabe
	Tomou vacina contra tétano	Sim; Não; Não se aplica; Não sabe
	Ameaça parto prematuro na gestação atual	Sim; Não; Não se aplica; Não sabe

	Ameaça de aborto na gestação atual	Não; Sim
	Motivo ida ao hospital	Sangramento vaginal; Perdeu líquido (água) vaginal; Sentiu contração ou dor ou cólica/ barriga endurecida; Febre/ infecção/infecção urinária; Bebê parou de mexer; A vinda foi agendada para esta data; O médico encaminhou; Outro; Não sabe
	Gravidez planejada	Sim; Não; Não se aplica; Não sabe
Saúde do recém-nascido	Sexo do RN	Masculino; Feminino
	Problema de saúde	Sim; Não; Não se aplica; Não sabe
	Defeito congênito	Sim; Não; Não se aplica; Não sabe
	Tipo de defeito congênito	Gastrosquise; Onfalocele; Defeitos do tubo neural- mielomeningocele; Defeitos do tubo neural-encefalocele; Trato genito-urinário- unilateral; Trato genito-urinário- bilateral; Cardíaco; Outro
	Nascimento pré-termo extremo	Sim; Não

Fonte: Instrumentos de coleta de dados – coortes BRISA (2010)

Análise dos dados e modelos preditivos

A presente pesquisa é proveniente de 2 bases de dados, composta por 6.414 mulheres/crianças em SL e 121 variáveis; em RP com 8.273 e 119 variáveis, totalizando 14.687 observações. Selecionamos com base em especialistas e na literatura 8 variáveis comuns de SL e RP, relacionadas a fatores socioeconômicos, demográficos, ambientais, reprodutivos, comportamentais e atenção à saúde materna, assistência pré-natal e ao parto, como potenciais preditoras ao NPT extremo. Criamos seis cenários distintos: três incluindo a variável ameaça de parto prematuro na gestação atual e outros três excluindo a variável ameaça de parto prematuro na gestação atual, composto por dados de SL e RP juntos (107 NPT extremo); apenas com SL (55 NPT extremo); apenas com RP (45 NPT extremo).

Realizou-se análise descritiva dos dados, com estimativa de médias e medianas para variáveis numéricas e frequências absolutas e relativas para variáveis categóricas, na população total e segundo os grupos de NPT extremo e não NPT extremo. A normalidade das variáveis quantitativas foi avaliada pelo teste de Shapiro-Wilk, considerando a distribuição dentro dos grupos definidos pelas variáveis independentes. Devido à ausência de normalidade, foram utilizados testes não paramétricos nas análises inferenciais, aplicando-se o teste exato de Fisher para variáveis qualitativas e o teste de soma de postos de Wilcoxon para variáveis quantitativas ou ordinais. Em todas as análises, a inferência estatística foi baseada no p-valor, adotando-se nível de significância de 5% ($\alpha = 0,05$).

Para a modelagem preditiva do nascimento pré-termo extremo (NPT extremo), variáveis quantitativas foram normalizadas pelo escore-z, aplicados para ajustar os valores a uma média de 0 e desvio-padrão de 1 e variáveis qualitativas foram transformadas convencionalmente em dummies (codificação one-hot encoding), formato binário o qual assume o valor 1 (um) para indicar a presença daquela categoria na observação e 0 (zero) para indicar sua ausência, garantindo que os modelos possam interpretá-las corretamente (Faceli *et al.*, 2011; Batista; Chiavegatto Filho, 2019).

Variáveis com ≥ 50 % de dados faltantes foram removidas, e aquelas com até 20 % de ausência foram imputadas por k-NN, algoritmo baseado em instâncias que classifica um ponto com base nos k vizinhos mais próximos, calculando a distância (geralmente euclidiana) entre o novo ponto e os dados de treino (Faceli *et al.*, 2011). Posteriormente, aplicou-se o algoritmo BORUTA para selecionar as variáveis mais importantes como preditoras do desfecho (Kursa; Rudnicki, 2010), resultando em 50 variáveis, e elencamos oito com maior número de indicações e maior importância na predição. Realizou-se uma análise de sensibilidade, com a variável “ameaça de parto prematuro na gestação atual” e diante de sua relevância clínica, foi destacada como forte preditor de parto pré-termo extremo, por refletir mecanismos fisiopatológicos já em curso, possivelmente relacionados ao início do desfecho e sua divisão em dois grupos (com e sem a condição) permitiu avaliar diretamente seu impacto no desempenho dos modelos e na estratificação de risco (Yu QY *et al.*, 2024).

A base foi dividida em conjuntos de treinamento (80%) e teste (20%), valores usualmente utilizados na literatura, com validação cruzada e otimização de hiperparâmetros por cross-validation e Grid Search (Kuhn & Johnson, 2013; Hastie, Tibshirani & Friedman, 2009). O processamento foi realizado no software R (versão 4.3.3), via RStudio Desktop, seguindo as diretrizes TRIPOD (*Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis*), que orientam o relato de estudos de modelos de predição, promovendo clareza e padronização na apresentação dos métodos e resultados (Moons *et al.*, 2015).

Durante o pré-processamento, verificaram-se consistência, padronização e ausência de dados; variáveis contínuas foram normalizadas e o desfecho desbalanceado foi ajustado por técnicas de undersampling (redução da classe majoritária). Foram ajustados e comparados diferentes algoritmos — regressão logística penalizada(*plr*), máquinas de vetor de suporte(*svmLinear* e *svmRadial*), florestas aleatórias(*rf*, *ranger* e *Rborist*), boosting(*glmboost* e *AdaBoost.M1*), métodos baseados em vizinhança(*Knn* e *KKnn*), redes neurais artificiais(*nnet*), árvores e modelos baseados em partição(*cforest* e *ctree2*) e modelos aditivos generalizados(*gam*)— selecionados por métricas de desempenho e interpretabilidade (Fergus

et al., 2013; Damaso et al., 2019; Gao et al., 2019; Carter et al., 2020). O desempenho dos modelos foi avaliado por métricas derivadas da matriz de confusão e da curva ROC: acurácia, sensibilidade, especificidade, valores preditivos positivo (VPP), valores preditivos negativo (VPN) e área sob a curva ROC (AUC), conforme Faceli et al. (2011); Badillo et al. (2020) e Batista & Chiavegatto Filho (2019).

Para a seleção do modelo final, os algoritmos foram ranqueados com base na AUC média, seguida da sensibilidade e do escore F1, priorizando modelos com melhor capacidade discriminativa e desempenho na classe minoritária. O modelo com melhor desempenho global foi selecionado para a etapa final.

RESULTADOS

Com 95% de intervalo de confiança e N de 107(SL-RP), a maioria das mães eram não brancas (74%), com uma média e mediana de idade (24 anos), residiam em SL (54%) e em RP (46%), não planejaram sua gestação (77%), pertencentes à classe D/E (61%), estavam desempregadas (57%), e dentre as que trabalhavam, a maioria trabalhava de qualificação não manual (21%), com renda per capita mensal (± 404.54 reais), realizaram em média (± 1.15) consulta de pré-natal no primeiro trimestre e (± 1.36) no segundo trimestre, dentre essas consultas (± 2.61) foram com o profissional médico e (± 1.27) com o enfermeiro, eram não hipertensas (82%), tiveram ameaça de parto prematuro na gestação atual (70%) e ameaça de aborto na gestação atual (21%). Observou-se que a prevalência de NPT extremo (< 28 semanas) variou conforme algumas características maternas e assistenciais, destacando-se estatisticamente significativas representada quando P-valor $< 0,05$ e, portanto, apresentou semelhança estatísticas entre a ameaça de parto prematuro na gestação atual (< 0.001), o planejamento da gestação ($p < 0.001$) e o acompanhamento ($p < 0,001$) pré-natal, e diferindo em ameaça de aborto na gestação atual (0.005) e cor da mãe autodeclarada (0.001). As gestantes que tiveram parto prematuro extremo iniciaram o pré-natal após o 1º mês de gestação (entre 2 a 4 meses) e realizaram menor número de consultas no segundo trimestre de gestação (1 à 3) em comparação às que tiveram partos não pré-termo extremo (2 à 3; $p < 0,001$). Além disso, entre aquelas que tiveram NPT extremo, (77%) não planejaram sua gestação em relação as demais (60% $p = < 0,001$), dentre essas (70%) tiveram ameaça de parto prematuro na gestação atual em relação às demais (13%; $p = < 0,001$) e (21%) tiveram ameaça de aborto na gestação atual em relação às demais (7.3%; $p = 0,005$), sugerindo possível influência de condições de saúde reprodutiva ou intercorrências clínicas. Em contrapartida, variáveis socioeconômicas, como classe econômica e ocupação da mãe, e de saúde materna, como hipertensão na gestação

e comportamentais como consumo de álcool antes ou durante a gestação, não apresentaram associação significativa com a ocorrência de NPT extremo (**Tabela 1**).

Tabela 1. Exemplo da distribuição das variáveis preditoras e desfecho de pré-natal e nascimento da Coorte BRISA de 2010, São Luís/MA e Ribeirão Preto/SP.

Característica/código	NPT extremo N = 107	Não NPT extremo N = 14.247	p-valor	Total N = 14.354
Consumo de álcool (álcool)			0.5 ¹	
Não	85 (79%)	11,677 (82%)		11,762 (82%)
Sim	22 (21%)	2,602 (18%)		2,624 (18%)
Ameaça de parto prematuro na gestação atual (ampt)			<0.001 ¹	
Não	31 (30%)	12,292 (87%)		12,323 (86%)
Sim	74 (70%)	1,894 (13%)		1,968 (14%)
Ignorado/Sem Informação	2	93		95
Classificação socioeconômica categorizada (Abepcat)			0.12 ¹	
D/E	61 (61%)	9,086 (68%)		9,147 (68%)
C	33 (33%)	3,223 (24%)		3,256 (24%)
A/B	6 (6.0%)	1,056 (7.9%)		1,062 (7.9%)
Ignorado/Sem informação	7	914		921
Ameaça de aborto na gestação atual (amaborto)			0.005 ¹	
Sim	12 (21%)	443 (7.3%)		455 (7.4%)
Não	46 (79%)	5,620 (93%)		5,666 (92%)
Não Sabe	0 (0%)	9 (0.1%)		9 (0.1%)
Ignorado/Sem Informação	49	8,207		8,256
Cidade			0.018 ¹	
Ribeirão Preto	49 (46%)	8,205 (57%)		8,254 (57%)
São Luís	58 (54%)	6,074 (43%)		6,132 (43%)
Cor da mãe autodeclarada (cormae2)			0.001 ¹	
Branca	27 (26%)	5,805 (41%)		5,832 (41%)
Não Branca	78 (74%)	8,356 (59%)		8,434 (59%)
Ignorado/Sem Informação	2	118		120
Ocupação da mãe (ocupa_mae2)			0.7 ¹	
Desempregado	60 (57%)	7,797 (55%)		7,857 (55%)
Manual Não Qualificado	9 (8.5%)	1,650 (12%)		1,659 (12%)
Qualificação	15 (14%)	1,723 (12%)		1,738 (12%)
Semiespecializado				
Qualificação Não Manual	22 (21%)	3,015 (21%)		3,037 (21%)
Ignorado/Sem Informação	1	94		95
Idade da mãe (idademae)			0.2 ²	
Median (Q1 - Q3)	24.00 (18.00 - 27.00)	25.00 (21.00 - 29.00)		25.00 (21.00 - 29.00)
Mean (± SD)	24.21 (± 6.32)	25.29 (± 5.93)		25.28 (± 5.93)
Ignorado/Sem informação	49	8,207		8,256
Infecção urinária na gestação			0.6 ¹	

Característica/código	NPT extremo N = 107	Não NPT extremo N = 14.247	p-valor	Total N = 14.354
(infecuri)				
Não	76 (73%)	9,973 (70%)		10,049 (70%)
Sim	28 (27%)	4,210 (30%)		4,238 (30%)
Ignorado/Sem informação	3	96		99
Diagnóstico de hipertensão antes ou durante a gestação (hipert2)			0.4 ¹	
Não	86 (82%)	12,032 (85%)		12,118 (85%)
Sim	19 (18%)	2,178 (15%)		2,197 (15%)
Ignorado/Sem informação	2	69		71
Mês de gestação iniciou o pré-natal (mes1cpn2)			0.2 ²	
Median (Q1 - Q3)	3.00 (2.00 - 4.00)	2.00 (2.00 - 3.00)		2.00 (2.00 - 3.00)
Mean (± SD)	8.07 (± 20.79)	3.38 (± 7.62)		3.42 (± 7.79)
Ignorado/Sem informação	11	489		500
Gravidez planejada (plangest)			<0.001 ¹	
Não	81 (77%)	8,448 (60%)		8,529 (60%)
Sim	24 (23%)	5,738 (40%)		5,762 (40%)
Ignorado/Sem informação	2	93		95
Sífilis diagnosticada na gestação (sifilisgest)			>0.9 ¹	
Não	104 (100%)	14,092 (99%)		14,196 (99%)
Sim	0 (0%)	83 (0.6%)		83 (0.6%)
Ignorado/Sem informação	3	104		107
Quantidade de consulta com enfermeiro(a) (qtcpenf2)			0.6 ²	
Median (Q1 - Q3)	0.00 (0.00 - 1.00)	0.00 (0.00 - 1.00)		0.00 (0.00 - 1.00)
Mean (± SD)	0.64 (± 1.27)	0.75 (± 1.55)		0.75 (± 1.55)
Ignorado/Sem informação	5	482		487
Quantidade de consulta com médico(a) (qtcpnmed2)			<0.001 ²	
Median (Q1 - Q3)	3.00 (1.00 - 5.00)	8.00 (5.00 - 10.00)		7.00 (5.00 - 10.00)
Mean (± SD)	3.22 (± 2.61)	7.21 (± 3.59)		7.18 (± 3.60)
Ignorado/Sem informação	4	466		470
Quantidade de consulta de pré-natal no 1º trimestre (qtcpn1t2)			0.050 ²	
Median (Q1 - Q3)	1.00 (0.00 - 2.00)	1.00 (1.00 - 2.00)		1.00 (1.00 - 2.00)
Mean (± SD)	1.23 (± 1.15)	1.44 (± 1.12)		1.43 (± 1.12)
Ignorado/Sem informação	6	777		783
Quantidade de consulta de pré-natal no 2º trimestre (Qtcpn2t2)			<0.001 ²	
Median (Q1 - Q3)	2.00 (1.00 - 3.00)	3.00 (2.00 - 3.00)		3.00 (2.00 - 3.00)
Mean (± SD)	1.98 (± 1.36)	2.64 (± 1.02)		2.63 (± 1.03)
Ignorado/Sem Informação	9	908		917
Renda per capita (Renda_Pc)			0.067 ²	
Median (Q1 - Q3)	377.50 (250.67 - 600.00)	450.00 (255.00 - 833.33)		450.00 (255.00 - 833.33)
Mean (± SD)	504.11 (± 404.54)	747.16 (± 976.56)		745.51 (± 974.01)

Característica/código	NPT extremo N = 107	Não NPT extremo N = 14.247	p-valor	Total N = 14.354
Ignorado/Sem Informação	26	2,422		2,448
Toxoplasmose na gestação (Toxogest)			>0.9 ¹	
Não	105 (100%)	14,057 (99%)		14,162 (99%)
Sim	0 (0%)	97 (0.7%)		97 (0.7%)
Ignorado/Sem Informação	2	125		127

Fonte: Elaborado pelos autores, 2026

¹ Teste Exato de Fisher (variáveis categóricas); ² Teste de soma de postos de Wilcoxon (variáveis quantitativas ou ordinais)

Dentre os algoritmos avaliados nos seis cenários, três incluindo a variável ameaça de parto prematuro na gestação atual (SL-RP, SL e RP) e três excluindo (SL-RP, SL e RP), observa-se que os modelos variaram na capacidade de identificar corretamente o NPT extremo, especialmente devido ao desbalanceamento dos dados. No cenário incluindo a variável (tabela 2), SL-RP o *knn* teve melhor desempenho na identificação de NPT extremo (sensibilidade=0.774; especificidade=0.767; AUC_ROC=0.806; F1-score=0.049), indicando boa discriminação entre casos e controles. O *plr* veio em seguida (sensibilidade=0.719; especificidade=0.842; AUC_ROC=0.801; F1-score=0.064). O *nnet* teve menor capacidade discriminativa (sensibilidade=0.719; especificidade=0.788; AUC_ROC=0.777), embora tenha apresentado igual sensibilidade ao *plr*. Em SL, o *svmRadial* foi o mais eficaz para detectar NPT extremo (sensibilidade=0.653; especificidade=0.794; AUC_ROC=0.732; F1-score=0.063). O *cforest* manteve bom desempenho (sensibilidade=0.6; especificidade=0.835; AUC_ROC=0.729; F1-score=0.066), enquanto o *nnet* apresentou resultados inferiores (sensibilidade=0.638; especificidade=0.66; AUC_ROC=0.666; F1-score=0.035). Já em RP, o *ranger* apresentou a melhor discriminação de NPT extremo (AUC_ROC=0.860; sensibilidade=0.771; especificidade=0.830; F1-score=0.053), seguido do *svmRadial* (sensibilidade=0.816; especificidade=0.805; AUC_ROC=0.854; F1-score=0.048). Já o *kkn* teve pior desempenho na identificação de NPT extremo (sensibilidade=0.773; AUC_ROC=0.787; F1-score=0.033).

Tabela 2. Comparação do desempenho dos modelos de aprendizado de máquina utilizados na predição do NPT extremo nas Coortes Brisa de 2010, São Luís/MA e Ribeirão Preto/SP, incluindo a variável ampt.

Modelos	Acurácia	Sensibilidade	Especificidade	Precisão	F1	AUC
Cenário 1: SL-RP						
AdaBoost	0.822 [0.747,	0.683 [0.652,	0.823 [0.748,	0.031 [0.024,	0.059 [0.046,	0.783 [0.768,
.M1	0.862]	0.721]	0.864]	0.035]	0.067]	0.806]
	0.853 [0.838,	0.71 [0.655,	0.854 [0.839,	0.036 [0.033,	0.068 [0.062,	0.798 [0.776,
cforest	0.87]	0.766]	0.872]	0.038]	0.073]	0.82]

ctree2	0.818 [0.758, 0.865]	0.692 [0.652, 0.749]	0.818 [0.758, 0.867]	0.03 [0.023, 0.037]	0.058 [0.044, 0.069]	0.777 [0.752, 0.805]
gam	0.831 [0.791, 0.86]	0.71 [0.67, 0.755]	0.832 [0.791, 0.861]	0.032 [0.026, 0.037]	0.061 [0.051, 0.071]	0.795 [0.763, 0.826]
glmboost	0.857 [0.834, 0.874]	0.701 [0.655, 0.752]	0.858 [0.835, 0.875]	0.036 [0.032, 0.04]	0.069 [0.061, 0.075]	0.799 [0.773, 0.829]
kknn	0.745 [0.716, 0.779]	0.765 [0.684, 0.851]	0.745 [0.715, 0.779]	0.022 [0.02, 0.025]	0.043 [0.038, 0.049]	0.799 [0.762, 0.84]
knn	0.767 [0.728, 0.821]	0.774 [0.705, 0.855]	0.767 [0.728, 0.822]	0.026 [0.021, 0.031]	0.049 [0.042, 0.059]	0.806 [0.781, 0.834]
nnet	0.788 [0.747, 0.828]	0.719 [0.67, 0.774]	0.788 [0.747, 0.829]	0.026 [0.022, 0.03]	0.05 [0.042, 0.057]	0.777 [0.739, 0.805]
plr	0.841 [0.812, 0.86]	0.719 [0.67, 0.769]	0.842 [0.812, 0.861]	0.034 [0.03, 0.037]	0.064 [0.057, 0.071]	0.801 [0.773, 0.829]
ranger	0.825 [0.794, 0.856]	0.683 [0.635, 0.746]	0.826 [0.795, 0.858]	0.029 [0.025, 0.033]	0.056 [0.048, 0.063]	0.788 [0.765, 0.811]
Rborist	0.771 [0.741, 0.813]	0.683 [0.635, 0.746]	0.772 [0.742, 0.814]	0.023 [0.02, 0.026]	0.044 [0.039, 0.051]	0.786 [0.765, 0.806]
rf	0.844 [0.819, 0.869]	0.664 [0.597, 0.737]	0.845 [0.82, 0.871]	0.032 [0.029, 0.036]	0.06 [0.055, 0.068]	0.788 [0.762, 0.811]
svmLinea r	0.865 [0.858, 0.874]	0.692 [0.652, 0.749]	0.866 [0.859, 0.876]	0.037 [0.035, 0.04]	0.071 [0.066, 0.076]	0.779 [0.737, 0.82]
svmRadia l	0.813 [0.79, 0.838]	0.719 [0.679, 0.765]	0.813 [0.79, 0.839]	0.028 [0.026, 0.031]	0.055 [0.051, 0.059]	0.791 [0.772, 0.804]

Cenário 2: SL

AdaBoost .M1	0.757 [0.697, 0.818]	0.635 [0.539, 0.705]	0.758 [0.698, 0.819]	0.028 [0.018, 0.039]	0.054 [0.035, 0.074]	0.693 [0.603, 0.754]
cforest	0.833 [0.805, 0.859]	0.6 [0.473, 0.7]	0.835 [0.806, 0.861]	0.035 [0.025, 0.044]	0.066 [0.047, 0.083]	0.729 [0.648, 0.803]
ctree2	0.784 [0.612, 0.872]	0.636 [0.579, 0.694]	0.786 [0.612, 0.875]	0.039 [0.024, 0.049]	0.073 [0.046, 0.092]	0.723 [0.662, 0.774]
gam	0.749 [0.72, 0.799]	0.633 [0.495, 0.733]	0.75 [0.72, 0.801]	0.025 [0.017, 0.032]	0.048 [0.033, 0.061]	0.726 [0.634, 0.791]
glmboost	0.815 [0.768, 0.851]	0.635 [0.539, 0.705]	0.817 [0.77, 0.852]	0.034 [0.025, 0.043]	0.065 [0.047, 0.081]	0.728 [0.641, 0.789]
kknn	0.617 [0.571, 0.657]	0.62 [0.545, 0.7]	0.617 [0.57, 0.658]	0.015 [0.013, 0.018]	0.03 [0.026, 0.035]	0.671 [0.614, 0.741]
knn	0.779 [0.724, 0.82]	0.635 [0.539, 0.705]	0.78 [0.724, 0.821]	0.028 [0.022, 0.034]	0.054 [0.043, 0.064]	0.726 [0.669, 0.781]
nnet	0.66 [0.596, 0.715]	0.638 [0.485, 0.795]	0.66 [0.595, 0.716]	0.018 [0.013, 0.023]	0.035 [0.026, 0.045]	0.666 [0.553, 0.746]
plr	0.764 [0.728, 0.807]	0.635 [0.497, 0.724]	0.765 [0.73, 0.809]	0.027 [0.019, 0.035]	0.051 [0.036, 0.066]	0.718 [0.628, 0.778]
ranger	0.718 [0.677, 0.745]	0.636 [0.485, 0.761]	0.719 [0.677, 0.746]	0.021 [0.016, 0.026]	0.041 [0.032, 0.05]	0.71 [0.64, 0.767]
Rborist	0.694 [0.677, 0.714]	0.62 [0.515, 0.724]	0.694 [0.677, 0.715]	0.019 [0.015, 0.023]	0.037 [0.03, 0.044]	0.703 [0.641, 0.761]
rf	0.761 [0.739, 0.784]	0.617 [0.485, 0.705]	0.763 [0.74, 0.786]	0.025 [0.019, 0.028]	0.047 [0.037, 0.054]	0.716 [0.64, 0.776]
svmLinea r	0.818 [0.748, 0.871]	0.582 [0.468, 0.683]	0.821 [0.749, 0.874]	0.034 [0.022, 0.047]	0.065 [0.042, 0.087]	0.722 [0.611, 0.833]
svmRadia l	0.793 [0.746, 0.839]	0.653 [0.552, 0.724]	0.794 [0.747, 0.84]	0.033 [0.022, 0.046]	0.063 [0.042, 0.085]	0.732 [0.656, 0.795]

Cenário 3: RP

AdaBoost .M1	0.781 [0.716, 0.845]	0.756 [0.656, 0.88]	0.781 [0.715, 0.846]	0.022 [0.017, 0.027]	0.042 [0.033, 0.051]	0.806 [0.754, 0.869]
cforest	0.856 [0.846, 0.867]	0.796 [0.676, 0.92]	0.857 [0.845, 0.867]	0.032 [0.028, 0.038]	0.062 [0.054, 0.074]	0.831 [0.772, 0.9]

ctree2	0.864 [0.858, 0.87]	0.796 [0.676, 0.92]	0.864 [0.858, 0.87]	0.034 [0.028, 0.04]	0.065 [0.055, 0.077]	0.838 [0.779, 0.896]
gam	0.783 [0.743, 0.822]	0.776 [0.676, 0.876]	0.783 [0.742, 0.822]	0.021 [0.019, 0.024]	0.041 [0.037, 0.046]	0.805 [0.767, 0.847]
glmboost	0.852 [0.828, 0.872]	0.776 [0.676, 0.876]	0.853 [0.829, 0.873]	0.031 [0.026, 0.038]	0.06 [0.049, 0.073]	0.817 [0.751, 0.888]
kknn	0.719 [0.68, 0.771]	0.773 [0.687, 0.86]	0.718 [0.679, 0.771]	0.017 [0.013, 0.021]	0.033 [0.026, 0.04]	0.787 [0.712, 0.862]
knn	0.822 [0.803, 0.839]	0.816 [0.696, 0.935]	0.822 [0.803, 0.839]	0.027 [0.024, 0.03]	0.052 [0.046, 0.059]	0.843 [0.769, 0.908]
nnet	0.717 [0.646, 0.798]	0.753 [0.654, 0.84]	0.717 [0.645, 0.799]	0.017 [0.014, 0.022]	0.033 [0.028, 0.043]	0.791 [0.758, 0.822]
plr	0.801 [0.751, 0.846]	0.733 [0.647, 0.82]	0.802 [0.751, 0.847]	0.023 [0.018, 0.029]	0.044 [0.035, 0.056]	0.819 [0.77, 0.873]
ranger	0.83 [0.805, 0.855]	0.771 [0.642, 0.92]	0.83 [0.805, 0.855]	0.028 [0.02, 0.037]	0.053 [0.039, 0.071]	0.86 [0.82, 0.906]
Rborist	0.795 [0.742, 0.834]	0.751 [0.611, 0.9]	0.795 [0.741, 0.835]	0.022 [0.019, 0.024]	0.042 [0.037, 0.047]	0.847 [0.817, 0.882]
rf	0.806 [0.757, 0.842]	0.771 [0.642, 0.92]	0.806 [0.756, 0.843]	0.024 [0.02, 0.027]	0.046 [0.039, 0.052]	0.836 [0.788, 0.883]
svmLinea r	0.864 [0.858, 0.87]	0.796 [0.676, 0.92]	0.864 [0.858, 0.87]	0.034 [0.028, 0.04]	0.065 [0.055, 0.077]	0.83 [0.77, 0.892]
svmRadial l	0.805 [0.778, 0.833]	0.816 [0.696, 0.935]	0.805 [0.777, 0.834]	0.025 [0.021, 0.031]	0.048 [0.041, 0.06]	0.854 [0.798, 0.907]

Legenda: *Foram avaliados algoritmos como Regressão logística penalizada(*plr*), máquinas de vetor de suporte (*svmLinear* e *svmRadial*), Florestas aleatórias (*rf*, *ranger* e *Rborist*), Boosting (*glmboost* e *AdaBoost.M1*), Métodos baseados em vizinhança (*Knn* e *KKnn*), Redes neurais artificiais (*nnet*), Árvores e modelos baseados em partição (*cforest* e *ctree2*) e Modelos aditivos generalizados(*gam*).com base em métricas como Precisão (VPP), sensibilidade, Valor preditivo negativo (VPN), sensibilidade, especificidade, F1-score, AUC_ROC.

No cenário excluindo a variável (Tabela 3), SL-RP, o *nnet* teve melhor desempenho na identificação de NPT extremo (sensibilidade=0.561; especificidade=0.577; AUC_ROC=0.648; F1-score=0.019). O *knn* veio em seguida (sensibilidade=0.624; especificidade=0.597; AUC_ROC=0.637; F1-score=0.022). O *svmRadial* teve menor capacidade discriminativa (sensibilidade=0.494; especificidade=0.62; AUC_ROC=0.581). Em SL, o *gam* foi o mais eficaz para detectar NPT extremo (sensibilidade=0.623; especificidade=0.534; AUC_ROC=0.587; F1-score=0.024). O *cforest* obteve bom desempenho (sensibilidade=0.588; especificidade=0.541; AUC_ROC=0.571; F1-score=0.024), enquanto o *svmLinear* apresentou resultados inferiores (sensibilidade=0.6388; especificidade=0.333; AUC_ROC=0.489; F1-score=0.019). Já em RP, o *Rborist* apresentou a melhor discriminação de NPT extremo (AUC_ROC=0.629; sensibilidade=0.573; especificidade=0.594; F1-score=0.016), seguido do *plr* (sensibilidade=0.531; especificidade=0.583; AUC_ROC=0.619; F1-score=0.015). Já o *AdaBoost.M1* teve pior desempenho na identificação de NPT extremo (sensibilidade=0.51; AUC_ROC=0.554; F1-score=0.013).

Tabela 3. Comparação do desempenho dos modelos de aprendizado de máquina utilizados na predição do NPT extremo nas Coortes Brisa de 2010, São Luís/MA e Ribeirão Preto/SP, excluindo a variável **ampt**.

Modelos	Acurácia	Sensibilidade	Especificidade	Precisão	F1	AUC
Cenário 1: SL-RP						
AdaBoost.M1	0.582 [0.534, 0.622]	0.587 [0.505, 0.705]	0.582 [0.533, 0.623]	0.01 [0.009, 0.011]	0.02 [0.018, 0.022]	0.627 [0.588, 0.68]
cforest	0.618 [0.577, 0.674]	0.543 [0.365, 0.683]	0.619 [0.577, 0.676]	0.01 [0.008, 0.012]	0.02 [0.016, 0.024]	0.628 [0.584, 0.692]
ctree2	0.546 [0.379, 0.713]	0.599 [0.358, 0.84]	0.546 [0.376, 0.715]	0.01 [0.008, 0.013]	0.02 [0.016, 0.025]	0.604 [0.539, 0.669]
gam	0.61 [0.591, 0.631]	0.514 [0.373, 0.664]	0.61 [0.591, 0.633]	0.01 [0.007, 0.012]	0.019 [0.014, 0.024]	0.622 [0.578, 0.678]
glmboost	0.609 [0.573, 0.66]	0.514 [0.346, 0.673]	0.61 [0.572, 0.663]	0.009 [0.007, 0.012]	0.019 [0.014, 0.023]	0.625 [0.583, 0.674]
kknn	0.561 [0.534, 0.585]	0.606 [0.533, 0.687]	0.561 [0.533, 0.585]	0.01 [0.009, 0.011]	0.02 [0.018, 0.022]	0.61 [0.577, 0.653]
knn	0.598 [0.562, 0.633]	0.624 [0.533, 0.756]	0.597 [0.561, 0.634]	0.011 [0.01, 0.013]	0.022 [0.02, 0.026]	0.637 [0.601, 0.685]
nnet	0.577 [0.524, 0.634]	0.561 [0.409, 0.709]	0.577 [0.523, 0.635]	0.01 [0.008, 0.011]	0.019 [0.016, 0.022]	0.648 [0.601, 0.702]
plr	0.606 [0.577, 0.645]	0.524 [0.383, 0.655]	0.607 [0.576, 0.646]	0.01 [0.008, 0.012]	0.019 [0.015, 0.023]	0.625 [0.586, 0.669]
ranger	0.621 [0.584, 0.657]	0.493 [0.367, 0.627]	0.622 [0.584, 0.658]	0.01 [0.007, 0.012]	0.019 [0.015, 0.024]	0.628 [0.595, 0.674]
Rborist	0.582 [0.565, 0.599]	0.577 [0.48, 0.686]	0.582 [0.564, 0.6]	0.01 [0.009, 0.012]	0.02 [0.017, 0.023]	0.609 [0.556, 0.674]
rf	0.626 [0.579, 0.674]	0.541 [0.406, 0.671]	0.627 [0.578, 0.676]	0.011 [0.009, 0.013]	0.021 [0.017, 0.025]	0.625 [0.579, 0.683]
svmLinear	0.561 [0.443, 0.68]	0.551 [0.334, 0.774]	0.561 [0.44, 0.683]	0.009 [0.007, 0.01]	0.017 [0.015, 0.02]	0.627 [0.566, 0.693]
svmRadial	0.619 [0.569, 0.683]	0.494 [0.365, 0.635]	0.62 [0.568, 0.685]	0.01 [0.008, 0.011]	0.019 [0.015, 0.022]	0.581 [0.518, 0.653]
Cenário 2: SL						
AdaBoost.M1	0.493 [0.447, 0.534]	0.586 [0.535, 0.638]	0.492 [0.446, 0.534]	0.011 [0.01, 0.012]	0.022 [0.019, 0.024]	0.546 [0.514, 0.578]
cforest	0.541 [0.499, 0.596]	0.588 [0.494, 0.679]	0.541 [0.499, 0.597]	0.012 [0.011, 0.013]	0.024 [0.021, 0.026]	0.571 [0.521, 0.613]
ctree2	0.379 [0.118, 0.667]	0.724 [0.409, 0.982]	0.376 [0.11, 0.67]	0.013 [0.01, 0.018]	0.025 [0.019, 0.033]	0.549 [0.48, 0.631]
gam	0.535 [0.485, 0.601]	0.623 [0.467, 0.761]	0.534 [0.483, 0.602]	0.012 [0.01, 0.015]	0.024 [0.02, 0.029]	0.587 [0.488, 0.666]
glmboost	0.494 [0.411, 0.583]	0.608 [0.45, 0.779]	0.493 [0.408, 0.584]	0.011 [0.01, 0.012]	0.022 [0.019, 0.024]	0.563 [0.49, 0.633]
kknn	0.493 [0.445, 0.525]	0.591 [0.467, 0.715]	0.492 [0.442, 0.525]	0.011 [0.009, 0.013]	0.021 [0.018, 0.025]	0.564 [0.524, 0.604]
knn	0.48 [0.449, 0.504]	0.588 [0.476, 0.7]	0.479 [0.447, 0.504]	0.011 [0.009, 0.013]	0.021 [0.017, 0.025]	0.555 [0.5, 0.627]
nnet	0.529 [0.425, 0.641]	0.505 [0.324, 0.691]	0.53 [0.424, 0.644]	0.01 [0.007, 0.014]	0.02 [0.014, 0.028]	0.546 [0.458, 0.636]
plr	0.491 [0.407, 0.577]	0.589 [0.43, 0.77]	0.49 [0.403, 0.579]	0.011 [0.009, 0.012]	0.021 [0.018, 0.024]	0.57 [0.497, 0.635]
ranger	0.518 [0.494, 0.542]	0.641 [0.535, 0.779]	0.517 [0.491, 0.542]	0.012 [0.011, 0.014]	0.024 [0.021, 0.028]	0.561 [0.515, 0.608]
Rborist	0.494 [0.462, 0.527]	0.589 [0.476, 0.709]	0.493 [0.46, 0.527]	0.011 [0.01, 0.012]	0.021 [0.019, 0.024]	0.557 [0.515, 0.599]

rf	0.493 [0.467, 0.526]	0.586 [0.505, 0.661]	0.492 [0.465, 0.526]	0.011 [0.01, 0.012]	0.021 [0.019, 0.023]	0.549 [0.513, 0.587]
svmLinea r	0.336 [0.191, 0.456]	0.688 [0.556, 0.811]	0.333 [0.185, 0.455]	0.01 [0.009, 0.011]	0.019 [0.017, 0.022]	0.489 [0.429, 0.574]
svmRadia l	0.585 [0.45, 0.739]	0.405 [0.25, 0.595]	0.586 [0.449, 0.743]	0.01 [0.007, 0.013]	0.019 [0.014, 0.025]	0.497 [0.427, 0.582]
Cenário 3: RP						
AdaBoost .M1	0.584 [0.534, 0.631]	0.451 [0.32, 0.56]	0.585 [0.535, 0.632]	0.007 [0.004, 0.009]	0.013 [0.009, 0.017]	0.554 [0.462, 0.629]
cforest	0.599 [0.562, 0.635]	0.489 [0.409, 0.569]	0.6 [0.563, 0.636]	0.007 [0.006, 0.009]	0.015 [0.011, 0.018]	0.586 [0.531, 0.638]
ctree2	0.59 [0.462, 0.702]	0.489 [0.329, 0.629]	0.591 [0.461, 0.703]	0.007 [0.005, 0.009]	0.014 [0.01, 0.018]	0.582 [0.561, 0.603]
gam	0.568 [0.514, 0.622]	0.551 [0.5, 0.631]	0.568 [0.513, 0.623]	0.008 [0.007, 0.009]	0.015 [0.013, 0.017]	0.602 [0.566, 0.65]
glmboost	0.589 [0.53, 0.647]	0.531 [0.5, 0.571]	0.589 [0.53, 0.648]	0.008 [0.007, 0.009]	0.015 [0.013, 0.018]	0.613 [0.577, 0.652]
kkn	0.569 [0.519, 0.611]	0.571 [0.533, 0.6]	0.569 [0.518, 0.611]	0.008 [0.007, 0.009]	0.016 [0.013, 0.018]	0.588 [0.566, 0.607]
knn	0.535 [0.51, 0.561]	0.613 [0.54, 0.667]	0.534 [0.51, 0.561]	0.008 [0.007, 0.009]	0.015 [0.014, 0.017]	0.6 [0.578, 0.627]
nnet	0.604 [0.56, 0.647]	0.491 [0.32, 0.631]	0.604 [0.56, 0.649]	0.007 [0.005, 0.01]	0.014 [0.01, 0.019]	0.611 [0.53, 0.691]
plr	0.582 [0.526, 0.639]	0.531 [0.5, 0.571]	0.583 [0.526, 0.639]	0.008 [0.007, 0.009]	0.015 [0.013, 0.017]	0.619 [0.584, 0.658]
ranger	0.59 [0.565, 0.611]	0.511 [0.402, 0.62]	0.59 [0.565, 0.611]	0.007 [0.006, 0.01]	0.015 [0.011, 0.019]	0.603 [0.538, 0.667]
Rborist	0.594 [0.562, 0.621]	0.573 [0.52, 0.627]	0.594 [0.562, 0.622]	0.008 [0.008, 0.009]	0.016 [0.015, 0.018]	0.629 [0.592, 0.666]
rf	0.616 [0.586, 0.644]	0.511 [0.4, 0.591]	0.616 [0.586, 0.645]	0.008 [0.006, 0.01]	0.016 [0.012, 0.019]	0.603 [0.542, 0.664]
svmLinea r	0.615 [0.515, 0.711]	0.489 [0.438, 0.56]	0.616 [0.515, 0.713]	0.008 [0.006, 0.01]	0.016 [0.012, 0.02]	0.594 [0.534, 0.649]
svmRadia l	0.6 [0.51, 0.69]	0.553 [0.46, 0.627]	0.6 [0.51, 0.691]	0.009 [0.006, 0.011]	0.017 [0.013, 0.022]	0.595 [0.537, 0.658]

Legenda: *Foram avaliados algoritmos como Regressão logística penalizada(*plr*), máquinas de vetor de suporte (*svmLinear* e *svmRadial*), Florestas aleatórias (*rf*, *ranger* e *Rborist*), Boosting (*glmboost* e *AdaBoost.M1*), Métodos baseados em vizinhança (*Knn* e *KKnn*), Redes neurais artificiais (*nnet*), Árvores e modelos baseados em partição (*cforest* e *ctree2*) e Modelos aditivos generalizados(*gam*).com base em métricas como Precisão (VPP), Valor preditivo negativo (VPN), sensibilidade, especificidade, F1-score, AUC_ROC.

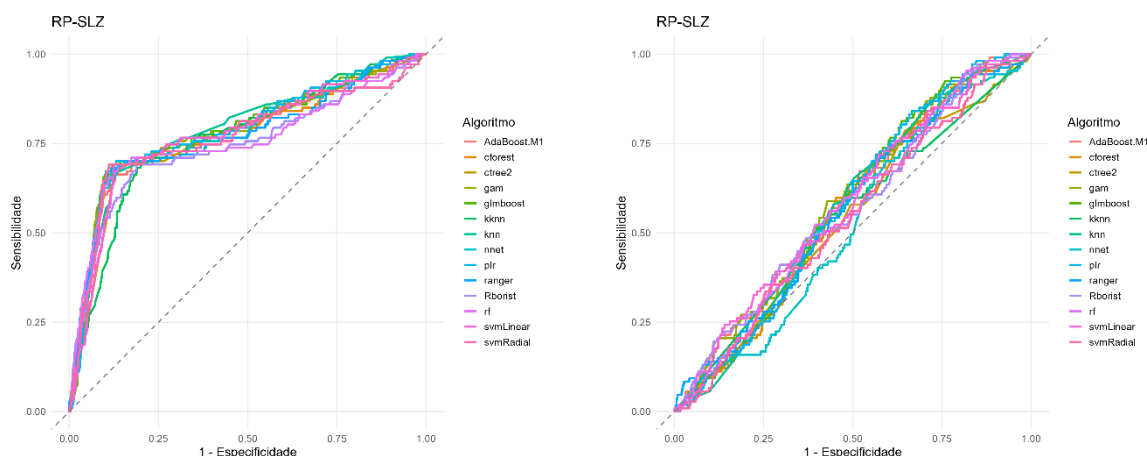
De modo geral, os modelos apresentaram variação em relação à precisão (valor preditivo positivo-VPP) (~0.007 a 0.036) e sensibilidade (~0.405 a 0.816), destacando o desafio inerente à predição de NPT extremo, especialmente em contextos de grande desbalanceamento entre as classes.

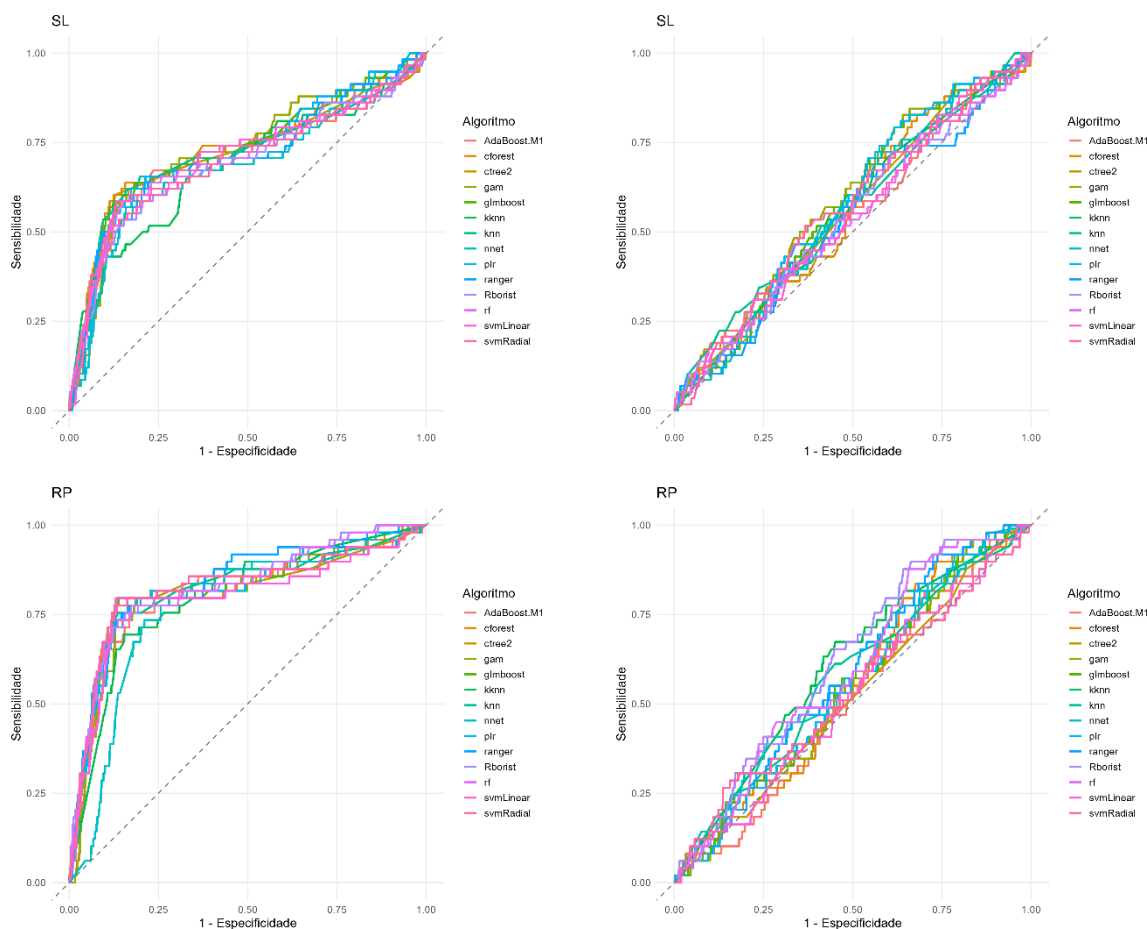
Portanto, nos seis cenários avaliados: três incluindo a variável ameaça de parto prematuro na gestação atual (SL-RP, SL e RP) e três excluindo (SL-RP, SL e RP), o desempenho dos modelos variou conforme as métricas, com destaque para *knn* ao incluir a variável e o *nnet* ao excluir a variável, apresentaram os melhores resultados em AUC-ROC e

sensibilidade, indicando maior capacidade de discriminar casos de nascimento pré-termo extremo. O modelo *ranger* também se destacou em alguns cenários, especialmente pela sensibilidade, enquanto o *nnet* ao incluir a variável e o *svmLinear* ao excluir tiveram desempenho inferior em AUC-ROC, apesar de variações no F1-score. Em todos os cenários, observou-se o impacto do desbalanceamento dos dados, com variação entre sensibilidade, precisão e F1-score, evidenciando a dificuldade de predição do desfecho de nascimento pré-termo extremo.

As curvas ROC foram obtidas a partir da média das probabilidades preditas em procedimento de validação cruzada. Observa-se que em relação aos modelos, tiveram pouca variação, entre eles o *knn*, *nnet*, *svmRadial*, *ranger*, *gam* e *Rborist* apresentaram melhores desempenho discriminatório, com curvas mais distantes da linha de aleatoriedade (linha cinza diagonal), indicando maior capacidade de distinguir entre os casos de NPT extremo e os não NPT extremo. Dentre eles, os modelos *nnet* e *gam* demonstraram as curvas mais favoráveis, sugerindo maior acurácia global. Em contrapartida, os modelos *kknn* e *AdaBoost.M1* apresentaram desempenho inferior, com curvas próximas da diagonal, o que indica fraca capacidade de classificação. Portanto, esses resultados apontam para a superioridade dos modelos *knn* e *nnet* na tarefa de predição do desfecho binário. (**Figura 2**).

Figura 2. Curva ROC dos modelos de aprendizado de máquina utilizados na predição do NPT extremo nas Coortes Brisa de 2010, São Luís/MA e Ribeirão Preto/SP, com “ampt” (esquerda) e sem “ampt” (direita).





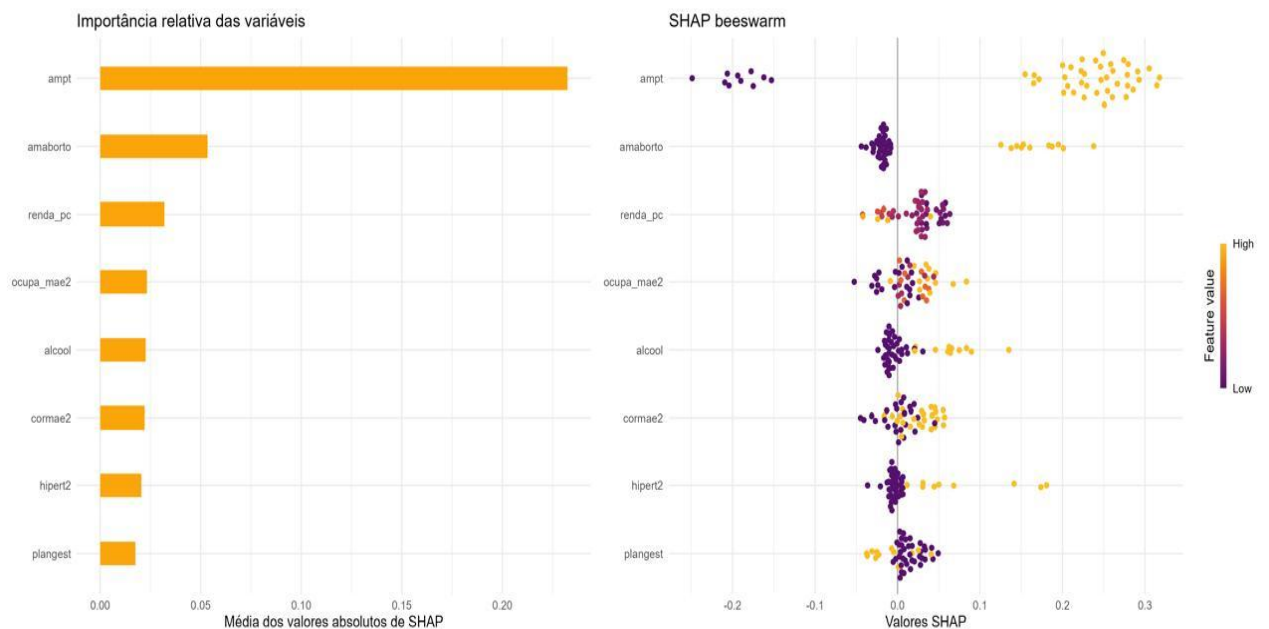
Fonte: Elaborado pelos autores, 2026

As Figuras 3 e 4 mostram a importância dos preditores para o NPT extremo, usando valores de Shap, técnica de interpretabilidade em aprendizado de máquina que explica a influência de cada variável nas previsões realizadas por um modelo. Para os modelos de melhor desempenho, quando incluindo a variável ameaça de parto prematuro na gestação atual o *knn*, *svmRadial* e *ranger* tiveram melhor desempenho; e ao excluir a variável, o *nnet*, *gam* e *Rborist* permitiram uma compreensão de como cada variável contribui para a predição, representada em histograma e valores dos atributos, esse último possui uma linha central onde à esquerda apresenta pouca influência ao desfecho e à direita elevada influência. Entre os preditores avaliados, as variáveis dos seis cenários com maior influência foram a ameaça de parto prematuro na gestação atual, ameaça de aborto na gestação atual, renda per capita, ocupação da mãe e cor da mãe indicando que exercem papel central na modelagem do risco. Em seguida, destaca-se gravidez planejada, sugerindo que fatores relacionados a saúde reprodutiva, influenciam significativamente a ocorrência do desfecho. Outras variáveis relevantes incluíram

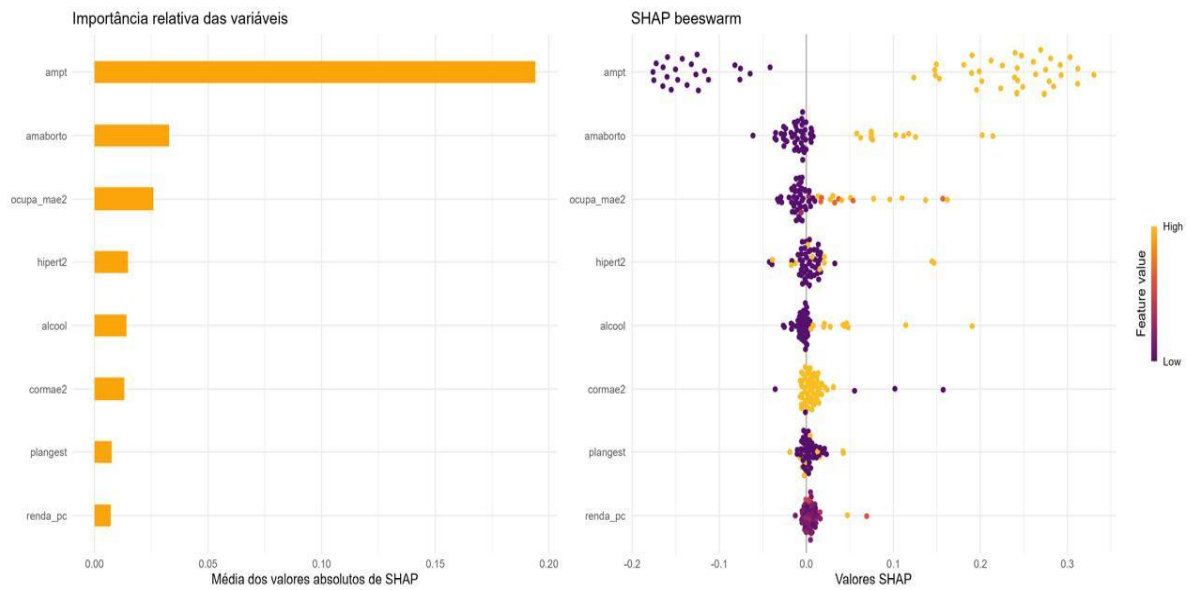
o uso de álcool e hipertensão, esta diagnosticada antes ou durante a gestação, reforçando a contribuição de determinantes comportamentais e saúde materna, bem como características da saúde da mulher, indicando sua influência no perfil de risco. A predominância de fatores clínicos e socioeconômicos sobre pré-natais e comportamentais sugere que fatores clínicos e a vulnerabilidade social podem ser componentes determinantes na predição do nascimento pré-termo extremo (**Figuras 3 e 4**).

Figura 3. Importância dos preditores de nascimento pré-termo extremo com base nos valores de Shapley incluindo a variável ameaça de parto prematuro na gestação atual (ampt) SL-RP (knn); SL (svmRadial); RP (ranger).

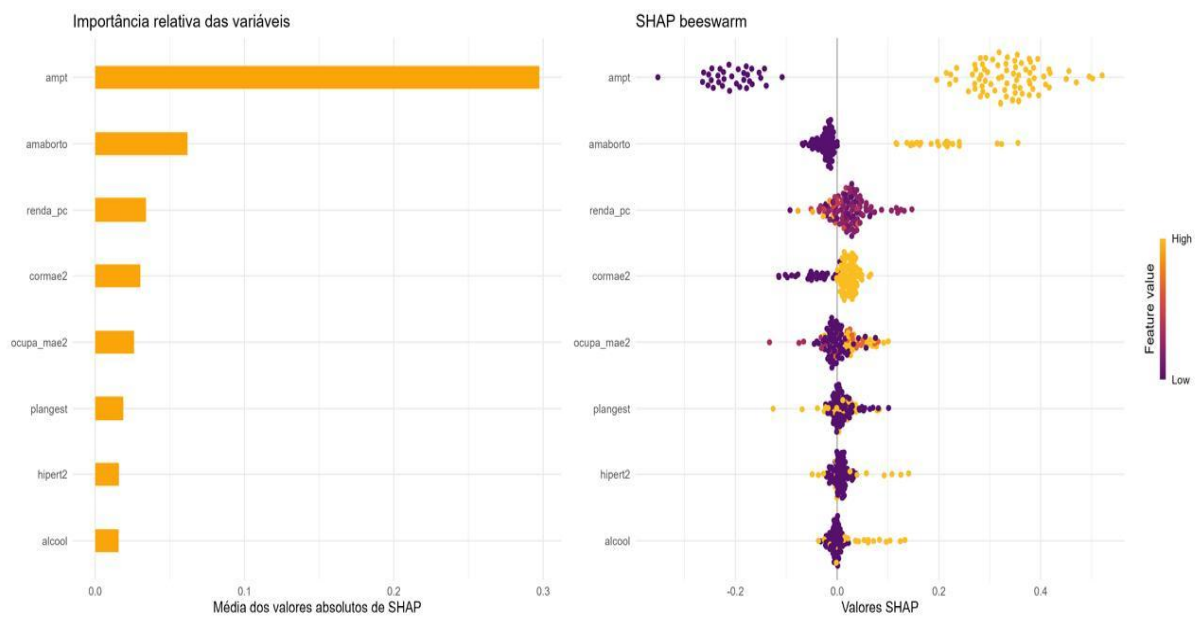
3a) SL-RP



3b) SL



3c) RP

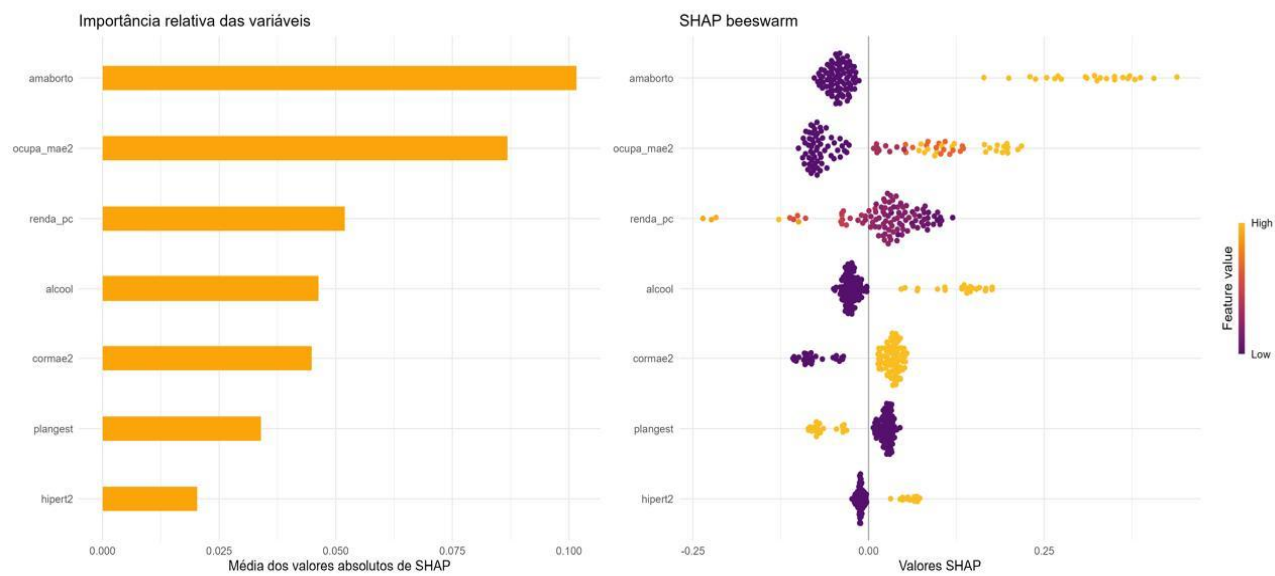


Fonte: Elaborados pelos autores, 2026

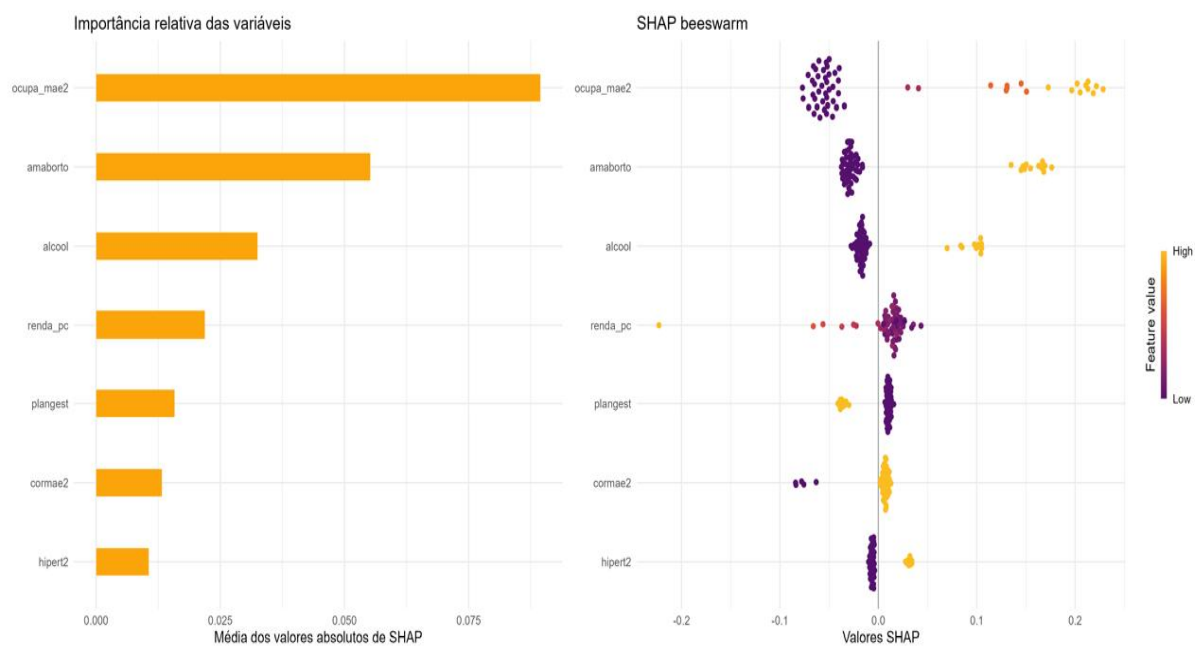
Legenda: **ampt**– ameaça de parto prematuro na gestação atual; **amaborto** – Ameaça de aborto na gestação atual; **ocupa_mae2**–Ocupação da mãe (qualificado não manual; manual semiespecializado; manual não qualificado; desempregado); **renda_pc** – Renda per capita mensal, renda média por pessoa da residência; **plangest** – A gravidez atual foi planejada; **cormae2** – cor da mãe autodeclarada; **alcool**- consumo de álcool na gestação; **hipert2** – Teve diagnóstico de hipertensão arterial sistêmica (HAS) (antes ou durante) a gestação.

Figura 4. Importância dos preditores de nascimento pré-termo extremo com base nos valores de Shapley, excluindo “ampt” SL-RP (nnet); SL (gam); RP (Rborist).

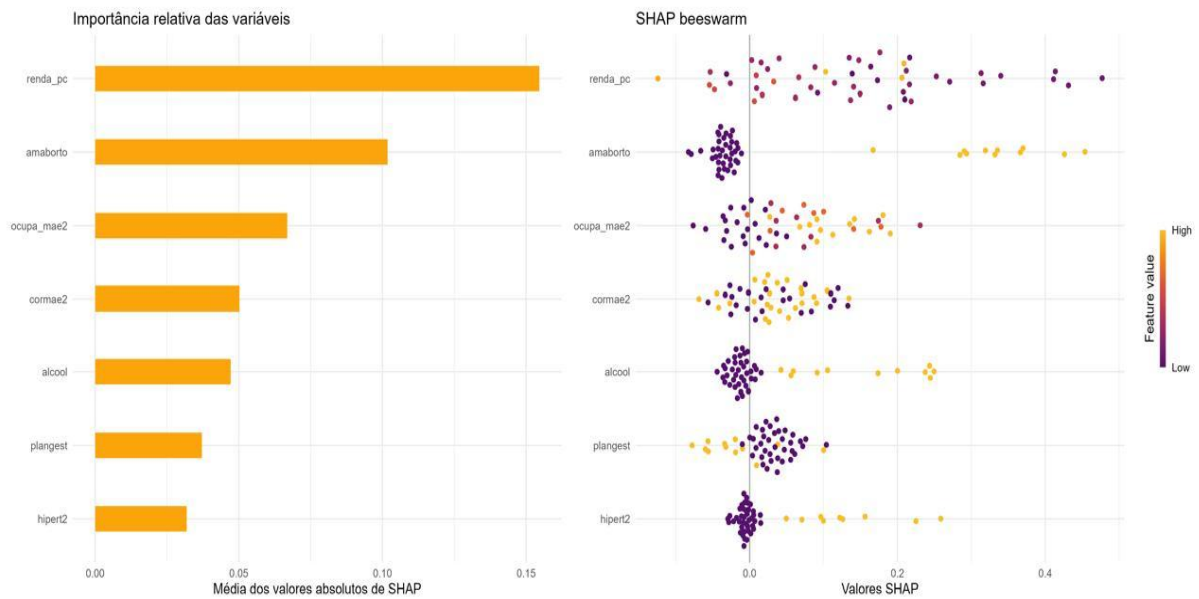
4a) SL-RP



4b) SL



4c) RP



Fonte: Elaborados pelos autores, 2026

Legenda: **amaborto** – Ameaça de aborto na gestação atual; **ocupa_mae2**–Ocupação da mãe (qualificado não manual; manual semiespecializado; manual não qualificado; desempregado); **renda_pc** – Renda per capita mensal, renda média por pessoa da residência; **plangest** – A gravidez atual foi planejada; **cormae2** – cor da mãe autodeclarada; **alcool**- consumo de álcool na gestação; **hipert2** – Teve diagnóstico de hipertensão arterial sistêmica (HAS) (antes ou durante) a gestação.

DISCUSSÃO

Das 14.687 mulheres/criança das coortes, 107 (0,72%) tiveram NPT extremo. Foram testados quatorze algoritmos de aprendizado de máquina — Regressão logística penalizada (*plr*), *Support Vector Machines* (*svmLinear* e *svmRadial*), *Random Forests* (*rf*, *ranger* e *Rborist*), Boosting (*glmboost* e *AdaBoost.M1*), Métodos baseados em vizinhança (*Knn* e *KKnn*), Redes neurais artificiais (*nnet*), Árvores e modelos baseados em partição (*cforest* e *ctree2*) e Modelos aditivos generalizados (*gam*). Entre os modelos testados, nos seis cenários (SL-RP, SL e RP com e sem a variável ameaça de parto prematuro na gestação atual), o desempenho variou — no cenário incluindo a variável, em SL-RP o *knn* apresentou sensibilidade de 0.774, AUC-ROC de 0.806 e F1-score de 0.049; SL, o modelo *svmRadial* obteve sensibilidade de 0.653, AUC-ROC de 0.732 e F1-score de 0.063; RP, o *ranger* apresentou os melhores resultados, com AUC-ROC de 0.860, sensibilidade de 0.771 e F1-score de 0.053 — no cenário excluindo a variável, em SL-RP o *nnet* apresentou sensibilidade de 0.561, AUC-ROC de 0.648 e F1-score de 0.019;

SL, o modelo *gam* obteve sensibilidade de 0.623, AUC-ROC de 0.587 e F1-score de 0.024; RP, o *Rborist* apresentou AUC-ROC de 0.629, sensibilidade de 0.573 e F1-score de 0.016.

Portanto, verifica-se que diante da relevância clínica ao incluirmos a variável ameaça de parto prematuro na gestação atual, obtivemos melhores resultados nas métricas de desempenho dos modelos. O *nnet*, *gam*, *Rborist*, *ranger*, *knn* e *svmRadial*, foram os algoritmos mais eficazes na tarefa de predição de NPT extremo, ao incluir e excluir a variável. Esses modelos também apresentaram um dos maiores valores de F1-score, refletindo um melhor equilíbrio entre precisão e sensibilidade, apresentando boa performance na tarefa de predição de NPT extremo.

Aplicou-se o algoritmo BORUTA para selecionar as variáveis mais importantes como preditoras do desfecho (Kursa; Rudnicki, 2010), resultando em 50 variáveis. Posteriormente elencamos oito variáveis independentes ou preditoras, com base em plausibilidade clínica e evidências epidemiológicas, as quais foram: ameaça de parto prematuro na gestação atual, ameaça de aborto na gestação atual, ocupação da mãe, cor da mãe autodeclarada, renda per capita, gravidez atual planejada, consumo de álcool na gestação, diagnóstico de hipertensão antes ou durante a gestação

Os resultados do presente estudo corroboram as evidências no tocante ao NPT extremo ser um evento raro, porém de alto impacto e com um elevado potencial de risco. Estudos anteriores identificaram também a baixa prevalência e o elevado potencial de risco da prematuridade extrema. Passini et al. (2014) publicaram um estudo multicêntrico brasileiro que investigou nascimentos em 20 hospitais de referência, com o objetivo de estimar a prevalência de partos pré-termo, resultando em 0,91% de NPT extremo dos 12,3% pré-termo. Taylor et al., (2022) publicaram uma revisão de duas grandes coortes de múltiplos centros, onde nascidos prematuros extremos representaram apenas cerca de 0,7% dos nascimentos nos Estados Unidos.

Os achados de Araújo et al. (2023) e Koivu e Sairanen, (2020) também coincidem com nosso estudo. Koivu e Sairanen (2020) desenvolveram um modelo de predição de parto prematuro baseado em aprendizado de máquina utilizando conjuntos de dados contendo gestações ocorridas entre 2013 e 2016 nos Estados Unidos, fornecidos pelo CDC, Centro Nacional de Estatísticas de Saúde, por meio do seu Sistema Nacional de Estatísticas Vitais e utilizou dados demográficos, histórico de gravidez e infecções maternas. Compararam três algoritmos de aprendizado de máquina, destacando-se a Regressão logística (glm) que apresentou o melhor desempenho (AUC = 0,74). Araújo et al. (2023) utilizou dados públicos e

técnicas de aprendizado de máquina para prever prematuridade e baixo peso ao nascer no Brasil. Foram testados três métodos, e o glm apresentou melhor desempenho (acurácia=0,91; F1-score=0,60). Gao et al. (2019) analisaram dados de partos no Vanderbilt University Medical Center, compararam diferentes algoritmos para prever parto prematuro extremo. O modelo de redes neurais recorrentes (RNN), treinado com dados balanceados artificialmente, alcançou o melhor desempenho (AUC = 0,827; sensibilidade = 0,965). Portanto os resultados reforçam que modelos de aprendizado de máquina são eficazes na análise de grandes bases de dados para prever NPT extremo, especialmente em contextos clínicos com desfechos raros e dados desbalanceados, contribuindo para a detecção precoce de riscos à saúde materno-infantil.

O uso dos valores de Shapley em nosso estudo forneceu *insights* sobre a interpretabilidade do modelo, permitindo uma compreensão de como cada variável contribuiu para a predição do NPT extremo, evidenciando que a ameaça de aborto na gestação atual, ameaça de parto prematuro, ocupação da mãe e renda per capita apresentaram maior influência na predição, seguido de cor da mãe autodeclarada, gravidez atual planejada, uso de álcool na gestação e hipertensão antes ou durante a gestação. Estes achados corroboram o que é encontrado na literatura, a exemplo do estudo de Yu QY et al. (2024), na China que prever parto prematuro (<37 semanas de gestação) em gestações únicas com diferentes intervalos gestacionais, provenientes de um estudo de coorte prospectivo baseado na população, realizado em 51 clínicas de obstetrícia e hospitais na cidade de Wenzhou, identificou que a adição de características fortemente associadas ao parto prematuro, como histórico de parto prematuro anterior, síndromes de hipertensão gestacional, diabetes gestacional e medidas ultrassonográficas, poderá levar a uma melhor avaliação do risco de parto prematuro durante a gravidez atual. Adhikari et al. (2019), no Canadá, desenvolveram modelo preditivo para NPT através do status sócio-econômico do bairro e identificaram dentre os fatores preditores, a renda familiar, educação, status de emprego, índice de massa corporal, consumo de fumo, renda pessoal mediana do bairro e acesso ao pré-natal, reforçando a importância da identificação precoce dessas variáveis que predizem gestações de risco, principalmente em contextos com recursos limitados.

Fowler et al. (2025) publicaram um artigo relativo a cuidados pré-parto no segundo e terceiro trimestres, reforçando a importância das consultas de pré-natal com profissionais médicos e enfermeiros, especialmente no segundo trimestre (idealmente a cada 4 semanas), quando são monitorados pressão arterial, peso, tamanho uterino, frequência cardíaca fetal, anemia, infecções sexualmente transmissíveis, rastreio de anomalias congênitas, para

identificar precocemente as alterações e permitir intervenções oportunas, reduzindo os riscos de parto prematuro.

Diante desse contexto, destacam-se importantes fortalezas deste estudo. Nossas descobertas são significativas e inovadoras por utilizarem modelos de aprendizado de máquina capazes de identificar, com precisão, variáveis preditoras de um evento raro, o NPT extremo, mesmo em bases de dados desbalanceadas e provenientes de cidades com diferentes contextos demográficos e socioeconômicos. Além disso, o modelo foi desenvolvido com apenas oito variáveis preditoras, todas de fácil coleta, baixo custo e disponíveis nos sistemas de informação do SUS, o que amplia sua aplicabilidade e viabiliza sua implementação na prática clínica e em serviços de saúde pública.

Esses achados podem ser particularmente úteis aos profissionais de saúde na identificação precoce de gestantes com maior risco para NPT extremo, favorecendo intervenções oportunas, como o reconhecimento antecipado de sintomas clínicos — ameaça de parto prematuro e abortamento na gestação atual — e o aconselhamento direcionado para modificações no estilo de vida. Tal abordagem é especialmente relevante em países de baixa e média renda, onde as taxas de prematuridade são mais elevadas e os recursos em saúde frequentemente limitados (Hassan et al., 2025).

Apesar dos resultados promissores, este estudo apresenta limitações que devem ser reconhecidas. As coortes analisadas possuem desequilíbrio entre as classes da base de dados, devido à baixa frequência do desfecho em relação ao número de observações. Esse desequilíbrio entre casos e não casos pode afetar o desempenho dos modelos, apesar da utilização de técnicas adequadas para o equilíbrio, reduzindo a capacidade de identificar corretamente os nascimentos pré-termo extremos. Além disso, possíveis vieses de informação e perdas de seguimento podem influenciar os resultados. Portanto, investigações futuras são necessárias para validar os modelos propostos com diferentes proporções de bases de dados em diferentes contextos geográficos e demográficos, a fim de confirmar sua confiabilidade e capacidade de generalização. Embora tenham sido utilizado dados de uma coorte, para este estudo foram incluídos somente dados do baseline, portanto, com recorte transversal. Todavia, as variáveis incluídas nos modelos foram somente aquelas relativas ao tempo anterior ao desfecho. Ademais, embora os modelos tenham demonstrado uma boa acurácia preditiva, sua

incorporação em fluxos de trabalho clínicos requer avaliações adicionais e esforços conjuntos da academia e serviços de saúde.

Estudos subsequentes devem examinar a viabilidade e o impacto clínico dessa integração, incluindo a aplicação combinada com tecnologias de saúde móvel (“mHealth”) como estratégia para otimizar a atenção pré-natal, especialmente em cenários com recursos limitados. Ademais, embora tenham sido utilizadas técnicas consistentes, os modelos ainda necessitam de validação em contextos clínicos reais. A aplicação prática de modelos de aprendizado de máquina em saúde enfrenta desafios importantes, incluindo questões de privacidade de dados, imparcialidade e mitigação de vieses, que devem ser cuidadosamente abordados antes de sua ampla adoção.

CONCLUSÃO

É possível prever o nascimento pré-termo extremo com boa performance, a partir de variáveis socioeconômicas, de acesso ao pré-natal, de saúde materna e paterna e de características fetais. A identificação precoce de gestantes e recém-nascidos em risco pode contribuir para intervenções preventivas mais eficazes e melhores desfechos maternos e neonatais. Os resultados deste estudo reforçam o potencial dos modelos preditivos baseados em aprendizado de máquina e indicam a necessidade de novas pesquisas que explorem diferentes técnicas e contextos populacionais para aprimorar a prevenção do nascimento pré-termo extremo.

REFERÊNCIAS

- ADAM, Taghreed *et al.* Restrições de capacidade para adoção de novas intervenções: tempo de consulta e a Gestão Integrada de Doenças Infantis no Brasil. **Política e planejamento de saúde**, v. 20, n. suppl_1, p. i49-i57. 2015.
- ADHIKARI, Kamala *et al.* Does neighborhood socioeconomic status predict the risk of preterm birth? A community-based Canadian cohort study. **BMJ Open**, v.9, n. 2, p. e025341, 2019
- ALBERTON, Marcos *et al.* Prevalência e tendência temporal da prematuridade no Brasil antes e durante a pandemia de covid-19: análise da série histórica 2011-2021. **Revista do SUS, Epidemiologia, serviços de saúde**, v. 32, n.2. 2023. DOI: doi.org/10.1590/S2237-96222023000200005
- ALLEMAN, Brandon W *et al.* Individual and center-level factors affecting mortality among extremely low birth weight infants. **Pediatrics**, v. 132, n. 1, p. e175-84. 2013.
- AMAZON. **Machine learning** [Internet]; 2022 [cited 2024, Agost 02]. Available from: <https://aws.amazon.com/pt/training/learn-about/machine-learning/>.
- ARAÚJO FILHO, José de Arimatéia Batista; NOMURA, Cesar Higa; ASSUNÇÃO JÚNIOR, Antonildes Nascimento; CHIAVEGATTO FILHO, Alexandre Dias Porto; HORVAT, Natally. Inteligência artificial na era da medicina de precisão: o que é preciso saber? **Arimed/PROCLIM, (capítulo). 2023**. Disponível em: <portal.secad.artmed.com.br/artigo/inteligencia-artificial-na-era-da-medicina-de-precisao-o-que-e-preciso-saber>. Acesso em: 14 jul. 2025.
- AYNALEM, Yared Asmare *et al.* The Magnitude of Neonatal Mortality and Its Predictors in Ethiopia: A Systematic Review and Meta-Analysis. **International Journal of Pediatrics**, 2021. DOI: 10.1155/2021/7478108.
- BADILLO, Solveig *et al.* An Introduction to Machine Learning. **Clinical Pharmacology and Therapeutics**, v. 107, n. 4, p. 871–885. 2020.
- BATISTA, André Filipe Moraes; CHIAVEGATTO FILHO, Alexandre Dias Porto. Machine Learning aplicado à Saúde. In: IN: ZIVIANI, Artur, FERNANDES, Natalia Castro, SAADE, Débora Cristina Muchaluate (ORG.). **Livro de Minicursos do SBCAS**. Sociedade Brasileira de Computação. Niterói, RJ: [s. n.]. 2019.

BAZO, Gabriel. Desenvolvimento e análise de modelos de estatística multivariada baseados em Aprendizado de Máquina: predição de partos pré-termo na coorte de nascimento BRISA em Ribeirão Preto. Ribeirão Preto. 2020. **Dissertação (Mestrado)- Universidade de São Paulo**, Ribeirão Preto. 2020. <https://www.teses.usp.br/teses/disponiveis/17/17144/tde-07092020-102951>.

BEHBOUDI-GANDEVANI S, Bidhendi-Yarandi R, Hossein Panahi M, Mardani A, Prinds C, Vaismoradi M, Glarcher M. Prevalence of preterm birth in Scandinavian countries: a systematic review and meta-analysis. **J Int Med Res.**, v. 51, n.10, e:3000605231203843. 2023. DOI: 10.1177/03000605231203843.

BILSTEEN JF, Andresen JB, Mortensen LH, et al. Educational disparities in perinatal health in Denmark in the first decade of the 21st century: a register-based cohort study. **BMJ Open** v.8, e023531. 2018. DOI: 10.1136/bmjopen-2018-023531.

BITTAR, R.E. Parto pré-termo. **Revista de Medicina**, v. 97, n. 2. P.195-207. 2018.

BRASIL. Ministério da Saúde (BR). Secretaria de Atenção à Saúde. **Departamento de Ações Programáticas Estratégicas**. Pré-natal e puerpério: atenção qualificada e humanizada: manual técnico [Internet]. Brasília: Ministério da Saúde; 2006.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS (DATASUS). Informações de Saúde TABNET: **Nascidos vivos – Brasil**. Disponível em: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinasc/cnv/nvuf.def>. Acesso em: 23 out. 2025.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS – DATASUS. **Sistema de Informações sobre Nascidos Vivos – SINASC**: dados tabulados por meio do TABWIN. Disponível em: <http://www2.datasus.gov.br>. Acesso em: 08 jun. 2025.

CARTER, J. *et al.* Development and validation of prediction models for the QUiPP App v.2: a tool for predicting preterm birth in women with symptoms of threatened preterm labor. **Ultrasound Obstet Gynecol.**, v.55, p. 357–367. 2020. DOI: doi.org/10.1002/uog.20422

CRUMP C, Friberg D, Li X, et al. Preterm birth and risk of sleep-disordered breathing from childhood into mid-adulthood. **Int J Epidemiol.**, v.48, p. 2039–2049. 2019. DOI: 10.1093/ije/dyz075.

DAVIS, Jesse; GOADRIC, Mark. A relação entre curvas de precisão-revoação e curvas ROC. Em: **ICML**, 2006.

DAMASO, Enio Luis, *et al.* Prediction of Preterm Birth by Maternal Characteristics and Medical History in the Brazilian Population. **J Pregnancy**, sep. 2019.

DEB-RINKER P, León JA, Gilbert NL et al. Diferenças entre mortalidade perinatal e infantil em países de alta renda: artefatos de registro de nascimento ou evidências de diferenças reais? **BMC Pediatr.**, v. 15, p.112. 2015. DOI: 10.1186/s12887-015-0430-8.

DIAS, Barbara Almeida Soares *et al.* Prematuridade recorrente: dados do estudo “Nascer no Brasil”. **Revista de Saúde Pública**, v. 56, p. 7, 2022.

DIAS, M. R. A importância do dia da Prematuridade: Análise da Lei 203/20 no Maranhão **Revista Brasileira de Saúde e Educação**, v. 10, n.1, pp. 45-60, 2021.

DING, L., Yin, X., Wen, G. et al. Prediction of preterm birth using machine learning: a comprehensive analysis based on large-scale preschool children survey data in Shenzhen of China. **BMC Pregnancy and Childbirth.**, v.24, p.810. 2024.

FACELI, K., Lorena, A.C., GAMA, J., CARVALHO, A.C.P.L.F. Inteligência Artificial: Uma abordagem de aprendizado de máquina. Rio de Janeiro: **LTC**, 2011.

FERGUS P *et al.* Prediction of preterm deliveries from EHG signals using machine learning. **PLoS One.**, v.8, n. 10, p. e77154, 2013. DOI: doi.org/101371/journal.pone.0077154.

FOWLER JR, Mahdy H, Vadakekut ES, Jack BW. Cuidados Anteparto no Segundo e Terceiro Trimestre. 18 de julho de 2024. In: *StatPearls*. Treasure Island (FL): **StatPearls Publishing**; 2025 jan–.

GAO S. C., *et al.*, 2019. Deep learning predicts extreme preterm birth from electronic health records. **Journal of Biomedical Informatics. Elsevier**. December 2019. DOI: doi.org/10.1016/j.jbi.2019.103334.

GETTE F, Aziz Ali S, Ho MSP, Richter LL, Chan ES, Yang CL, Kieran E, Mammen C, Roberts A, Kang KT, Wong J, Rassekh SR, Castaldo M, Harris KC, Lee J, Lam CKL, Chan NH, Lisonkova S, Ting JY. Long-term health outcomes of preterm birth: a narrative review. **Front Pediatr**. 2025 Apr 23;13:1565897. DOI: 10.3389/fped.2025.1565897.

GÉRON, Aurélien. Aprendizado de Máquina Prático com Scikit-Learn, Keras e TensorFlow. 2ª edição. **O'Reilly**, 2019.

GOVERNO DO ESTADO DO MARANHÃO. Plano Maranhão 2050: Plano Estratégico de Longo Prazo do Estado do Maranhão. São Luís, MA, 2024. Disponível em: <https://maranhao2050.ma.gov.br/>. Acesso em: dez. 2025.

GRÉTARSDÓTTIR ÁS, Aspelund T, Steingrímsdóttir W, et al. Preterm births in Iceland 1997-2016: Preterm birth rates by gestational age groups and type of preterm birth. **Birth** 2020; 47: 105–114. 20191119. DOI: 10.1111/birt.12467.

GUL, A., Elrasheed, F., Logman, H., et al. Role of Artificial Intelligence in Risk Prediction of Preterm Labor: Narrative Review. **South Eastern European Journal of Public Health**. (2024) Vol XXV S2.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. Nova York: springer, 2009.

HASSAN, A., Nawaz, S., Tahira, S., & Ahmed, A. “Preterm Birth Prediction Using an Explainable Machine Learning Approach.” **Artificial Intelligence and Applications**, 2025 DOI:10.47852 /bonviewAIA52024517.

HE, Haibo; GARCIA, Eduardo A. Aprendizado a partir de dados desbalanceados. **IEEE Transactions on Knowledge and Data Engineering**, 2009.

HOGLUND B, Lindgren P, Larsson M. Pregnancy and birth outcomes of women with intellectual disability in Sweden: a national register study. **Acta Obstet Gynecol Scand** 2012; 91: 1381–1387. 20120918. DOI: 10.1111/j.1600-0412.2012.01509.x.

KUHN, Max; JOHNSON, Kjell. *Applied Predictive Modeling*. New York: **Springer**, 2013. DOI: 10.1007/978-1-4614-6849-3.

KURSA, Miron B.; RUDNICKI, Witold R. Feature selection with the Boruta package. **Journal of statistical software**, v. 36, p. 1-13, 2010.

KOIVU, A., Sairanen, M. Previsão do risco de natimortos e gestações prematuras com aprendizado de máquina. **Sistemas de Ciência da Informação em Saúde**, v. 8, n. 1, p.14. 2020. DOI:10.1007/s13755-020-00105-9.

LACY, Nina *et al.* Predicting individual cases of major adolescent psychiatric conditions with artificial intelligence. **Translational psychiatry**, v. 13, n. 1, p. 314, 2023.

MCDERMOTT, Matthew B. A. *et al.* A Closer Look at AUROC and AUPRC under Class Imbalance.archivePrefix 2401.0609.1v1 [cs.LG] 11 Jan 2024.

<https://arxiv.org/abs/2401.06091>. Acesso em: marc. 2026.

MARTINELLI, Katrini Guidolini *et al.* Prematuridade no Brasil entre 2012 e 2019: dados do Sistema de Informações sobre Nascidos Vivos. **Revista Brasileira de Estudos de População**, v. 38, p. e0173, 2021.

MOONS, Karel GM *et al.* Transparent Reporting of a multivariable prediction model for Individual Prognosis or Diagnosis (TRIPOD): explanation and elaboration. **Annals of internal medicine**, v. 162, n. 1, p. W1-W73, 2015.

MORGEN CS, Bjørk C, Andersen PK, *et al.* Posição socioeconômica e o risco de parto prematuro – um estudo dentro da Coorte Nacional de Nascimento da Dinamarca. **International Journal of Epidemiology**., v.37, p.1109–1120. 2008. Acesso abril 13, 2026

OLIVEIRA, Laura Leismann de *et al.* Fatores maternos e neonatais relacionados à prematuridade. **Revista da Escola de Enfermagem da USP**, v. 50, n. 03, pág. 382-389, 2016.

OUATTARA, A. *et al.* Les facteurs associés à l'accouchement prématuré en milieu urbain africain À propos d'une étude cas témoins au CHU-YO et au centre médical Saint-Camille de Ouagadougou. **Médecine et Santé Tropicales**, v. 25, n. 3, p. 296-299, 2015.

PAIXÃO, Gabriela Miana de Mattos *et al.* Machine Learning na Medicina: Revisão e Aplicabilidade. **Arquivos Brasileiros de Cardiologia**, vol. 118, p. 95-102, 2022.

PASSINI R Jr, Cecatti JG, Lajos GJ, Tedesco RP, Nomura ML, Dias TZ, Haddad SM, Rehder PM, Pacagnella RC, Costa ML, Sousa MH; Brazilian Multicentre Study on Preterm Birth Study Group. Brazilian Multicentre Study on Preterm Birth (EMIP): Prevalence and factors associated with spontaneous preterm birth. PLoS ONE. 2014 Oct 9;9(10):e109069. DOI: 10.1371/journal.pone.0109069

PEREIRA, Thaís Basso de Brito *et al.* Regulatory Cytokine Expression and Preterm Birth: case-control study in a cohort. **PLOS ONE**, v. 11, n.8, 2016. DOI: doi.org/10.1371/journal.pone.0158380.

PFOB, André; LU, Sheng Chieh; SIDEY-GIBBONS, Chris. Machine learning in medicine: a practical introduction to techniques for data pre-processing, hyperparameter tuning, and model comparison. **BMC Medical Research Methodology**, v. 22, n. 1, 2022.

ROCHA, P. R. H. **Influência da prematuridade e restrição do crescimento intrauterino na habilidade motora na infância [tese]**. Universidade de São Paulo, Ribeirão Preto, 2019.

ROKACH, Lior; MAIMON, Oded. Data mining with decision trees: theory and applications. 2nd edition). **Series in Machine Perception and Artificial Intelligence**, v. 81, 2015.

ROWE, Michael. **An Introduction to Machine Learning for Clinicians**. [S. l.]: Lippincott Williams and Wilkins, 2019.

SAITO, Takaya; REHMSMEIER, Marc. PLOS ONE, 2015. O gráfico de precisão-revoação é mais informativo do que o gráfico ROC ao avaliar classificadores binários em conjuntos de dados desbalanceados. **PLOS ONE**, 2015

SANTOS, Hellen Geremias dos *et al.* Machine learning para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil. **Cadernos De saúde pública**, v. 35, p. e00050818, 2019.

SIDEY-GIBBONS, Jenni A.M.; SIDEY-GIBBONS, Chris J. Machine learning in medicine: a practical introduction. **BMC Medical Research Methodology**, v. 19, n. 1, 2019.

SILVA, ANTÔNIO AUGUSTO DA; SIMÕES, VANDA MARIA; Barbieri, Marco Antonio; CARDOSO, VIVIANE CUNHA; ALVES, CLAUDIA MARIA; **THOMAZ, ERIKA BÁRBARA**; QUEIROZ, REJANE CHRISTINE; CAVALLI, RICARDO CARVALHO; BATISTA, ROSÂNGELA FERNANDES; Bettiol, Heloísa. A protocol to identify non-classical risk factors for preterm births: the Brazilian Ribeirao Preto and Sao Luis prenatal cohort (BRISA). **Reproductive Health JCR**, v. 11, p. 79, 2014.

SIMHAN, Hyagriv N., & CARITIS, Steve N. Prevention of preterm delivery. *New England Journal of Medicine*, v. 357, n. 5, p. 477-48, 2007. DOI: 10.1056/NEJMra050435.

STEYERBERG, Ewout W. **Applications of prediction models**. Springer New York, 2009.

STEYERBERG, E. W. **Clinical Prediction Models: A Practical Approach to Development, Validation, and Updating**. 2nd ed. Cham: Springer, 2019.

SUFF, Natalie; STORY, Lisa; SHENNAN, Andrew. A previsão do parto prematuro: o que há de novo? In: **Seminários em Medicina Fetal e Neonatal**, p.27-32, WB Saunders, 2019.

SUN, Qi; ZOU, Xiaoxuan; YAN, Yousheng; ZHANG, Hongguang; WANG, Shuo; GAO, Yongmei; LIU, Haiyan; LIU, Shuyu; LU, Jianbo; YANG, Ying; MA, Xu. Machine learning-based prediction model of preterm birth using electronic health record. **Journal of Healthcare Engineering**, v. 2022, p. 9635526, 2022. DOI: 10.1155/2022/9635526.

SWAMY GK, Ostbye T, Skjaerven R. Associação do parto prematuro com sobrevivência a longo prazo, reprodução e parto prematuro de próxima geração. **Jama**, v.299, p.1429–1436. 2008. DOI: 10.1001/jama.299.12.1429. DOI: doi.org/10.1001/jama.299.12.1429

TAYLOR GL, O'Shea TM. Extreme Prematurity: Risk and Resiliency. **Curr Probl Pediatr Adolesc Health Care.**, v. 52, n.2, e:101132. 2022. DOI: 10.1016/j.cppeds.2022.101132

TINGLEFF T, Räisänen S, Vikanes Å et al. Associação entre país de nascimento materno e nascimento prematuro: um estudo de registro populacional de 910.752 partos. **Scand J Public Health**, v.49, p.904–913. 2021. DOI: doi.org/10.1177/1403494821992894

VON Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. **Lancet.**, 370(9596):1453–1457. 2007

WEBER, Ann *et al.* Application of machine-learning to predict early spontaneous preterm birth among nulliparous non-Hispanic black and white women. **Ann Epidemiol**, v.28, n.11, p.783-789, 2018.

World Health Organization. **Born too soon: the global action report on preterm birth.**

Geneva: 2012. Disponível em:

https://www.who.int/pmnch/media/news/2012/201204_borntoosoon-report.pdf. Acesso em 30 de julho de 2024.

World Health Organization. **Preterm birth.** 2023. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/preterm-birth>. Acesso em 30 de julho de 2024.

World Health Organization. **Preterm birth: Key facts.** Geneva: WHO, 2019. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/preterm-birth>. Acesso em: 14 jul. 2024.

World Health Organization.; UNICEF; LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE. **Born Too Soon: Decade of Action on Preterm Birth.** Geneva: World Health

Organization, 2023. Disponível em: <https://www.who.int/news/item/09-05-2023-152-million-babies-born-preterm-in-the-last-decade>. Acesso em: 9 jun. 2025.

YU, Q. Y. et al. Predicting risk of preterm birth in singleton pregnancies using machine learning algorithms. **Frontiers in Big Data**, v. 7, 2024. Disponível em: <https://doi.org/10.3389/fdata.2024.1291196>. Acesso em: 04 maio 2026.

ZHOU, Zhi-Hua. **Aprendizado de máquina**. Springer Nature, 2021.

ZUGAIB, Marcelo. **Zugaib Obstetrícia**. 2º ed. Barueri, São Paulo: Manole, 2010.

7 CONSIDERAÇÕES FINAIS

Com base nos resultados obtidos, este estudo evidencia o potencial significativo do uso de modelos de aprendizado de máquina na predição de desfechos de nascimento pré-termo extremo (NPT extremo). Modelos como regressão logística penalizada (plr), suporte vector machine (svmLinear e svmRadial), random forest (rf, ranger e Rborist), boosting (glmboost e AdaBoost.M1), métodos baseados em vizinhança (Knn e KKnn), redes neurais artificiais (nnet), árvores e modelos baseados em partição (cforest e ctree2) e modelos aditivos generalizados (gam) demonstraram desempenho promissor, mesmo diante do desbalanceamento da base de dados, sendo capazes de discriminar com eficácia neonatos em risco e permitir intervenções clínicas antecipadas.

Os modelos avaliados no cenário que incluiu a variável ameaça de parto prematuro na gestação atual apresentaram os melhores desempenhos. No cenário combinado São Luís–Ribeirão Preto, o Knn destacou-se com sensibilidade de 0,774, AUC-ROC de 0,806 e F1-score de 0,049. Em São Luís, o svmRadial apresentou sensibilidade de 0,653, AUC-ROC de 0,732 e F1-score de 0,063. Já em Ribeirão Preto, o ranger obteve AUC-ROC de 0,860, sensibilidade de 0,771 e F1-score de 0,053. No cenário sem a inclusão dessa variável, os melhores resultados foram observados para o nnet em São Luís–Ribeirão Preto (sensibilidade de 0,561, AUC-ROC de 0,648 e F1-score de 0,019), para o gam em São Luís (sensibilidade de 0,623, AUC-ROC de 0,587 e F1-score de 0,024) e para o Rborist em Ribeirão Preto (AUC-ROC de 0,629, sensibilidade de 0,573 e F1-score de 0,016). Esses achados reforçam a importância da seleção da técnica mais adequada às características da população estudada e aos objetivos clínicos, buscando o melhor equilíbrio entre sensibilidade e capacidade discriminatória.

A incorporação dessas ferramentas em sistemas de saúde, especialmente em contextos com recursos limitados, pode representar uma estratégia inovadora para o aprimoramento da assistência pré-natal. Além de contribuir para a otimização dos recursos do sistema público de saúde, a aplicação desses modelos possui potencial para qualificar a tomada de decisão clínica, favorecer intervenções precoces e melhorar os desfechos maternos e neonatais, em consonância com as diretrizes globais voltadas ao fortalecimento da saúde perinatal.

Adicionalmente, os resultados desta pesquisa podem servir como base para novos estudos que explorem diferentes algoritmos, populações e contextos assistenciais. A

continuidade dessas investigações é fundamental para o avanço da medicina preditiva e para a consolidação de soluções tecnológicas capazes de apoiar a prevenção e o manejo de condições associadas à prematuridade extrema.

Por fim, a conclusão deste trabalho representa não apenas uma contribuição científica para a área da saúde materno-infantil, mas também a concretização de uma importante trajetória pessoal e profissional. O desenvolvimento desta pesquisa exigiu dedicação, disciplina e resiliência diante dos desafios de conciliar as demandas do mestrado com as responsabilidades profissionais e familiares. Cada etapa foi marcada por aprendizados, superações e pela busca constante do equilíbrio entre diferentes papéis e compromissos. Assim, a realização desta dissertação simboliza não apenas a obtenção do título de mestre, mas também a confirmação de que a persistência, o esforço contínuo e o apoio da família, amigos, professores e colegas são fundamentais para transformar desafios em conquistas e sonhos em realidade.

REFERÊNCIAS

- ABREU, Lidiene Penha *et al.* Violência por parceiro íntimo na gestação e tempo de retorno das atividades sexuais após o parto: análise da coorte de pré-natal BRISA. **Cadernos de Saúde Pública**, v.40, n. 5, 2023 <https://doi.org/10.1590/0102-311XEN094223>
- ADAM, Taghreed *et al.* Restrições de capacidade para adoção de novas intervenções: tempo de consulta e a Gestão Integrada de Doenças Infantis no Brasil. **Política e planejamento de saúde**, v. 20, n. suppl_1, p. i49-i57, 2015.
- ADHIKARI, Kamala *et al.* Does neighborhood socioeconomic status predict the risk of preterm birth? A community-based Canadian cohort study. **BMJ Open**, v.9, n. 2, p. e025341, 2019. <https://doi.org/10.1136/bmjopen-2018-025341>
- AHUMADA-BARRIOS, Margarita E.; ALVARADO, German F. Fatores de Risco para parto prematuro em um hospital. **Revista Latino-Americana de Enfermagem**, v. 24, 2016.
- ALBERTON, Marcos *et al.* Prevalência e tendência temporal da prematuridade no Brasil antes e durante a pandemia de covid-19: análise da série histórica 2011-2021. *Revista do SUS, Epidemiologia, serviços de saúde* v. 32, n.2, 2023. <https://doi.org/10.1590/S2237-96222023000200005>
- ALLEMAN, Brandon W *et al.* Individual and center-level factors affecting mortality among extremely low birth weight infants. **Pediatrics**, v. 132, n. 1, p. e175-84, 2013.
- ALMEIDA, Bruna de; COUTO, Rafael Henrique Mastella; TRAPANI JUNIOR, Alberto. Prevalência e fatores associados aos óbitos em prematuros internados. **Arquivos Catarinenses de Medicina**, v. 48, n. 4, p. 35-50, 2019.
- ALMEIDA, Joana Patricia Rodrigues Félix Peixoto de; BARTOSCH, Carla Maria Magno; MACEDO, Alexandra Matias Pereira Cunha Coelho. O impacto da medição por rotina do comprimento cervical por ecografia transvaginal na predição de parto pré-termo: Um estudo retrospectivo num hospital terciário. **Revista Brasileira de Ginecologia e Obstetrícia**, v. 43, p. 264-274, 2021.
- AMAZON. Machine learning [Internet]; 2022 [cited 2024, Agost 02]. Available from: <https://aws.amazon.com/pt/training/learn-about/machine-learning/>.
- ARAÚJO FILHO, José de Arimatéia Batista; NOMURA, Cesar Higa; ASSUNÇÃO JÚNIOR, Antonildes Nascimento; CHIAVEGATTO FILHO, Alexandre Dias Porto; HORVAT,

Natally. Inteligência artificial na era da medicina de precisão: o que é preciso saber?

Arimed/PROCLIM, (capítulo), 2023. Disponível em:

<portal.secad.artmed.com.br/artigo/inteligencia-artificial-na-era-da-medicina-de-precisao-o-que-e-preciso-saber>. Acesso em: 14 jul. 2025.

ARAÚJO, Saulo A.; SANTOS, João V. R. dos; GANDOLFI, Vinícius C.; TOREZZAN, Cristiano. Métodos de aprendizado de máquina aplicados para a previsão de prematuridade e do peso ao nascer no Brasil com base em dados públicos. *In: Proceeding Series of the Brazilian Society of Computational and Applied Mathematics*, v. 10, n. 1, CNMAC 2023, 18 dez. 2023.

AREDES, Natália Del Angelo; SANTOS, Raionara Cristina de Araújo; FONSECA, Luciana Mara Monti. Cuidados com a pele do recém-nascido prematuro: revisão integrativa. **Revista Eletrônica de Enfermagem**, v. 19, p. a59-a59, 2017.

Associação Brasileira de Pais, Familiares, Amigos e Cuidadores de Bebês Prematuros – ONG Prematuridade.com. “**Novembro roxo: conscientização sobre a prematuridade**”, 2022 – 2024. <https://prematuridade.com>

AYNALEM, Yared Asmare *et al.* The Magnitude of Neonatal Mortality and Its Predictors in Ethiopia: A Systematic Review and Meta-Analysis. **International Journal of Pediatrics**, 2021. Doi: 10.1155/2021/7478108

AZIZ, Khalid *et al.* Part 5: Neonatal Resuscitation: 2020 American Heart Association Guidelines for Cardiopulmonary Resuscitation and Emergency Cardiovascular Care **Circulation**, v. 142, 16 suppl 2, p. S524-S550, 2020. doi: 10.1161/CIR.0000000000000902

BADILLO, Solveig *et al.* An Introduction to Machine Learning. **Clinical Pharmacology and Therapeutics**, v. 107, n. 4, p. 871–885, 2020.

BARBOSA, D. C. M. Fatores associados com altas taxas de cesáreas na coorte de nascimentos de Ribeirão Preto em 2010: **projeto BRISA**. Ribeirão Preto: Universidade de São Paulo, 2020.

BARFIELD, Wanda D. Public Health Implications of Very Preterm Birth. **Clinical Perinatology**, v. 45, n.5, p: 565-577, 2018. doi: 10.1016/j.clp.2018.05.007.

BARROS, Fernando C *et al.* Caesarean sections and the prevalence of preterm and early-term births in Brazil: secondary analyses of national birth registration. **BMJ Open**, v. 8, 2018.

BARTMAN, Thomas *et al.* Apgar Score at 5 Minutes Is Associated with Mortality in Extremely Preterm Infants Even after Transfer to an All Referral NICU. **American Journal of Perinatology**, v. 32, n 13, p. 1268-1272, 2015.

BATISTA, André Filipe Moraes; CHIAVEGATTO FILHO, Alexandre Dias Porto. Machine Learning aplicado à Saúde. *In*: IN: ZIVIANI, Artur, FERNANDES, Natalia Castro, SAADE, Débora Cristina Muchaluate (ORG.). **Livro de Minicursos do SBCAS**. Sociedade Brasileira de Computação. Niteroi, RJ: [s. n.], 2019.

BAZO, Gabriel. Desenvolvimento e análise de modelos de estatística multivariada baseados em Aprendizado de Máquina: predição de partos pré-termo na coorte de nascimento BRISA em Ribeirão Preto. Ribeirão Preto. 2020. **Dissertação (Mestrado)- Universidade de São Paulo**, Ribeirão Preto, 2020. <https://www.teses.usp.br/teses/disponiveis/17/17144/tde-07092020-102951>.

BEHBOUDI-GANDEVANI S, Bidhendi-Yarandi R, Hossein Panahi M, Mardani A, Prinds C, Vaismoradi M, Glarcher M. Prevalence of preterm birth in Scandinavian countries: a systematic review and meta-analysis. **J Int Med Res.** 2023 Oct; 51(10):3000605231203843. doi: 10.1177/03000605231203843. PMID: 37843530; PMCID: PMC10683576.

BILSTEEN JF, Andresen JB, Mortensen LH, et al. Educational disparities in perinatal health in Denmark in the first decade of the 21st century: a register-based cohort study. **BMJ Open** 2018; 8: e023531. 20181108. DOI: 10.1136/bmjopen-2018-023531. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

BITTAR, R.E . Parto pré-termo. **Revista de Medicina**, v. 97, n. 2. P.195-207, 2018.

BLENCOWE, Hannah *et al.* Estimativas nacionais, regionais e mundiais de taxas de nascimentos prematuros no ano de 2010 com tendências temporais desde 1990 para países selecionados: uma análise sistemática e implicações. **The Lancet**, v. 379, n. 9832, p. 2162-2172, 2012.

BORTOLETTO, Thaísa Guedes *et al.* Parto pré-termo espontâneo sem rotura de membranas: prevalência, fatores de risco e o papel do colo uterino. **Femina**, v. 48, n. 9, p. 568-73, 2020.

BOUZADA, Maria Candida Ferraz *et al.* Resposta aos procedimentos de reanimação neonatal no quinto minuto de vida em recém-nascidos Apgar ≤ 3 no primeiro minuto. **Revista Médica de Minas Gerais**, v. 28, suppl 6, 2018. <https://dx.doi.org/10.5935/2238-3182.20180100>

BRASIL- Ministério da Saúde. **Departamento de Informática do SUS. Datasus.** Estatísticas Vitais. Brasília (DF): Ministério da Saúde; 2019.

BRASIL, Ministério da Saúde. Secretaria de Atenção à Saúde, Departamento de Ações Programáticas Estratégicas. **Atenção humanizada ao recém-nascido: Método canguru**[internet]. Brasília: Ministério da Saúde, 2017. Available from: https://bvsmms.saude.gov.br/bvs/publicacoes/atencao_humanizada_metodo_canguru_manual_3ed.pdf

BRASIL. Ministério da Saúde (BR). Secretaria de Atenção à Saúde. **Departamento de Ações Programáticas Estratégicas.** Pré-natal e puerpério: atenção qualificada e humanizada: manual técnico [Internet]. Brasília: Ministério da Saúde; 2006.

BRASIL. Ministério da Saúde. Altera no âmbito do sistema único de saúde, a portaria GM/MS nº3 de 28/09/2017-SUS- para **A Rede Alyne**. Brasília (DF): Ministério da Saúde, 2024. https://bvsmms.saude.gov.br/saudeelegis/gm/prt5330_13_09_2024.html

BRASIL. Ministério da Saúde. **Atenção à Saúde do Recém-Nascido: Guia para os Profissionais de Saúde – Cuidados com o Recém-Nascido Pré-Termo.** Brasília: Ministério da Saúde, 2011. Disponível em: <https://bvsmms.saude.gov.br>.

BRASIL. Ministério da Saúde. **Atenção ao pré-natal de baixo risco. Cadernos de Atenção Básica**, n.º 32. 1. ed., Brasília: Editora do Ministério da Saúde, 2013.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS – DATASUS. **Sistema de Informações sobre Nascidos Vivos – SINASC:** dados tabulados por meio do TABWIN. Disponível em: <http://www2.datasus.gov.br>. Acesso em: 01 maio. 2026.

BRASIL. Ministério da Saúde. Departamento de Informática do SUS (DATASUS). Informações de Saúde TABNET: **Nascidos vivos – Brasil**. Disponível em: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinasc/cnv/nvuf.def>. Acesso em: 08 out. 2025.

BRASIL. Ministério da Saúde. **Política Nacional de Atenção Integral à Saúde da Criança: orientações para implementação.** Brasília: Ministério da Saúde, 2018.

BRASIL. Ministério da Saúde. **Portaria nº 1.130, de 5 de agosto de 2015.** Institui a Política Nacional de Atenção Integral à Saúde da Criança (PNAISC) no âmbito do Sistema Único de Saúde (SUS). Diário Oficial da União: seção 1, Brasília, DF, 6 ago. 2015.

BREIMAN, L. Random forest. **Machine learning**, v. 45, n.1, p. 5-32, 2001.

CARTER, J. *et al.* Development and validation of prediction models for the QUiPP App v.2: a tool for predicting preterm birth in women with symptoms of threatened preterm labor.

Ultrasound Obstet Gynecol 2020; 55: 357–367. <https://doi.org/10.1002/uog.20422>

CARVALHO, Fernanda C. *et al.* Fatores de risco maternos mais prevalentes relacionados à ocorrência de partos prematuros: Revisão de Literatura. **Brazilian Journal Of Surgery And Clinical Research**, v. 36, n. 1, p. 112-123, 2021.

CARVALHO, Jovanka Bittencourt Leite de *et al.* Condições socioeconômicas da gestação de bebês prematuros. **Rev. enferm. UFPE on line**, p. 386-390, 2018.

CHERMONT, Aurimery Gomes *et al.* Fatores de risco associados à prematuridade e baixo peso ao nascer nos extremos da vida reprodutiva em uma maternidade privada. **Revista Eletrônica Acervo Saúde**, n. 39, p. e2110-e2110, 2020.

CHRISTEN, P. Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection. **Springer**, 2012.

COLLINS Gary S *et al.* Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence. **BMJ Open**, vol 11, p e048008, 2021.

CRAIG ED, THOMPSON JM, MITCHELL EA. Socioeconomic status and preterm birth: New Zealand trends, 1980 to 1999. **Arch Dis Child Fetal Neonatal Ed.**, v.86, p. F142-6, 2002. <https://doi.org/10.1136/fn.86.3.f142>

CRUMP C, Friberg D, Li X, et al. Preterm birth and risk of sleep-disordered breathing from childhood into mid-adulthood. **Int J Epidemiol** 2019; 48: 2039–2049. DOI: 10.1093/ije/dyz075.

DAMASO, Enio Luis, *et al.* Prediction of Preterm Birth by Maternal Characteristics and Medical History in the Brazilian Population. **J Pregnancy**, sep. 2019. <https://doi.org/10.1155/2019/4395217>

DAVIS, Jesse; GOADRIC, Mark. A relação entre curvas de precisão-revocação e curvas ROC. Em: **ICML**, 2006.

DEB-RINKER P, León JA, Gilbert NL et al. Diferenças entre mortalidade perinatal e infantil em países de alta renda: artefatos de registro de nascimento ou evidências de diferenças reais? **BMC Pediatr** 2015; 15: 112. 20150904. DOI: 10.1186/s12887-015-0430-8.

- DEFILIPO, Érica Cesário *et al.* Fatores associados à prematuridade: estudo caso-controle. **Revista Paulista de Pediatria**, v. 40, p. e2020486, 2022.
- DIAS, Barbara Almeida Soares *et al.* Prematuridade recorrente: dados do estudo “Nascer no Brasil”. **Revista de Saúde Pública**, v. 56, p. 7, 2022.
- DIAS, M. R. A importância do dia da Prematuridade: Análise da Lei 203/20 no Maranhão **Revista Brasileira de Saúde e Educação**, v. 10, n.1, pp. 45-60, 2021.
- DING, L., Yin, X., Wen, G. et al. Prediction of preterm birth using machine learning: a comprehensive analysis based on large-scale preschool children survey data in Shenzhen of China. **BMC Pregnancy and Childbirth**. 2024 24:810.
- FACELI, K., Lorena, A.C., GAMA, J., CARVALHO, A.C.P.L.F. Inteligência Artificial: Uma abordagem de aprendizado de máquina. Rio de Janeiro: **LTC**, 2011.
- FACELI, Katti; LORENA, Ana Carolina; GAMA, João; ALMEIDA, Tiago Agostinho de; CARVALHO, André Carlos Ponce de Leon Ferreira de. **Inteligência artificial: uma abordagem de aprendizado de máquina**. 2. ed. Rio de Janeiro: LTC, 2021. ISBN 978-8521637349.
- FERGUS P et al. Prediction of preterm deliveries from EHG signals using machine learning. **PLoS One.**, v.8, n. 10, p. e77154, 2013. <https://doi.org/10.1371/journal.pone.0077154>.
- FOWLER JR, Mahdy H, Vadakekut ES, Jack BW. Cuidados Anteparto no Segundo e Terceiro Trimestre. 18 de julho de 2024. In: *StatPearls*. Treasure Island (FL): **StatPearls Publishing**; 2025 jan—. PMID: 28846223.
- FREITAS, Paulo Fontoura; ARAÚJO, Rafael Rodrigues de. Prematuridade e fatores associados em Santa Catarina, Brasil: análise após alteração do campo idade gestacional na Declaração de Nascidos Vivos. **Revista Brasileira de Saúde Materno Infantil**, v. 15, p. 309-316, 2015.
- GAÍVA, Maria Aparecida; MODES, Priscilla Shirley Siniak dos Anjos.; STEVANATO, Jaquelline Monte. Fatores associados à prematuridade: análise dos dados do sistema de informações sobre nascidos vivos. **Rev enfermagem UFPE online**, v. 9, n. 12, p. 1110-7, 2015.

GAO S. C., *et al.*, 2019. Deep learning predicts extreme preterm birth from electronic health records. **Journal of Biomedical Informatics. Elsevier**. December 2019.
<https://doi.org/10.1016/j.jbi.2019.103334>.

GÉRON, Aurélien. *Aprendizado de Máquina Prático com Scikit-Learn, Keras e TensorFlow*. 2ª edição. **O'Reilly**, 2019.

GETTE F, Aziz Ali S, Ho MSP, Richter LL, Chan ES, Yang CL, Kieran E, Mammen C, Roberts A, Kang KT, Wong J, Rassekh SR, Castaldo M, Harris KC, Lee J, Lam CKL, Chan NH, Lisonkova S, Ting JY. Long-term health outcomes of preterm birth: a narrative review. **Front Pediatr**. 2025 Apr 23;13:1565897. doi: 10.3389/fped.2025.1565897. PMID: 40336800; PMCID: PMC12055779.

GINAR, Daniele Olivo. Prevalência e fatores associados à prematuridade, em Santa Catarina, no período de 2016 a 2018. **Medicina-Pedra Branca**, 2020.

GRÉTARSDÓTTIR ÁS, Aspelund T, Steingrimsdóttir W, et al. Preterm births in Iceland 1997-2016: Preterm birth rates by gestational age groups and type of preterm birth. **Birth** 2020; 47: 105–114. 20191119. DOI: 10.1111/birt.12467. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

GUIMARÃES, Eliete Albano de Azevedo *et al.* Prevalência e fatores associados à prematuridade em Divinópolis, Minas Gerais, 2008-2011: análise do Sistema de Informações sobre Nascidos Vivos. **Epidemiologia e Serviços de saúde**, v. 26, p. 91-98, 2017.

GUL, A., Elrasheed, F., Logman, H., et al. Role of Artificial Intelligence in Risk Prediction of Preterm Labor: Narrative Review. **South Eastern European Journal of Public Health**. (2024) Vol XXV S2.

HARRISON, Conrad J.; SIDEY-GIBBONS, Chris J. Machine learning in medicine: a practical introduction to natural language processing. **BMC medical research methodology**, v. 21, n. 1, p. 158, 2021.

HASTIE, Trevor; TIBSHIRANI, Robert; FRIEDMAN, Jerome. **The Elements of Statistical Learning: Data Mining, Inference, and Prediction**. Nova York: springer, 2009.

HE, Haibo; GARCIA, Eduardo A. *Aprendizado a partir de dados desbalanceados*. **IEEE Transactions on Knowledge and Data Engineering**, 2009.

HOGLUND B, Lindgren P, Larsson M. Pregnancy and birth outcomes of women with intellectual disability in Sweden: a national register study. **Acta Obstet Gynecol Scand** 2012; 91: 1381–1387. 20120918. DOI: 10.1111/j.1600-0412.2012.01509.x.

HUTCHEON, J. A. et al. Small size at birth or abnormal intrauterine growth trajectory: Which matters more for child growth? **American Journal of Epidemiology**, v. 183, n. 12, p. 1107–1113, 2016.

KOIVU, A., Sairanen, M. Previsão do risco de natimortos e gestações prematuras com aprendizado de máquina. **Health Inf Sci Syst** v.8, n.1, p.14. 2020. DOI: 10.1007/s13755-020-00105-9

LACY, Nina *et al.* Predicting individual cases of major adolescent psychiatric conditions with artificial intelligence. **Translational psychiatry**, v. 13, n. 1, p. 314, 2023.

LIU, L; WANG, Y; ZHANG, J. New Machine Learning Algorithm: Random Forest. In: Liu, B., Ma, M., Chang, J. (eds) Information Computing and Applications. ICICA 2012. Lecture Notes in Computer Science, vol 7473. Springer, Berlin, Heidelberg. 2012

LIU, Li *et al.* Causas globais, regionais e nacionais da mortalidade de menores de 5 anos em 2000–15: uma análise sistemática atualizada com implicações para os Objetivos de Desenvolvimento Sustentável. **The Lancet**, v. 388, n. 10063, p. 3027-3035, 2016.

MARANHÃO. Plano Maranhão 2050: planejamento estratégico de longo prazo para o desenvolvimento sustentável do Estado do Maranhão. São Luís: **Governo do Estado do Maranhão, 2024**. Disponível em: <<https://maranhao2050.ma.gov.br/>>. Acesso em: 02 fev. 2026.

MARCONDES, E. *et al.* Pediatria: Fundamentos e Prática. 11. ed. **São Paulo: Sarvier**, p. 287-305, 2020.

MARTINELLI, Katrini Guidolini *et al.* Prematuridade no Brasil entre 2012 e 2019: dados do Sistema de Informações sobre Nascidos Vivos. **Revista Brasileira de Estudos de População**, v. 38, p. e0173, 2021.

MEDEIROS, M. N. L. et al. Validity of pre and post-term birth rates based on the date of last menstrual period compared to early obstetric ultrasonography. **Cadernos de saúde pública**, v. 31, n. 4, p. 885–90, 2015.

MONTEIRO Laranjeira, P. F., Batista dos Santos, V. C., da Silva Nogueira, J., Monteiro Laranjeira, A. C., Monteiro Amorim, G., Monteiro Laranjeira, R., & Nóbrega Barbosa, K. G. (2022). Prematuridade extrema em uma unidade pública de referência: morbidade, viabilidade e mortalidade. *Research, Society and Development*, v.11, n. 6, e58311629468. DOI: <http://dx.doi.org/10.33448/rsd-v11i6.29468>

MORGEN CS, Bjørk C, Andersen PK, et al. Socioeconomic position and the risk of preterm birth—a study within the Danish National Birth Cohort. *Int J Epidemiol.*, v.37, p.1109–1120. 2008. DOI: 10.1093/ije/dyn112.

OLIVEIRA, Laura Leismann de *et al.* Fatores maternos e neonatais relacionados à prematuridade. *Revista da Escola de Enfermagem da USP*, v. 50, n. 03, p. 382-389, 2016.

OSCHINA. **K-Fold Cross Validation** [2019]. Disponível em: <https://my.oschina.net/Bettyty/blog/751627>. Acesso em: 27 jun. 2025.

OUATTARA, A. et al. Les facteurs associés à l'accouchement prématuré en milieu urbain africain À propos d'une étude cas témoins au CHU-YO et au centre médical Saint-Camille de Ouagadougou. *Médecine et Santé Tropicales*, v. 25, n. 3, p. 296-299, 2015.

PAIXÃO, Gabriela Miana de Mattos *et al.* Machine Learning na Medicina: Revisão e Aplicabilidade. *Arquivos Brasileiros de Cardiologia*, vol. 118, p. 95-102, 2022.

PASSINI R Jr, Cecatti JG, Lajos GJ, Tedesco RP, Nomura ML, Dias TZ, Haddad SM, Rehder PM, Pacagnella RC, Costa ML, Sousa MH; Brazilian Multicentre Study on Preterm Birth Study Group. Brazilian Multicentre Study on Preterm Birth (EMIP): Prevalence and factors associated with spontaneous preterm birth. *PLoS ONE*, v. 9, n. 10, e109069. 2014. DOI:10.1371/journal.pone.0109069

PEREIRA, Thaís Basso de Brito *et al.* Regulatory Cytokine Expression and Preterm Birth: case-control study in a cohort. *PLOS ONE*, v. 11, n.8, 2016. DOI: 10.1371/journal.pone.0158380.

PFOB, André; LU, Sheng Chieh; SIDEY-GIBBONS, Chris. Machine learning in medicine: a practical introduction to techniques for data pre-processing, hyperparameter tuning, and model comparison. *BMC Medical Research Methodology*, v. 22, n. 1, 2022.

PITILIN, Érica de Brito *et al.* Fatores perinatais associados à prematuridade em unidade de terapia intensiva neonatal. *Texto & Contexto-Enfermagem*, v. 30, 2021.

RIBEIRO, Marizélia Rodrigues Costa *et al.* PESQUISA BRISA: estudo de coorte para investigação de nascimento pré-termo e saúde da criança. **VII Conferência Internacional de Políticas Públicas**. UFMA, 2015.

ROCHA, P. R. H. **Influência da prematuridade e restrição do crescimento intrauterino na habilidade motora na infância [tese]**. Universidade de São Paulo, Ribeirão Preto, 2019.

ROKACH, Lior; MAIMON, Oded. Data mining with decision trees: theory and applications. 2nd edition). **Series in Machine Perception and Artificial Intelligence**, v. 81, 2015.

ROWE, Michael. **An Introduction to Machine Learning for Clinicians**. [S. l.]: Lippincott Williams and Wilkins, 2019.

SAITO, Takaya; REHMSMEIER, Marc. PLOS ONE, 2015. O gráfico de precisão-revocação é mais informativo do que o gráfico ROC ao avaliar classificadores binários em conjuntos de dados desbalanceados. **PLOS ONE**, 2015

SALUSTIANO, Letícia Pacífico de Queiroz *et al.* Factors associated with duration of breastfeeding in children under six months. **Revista Brasileira de Ginecologia e Obstetrícia**, v. 34, p. 28-33, 2012.

SANTOS, Hellen Geremias dos *et al.* Machine learning para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil. **Cadernos De saúde pública**, v. 35, p. e00050818, 2019.

SHARIFI-HERIS, Zahra *et al.* Machine learning approach for preterm birth prediction using health records: a systematic review. **Journal of Biomedical Informatics**, v. 129, p. 104040, 2022.

SIDEY-GIBBONS, Jenni A.M.; SIDEY-GIBBONS, Chris J. Machine learning in medicine: a practical introduction. **BMC Medical Research Methodology**, v. 19, n. 1, 2019.

SILVA AA, Simões VM, Barbieri MA, Cardoso VC, Alves CM, Thomaz EB, de Sousa Queiroz R. A protocol to identify non-classical risk factors for preterm births: the Brazilian Ribeirão Preto and São Luís prenatal cohort (BRISA). **Reprod Health**. v.11, n.1, p.79. 2014. DOI: 10.1186/1742-4755-11-79.

SILVA, Ana Maria Rigo *et al.* Risk factors for pre-term birth in Londrina, Paraná State, Brazil. **Cadernos de Saúde Pública**, v. 25, p. 2125-2138, 2009.

SIMHAN, Hyagriv N., & CARITIS, Steve N. Prevention of preterm delivery. *New England Journal of Medicine*, v. 357, n. 5, p. 477-48, 2007. doi: 10.1056/NEJMra050435.

SOUSA, Derrijulie Siqueira *et al.* Morbidade em recém- nascidos prematuros de extremo baixo peso em unidade de terapia intensiva neonatal. **Rev. Bras. Saúde Matern. Infant.**, Recife, 17 (1): 149-157 jan-mar., 2017. DOI: 10.1590/1806-9304201700010000

SOUZA, Daniel Miranda Lopes de *et al.* Prevalência de prematuridade e fatores associados no estado do Rio Grande do Sul. **Brazilian Journal of Health Review**, v. 2, n. 5, p. 4052-4070, 2019.

STEYERBERG, Ewout W. **Applications of prediction models**. Springer New York, 2009.

SUFF, Natalie; STORY, Lisa; SHENNAN, Andrew. A previsão do parto prematuro: o que há de novo?. In: **Seminários em Medicina Fetal e Neonatal.**, p.27-32, WB Saunders, 2019.

SUN, Qi; ZOU, Xiaoxuan; YAN, Yousheng; ZHANG, Hongguang; WANG, Shuo; GAO, Yongmei; LIU, Haiyan; LIU, Shuyu; LU, Jianbo; YANG, Ying; MA, Xu. Machine learning-based prediction model of preterm birth using electronic health record. **Journal of Healthcare Engineering**, v. 2022, p. 9635526, 2022. DOI: 10.1155/2022/9635526.

SUN, Yanmin et al. Cost-sensitive boosting for classification of imbalanced data. **Pattern Recognition**, v. 40, n. 12, p. 3358–3378, 2007. DOI: 10.1016/j.patcog.2007.04.009.

SWAMY GK, Ostbye T, Skjaerven R. **Associação do parto prematuro com sobrevivência a longo prazo, reprodução e parto prematuro de próxima geração**. *Jama* 2008; 299: 1429–1436. DOI: 10.1001/jama.299.12.1429.

TINGLEFF T, Räisänen S, Vikanes Å et al. Associação entre país de nascimento materno e nascimento prematuro: um estudo de registro populacional de 910.752 partos. **Scand J Public Health** 2021; 49: 904–913. 20210215. DOI: 10.1177/1403494821992894.

TUON, Rogerio Antonio *et al.* Impacto do monitoramento telefônico de gestantes na prevalência da prematuridade e análise dos fatores de risco associados em Piracicaba, São Paulo, Brasil. **Cadernos de Saúde Pública**, v. 32, 2016.

United Nations. **Nações Unidas Transformando nosso mundo: A Agenda 2030 para o desenvolvimento sustentável** [2015]. Disponível em: <https://sustainabledevelopment.un.org/post2015/transformingourworld/>. Acesso em 25 de Julho de 2024.

VANIN, Luísa Krusser *et al.* Fatores de risco materno-fetais associados à prematuridade tardia. **Revista Paulista de Pediatria**, v. 38, 2019.

VON Elm E, Altman DG, Egger M, et al. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. **Lancet**. 2007;370(9596):1453–1457.

WEBER, Ann *et al.* Application of machine-learning to predict early spontaneous preterm birth among nulliparous non-Hispanic black and white women. **Ann Epidemiol**, v.28, n.11, p.783-789, 2018.

World Health Organization. **Born too soon: the global action report on preterm birth**.

Geneva: 2012. Disponível em:

https://www.who.int/pmnch/media/news/2012/201204_borntoosoon-report.pdf. Acesso em 30 de julho de 2024.

World Health Organization. **Preterm birth**. 2023. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/preterm-birth>. Acesso em 30 de julho de 2024.

World Health Organization. **Preterm birth: Key facts**. Geneva: WHO, 2019. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/preterm-birth>. Acesso em: 14 jul. 2024.

World Health Organization.; UNICEF; LONDON SCHOOL OF HYGIENE & TROPICAL MEDICINE. **Born Too Soon: Decade of Action on Preterm Birth**. Geneva: World Health Organization, 2023. Disponível em: <https://www.who.int/news/item/09-05-2023-152-million-babies-born-preterm-in-the-last-decade>. Acesso em: 9 jun. 2025.

XU, Haiqing *et al.* Time trends and risk factor associated with premature birth and infants deaths due to prematurity in Hubei Province, China from 2001 to 2012. **BMC pregnancy and childbirth**, v. 15, n. 1, p. 1-9, 2015.

XU, Xu K. *et al.* Fatores de risco associados ao nascimento prematuro entre filhos únicos após a tecnologia de reprodução assistida na Austrália 2007–2009 – um estudo retrospectivo de base populacional. **BMC pregnancy and childbirth**, v. 14, n. 1, pág. 1-11, 2014.

YU, Q. Y. et al. Predicting risk of preterm birth in singleton pregnancies using machine learning algorithms. **Frontiers in Big Data**, v. 7, 2024. Disponível em: <https://doi.org/10.3389/fdata.2024.1291196>. Acesso em: 04 maio 2026.

ZHOU, Zhi-Hua. **Aprendizado de máquina**. Springer Nature, 2021.

ZUGAIB, Marcelo. **Zugaib Obstetrícia**. 2º ed. Barueri, São Paulo: Manole, 2010.

ANEXOS

Os anexos contêm material de apoio utilizado para o desenvolvimento do trabalho que não foram construídos pela equipe de execução deste trabalho, tais como: questionários utilizados para extração das variáveis utilizadas, documentos com Pareceres de Aprovação do trabalho pelos Comitês de Ética em Pesquisa (CEP) de Ribeirão Preto e São Luís e Normas da revista.

ANEXO A- QUESTIONÁRIO PRÉ-NATAL ENTREVISTA



QUESTIONÁRIO DO PRÉ-NATAL ENTREVISTA

BLOCO A – DADOS DE IDENTIFICAÇÃO

1A. Número de identificação: _____

1ª casela: 1 Ribeirão Preto

2 São Luís

2ª casela: 1 Pré-natal

2 Nascimento

3 1º ano

3ª casela: M. Avaliação no pré-natal

A. Avaliação no nascimento RN 1

B. Avaliação no nascimento RN 2

C. Avaliação no nascimento RN 3

D. Avaliação no nascimento RN 4

4ª e 5ª caselas: QM. Questionário da mãe

QC. Questionário do RN

SC. Saliva da criança

CO. Cordão umbilical

6ª à 9ª. caselas: número seqüencial para cada cidade

NUMERO

2A. Cidade:

1. ☐ Ribeirão Preto2. ☐ São LuísCIDADEP

3A. Data da Entrevista (DD/MM/AAAA): __/__/____

DATAENTP

Entrevistador (a) : _____

4A. Nome completo da gestante (não abreviar):

NOME G

5A. Data de nascimento da gestante (DD/MM/AAAA):

__/__/____

DNGEST

6A. Idade da gestante __

99. ☐ Não sabeIDAGEG

7A. Data da ultrassonografia do recrutamento

(DD/MM/AAAA): __/__/____

DATUSR 99999999. ☐ Não sabe

8A. Idade gestacional pela ultrassonografia do recrutamento: __ semanas

IDGESTUSR 99. ☐ Não sabe9A. Idade gestacional atual (pela USG):
__ semanasIDGESTAT 99. ☐ Não sabe

BLOCO C – DADOS SOCIOECONÔMICOS E DEMOGRÁFICOS

1C. A sra. sabe ler e escrever?

1. ☐ Sim
 2. ☐ Não
 9. ☐ Não sabe

LERG

☐

2C. A sra. frequenta ou frequentou escola?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 7C**
 9. ☐ Não sabe

ESCOLG

☐

3C. Qual o último curso que a sra. frequentou ou frequenta?

1. ☐ Alfabetização de jovens e adultos
 2. ☐ Ensino fundamental ou 1o grau
 3. ☐ Ensino médio ou 2o grau
 4. ☐ Superior graduação incompleto **Passe para a questão 5C**
 5. ☐ Superior graduação completo **Passe para a questão 5C**
 8. ☐ Não se aplica
 9. ☐ Não sabe

CURSOG

☐

4C. Qual a série que a sra. frequenta ou até que série a sra. estudou?

1. ☐ Primeira
 2. ☐ Segunda
 3. ☐ Terceira
 4. ☐ Quarta
 5. ☐ Quinta
 6. ☐ Sexta
 7. ☐ Sétima
 8. ☐ Oitava
 88. ☐ Não se aplica
 99. ☐ Não sabe

SERIEG

☐☐

5C. A sra estava estudando quando ficou grávida?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 7C**
 8. ☐ Não se aplica
 9. ☐ Não sabe

ESTUDGRAV

☐

6C. A sra. parou de estudar porque ficou grávida?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

PAROUEST

☐

7C. Qual a situação conjugal atual da sra.?

1. ☐ Casada
2. ☐ União consensual (Mora junto)
3. ☐ Solteira
4. ☐ Separada/desquitada/divorciada
5. ☐ Viúva
9. ☐ Não sabe

SITCONG

☐

8C. Quantas pessoas vivem atualmente na casa onde a sra. mora? (Considere apenas as pessoas que estão morando na casa há pelo menos 3 meses, e que não são temporários, como um tio que está temporariamente vivendo com a sra. por menos de 3 meses ou visitantes). __

99. ☐ Não sabe

PESSOASP

☐☐

9C. A sra. mora atualmente com o marido ou companheiro?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

MORACOMPP

☐

10C. A sra. mora atualmente com filhos (biológicos ou não)?

1. ☐ Sim
2. ☐ Não **Passe para a questão 12C**
9. ☐ Não sabe

MORAFILHOP

☐

11C. Caso sim, com quantos filhos? __

88. ☐ Não se aplica
99. ☐ Não sabe

QTFILHOSP

☐☐

12C. De onde vem a água da casa usada para beber?

1. ☐ Rede pública/água encanada
2. ☐ Poço artesiano
3. ☐ Poço/cacimba
5. ☐ Outro _____
4. ☐ Rio/riacho/lagoa
9. ☐ Não sabe

AGUABEBER

☐

13C. Quantos cômodos têm na sua casa? _____ cômodos

99. ☐ Não sabe

COMODOS

☐☐

14C. Quantos cômodos servem como dormitório? _____ cômodos

9. ☐ Não sabe

DORMITOR

☐

15C. A sra. exerce alguma atividade remunerada dentro ou fora de casa?

1. ☐ Sim
2. ☐ Não **Passe para a questão 19C**
9. ☐ Não sabe

ATIVREMP

☐

16C. Qual a sua ocupação (o que faz atualmente no trabalho)?

88. ☐ Não se aplica

99. ☐ Não sabe

OCUPG

☐☐

17C. Qual a sua relação de trabalho?

1. ☐ Trabalha por conta própria

2. ☐ Assalariado ou empregado

3. ☐ Dono de empresa-empregador

4. ☐ Faz bico

8. ☐ Não se aplica

9. ☐ Não sabe

RELACAOP

☐

18C. A sra. parou de trabalhar porque ficou grávida?

1. ☐ Sim

2. ☐ Não

8. ☐ Não se aplica

9. ☐ Não sabe

PAROUTRAB

☐

19C. Quem é a pessoa da família com maior renda atualmente?
(considerar chefe da família aquele de maior renda)

1. ☐ A entrevistada **Passe para a questão 28C**

2. ☐ Companheiro

3. ☐ Mãe

4. ☐ Pai

5. ☐ Avó

6. ☐ Avô

7. ☐ Madrasta

8. ☐ Padrasto

9. ☐ Tia

10. ☐ Tio

11. ☐ Irmã

12. ☐ Irmão

13. ☐ Outro _____

99. ☐ Não sabe

CHEFEP

☐☐

20C. Qual o sexo da pessoa da família com maior renda?

1. ☐ Masculino

2. ☐ Feminino

8. ☐ Não se aplica

9. ☐ Não sabe

SEXOCHEFEP

☐

21C. Qual a idade da pessoa da família com maior renda (anos completos)? __

88. ☐ Não se aplica

99. ☐ Não sabe

IDCHEFEP

☐☐

22C. Essa pessoa sabe ler e escrever?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

LERCHEFEP

☐

23C. Essa pessoa frequenta ou frequentou escola?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 26C**
 8. ☐ Não se aplica
 9. ☐ Não sabe

ESCCHEFEP

☐

24C. Qual foi o último curso que essa pessoa frequentou ou frequenta?

1. ☐ Alfabetização de jovens e adultos
 2. ☐ Ensino fundamental ou 1o grau
 3. ☐ Ensino médio ou 2o grau
 4. ☐ Superior graduação incompleto **Passe para a questão 26C**
 5. ☐ Superior graduação completo **Passe para a questão 26C**
 8. ☐ Não se aplica
 9. ☐ Não sabe

CURSOCHFEP

☐

25C. Qual a série que essa pessoa frequenta ou até que série frequentou?

1. ☐ Primeira
 2. ☐ Segunda
 3. ☐ Terceira
 4. ☐ Quarta
 5. ☐ Quinta
 6. ☐ Sexta
 7. ☐ Sétima
 8. ☐ Oitava
 88. ☐ Não se aplica
 99. ☐ Não sabe

SERIECHFEP

☐ ☐

26C. Qual a ocupação atual (ou no que trabalha) a pessoa com a maior renda da família? (Descreva a ocupação. Caso seja aposentado, colocar a última atividade que exerceu).

88. ☐ Não se aplica
 99. ☐ Não sabe

OCUPCHFEP

☐ ☐

38C. Grau de Instrução da pessoa com maior renda

Analfabeto/Primário incompleto/ Até 3ª Série Fundamental	0
Primário completo/ Até 4ª Série Fundamental/Ginasial incompleto	1
Ginasial completo/ Fundamental completo/Colegial incompleto	2
Colegial completo/ Médio completo/Superior incompleto	4
Superior completo	8

INSTRUCAOP

☐**BLOCO D – HÁBITOS DE VIDA****Agora vamos conversar um pouco sobre o consumo de bebida alcoólica.**

1D. Durante a gravidez, a sra. tomou cerveja?

1. ☐ Sim2. ☐ Não **Passe para a questão 5D**9. ☐ Não sabe

CERVEJAG

☐

2D.Quantos dias por semana? _

8. ☐ Não se aplica9. ☐ Não sabe

DIACERVG

☐

3D.Quanto tomava por dia (número de vasilhas) __

88. ☐ Não se aplica99. ☐ Não sabe

QTCERVG

☐

4D.Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)2. ☐ Lata (350ml)3. ☐ Garrafa pequena (300ml) – long neck4. ☐ Garrafa (600-720ml)5. ☐ Outro8. ☐ Não se aplica9. ☐ Não sabe

TIPOCERVG

☐

5D.Durante a gravidez a sra tomou vinho?

1. ☐ Sim2. ☐ Não **Passe para a questão 9D**9. ☐ Não sabe

VINHOG

☐

6D. Quantos dias por semana? _

8. ☐ Não se aplica9. ☐ Não sabe

DIAVINHOG

☐

7D.Quanto tomava por dia (número de vasilhas) __

88. ☐ Não se aplica99. ☐ Não sabe

QTVINHOG

☐

8D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)
2. ☐ Cálice ou taça (400 ml)
3. ☐ Garrafa pequena (300ml)
4. ☐ Garrafa (600-720ml)
5. ☐ Outro
8. ☐ Não se aplica
9. ☐ Não sabe

TIPOVINHOG

☐

9D. Durante a gravidez a sra. tomou algum outro tipo de bebida como uísque, vodka, gim, rum, cachaça, caipirinha, batida?

1. ☐ Sim
2. ☐ Não **Passe para a questão 13D**
9. ☐ Não sabe

DESTG

☐

10D. Quantos dias por semana? __

8. ☐ Não se aplica
9. ☐ Não sabe

DIADDESTG

☐

11D. Quanto tomava por dia (número de vasilhas) __

88. ☐ Não se aplica
99. ☐ Não sabe

QTDESTG

☐ ☐

12D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)
2. ☐ Cálice, taça (400 ml)
3. ☐ Martelo, copo de pinga (100ml)
4. ☐ Lata (350ml) retirar
5. ☐ Garrafa pequena (300ml)
6. ☐ Garrafa (600-720ml)
7. ☐ Outro
8. ☐ Não se aplica
9. ☐ Não sabe

TIPODESTG

☐

13D. Durante a gravidez com que frequência a sra. consumiu cinco ou mais unidades de bebida alcoólica em uma única ocasião?

1. ☐ Todos os dias
2. ☐ Quase todos os dias
3. ☐ 1 a 4 vezes na semana
4. ☐ 1 a 3 vezes por mês
5. ☐ Raramente
6. ☐ Nunca
8. ☐ Não se aplica
9. ☐ Não sabe

FREQBEB

☐

Agora vamos conversar um pouco sobre o hábito de fumar.

14D. A sra. fumou ou fuma durante esta gravidez?

1. ☐ Sim
 2. ☐ Não **Passe para questão 16D**
 9. ☐ Não sabe

FUMOGRAVP

☐

15D. Quantos cigarros a sra. fuma ou fumou por dia durante a gravidez? ____

88. ☐ Não se aplica
 99. ☐ Não sabe

QTFUMOP

☐ ☐

Agora vamos conversar um pouco sobre o consumo de bebidas com cafeína

16D. Desde que ficou grávida a sra. tem tomado café (café preto, café instantâneo tipo nescafé ou capuccino)?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 20D**
 9. ☐ Não sabe

CAFE

☐

17D. Quantos dias por semana a sra. toma café? __

8. ☐ Não se aplica
 9. ☐ Não sabe

DIASCAFE

☐

18D. Quantas vezes por dia a sra. toma café? __

88. ☐ Não se aplica
 99. ☐ Não sabe

VEZESCAFE

☐ ☐

19D. Qual o tipo de vasilha em que a sra. costuma tomar café?

1. ☐ xícara de chá
 2. ☐ xícara de cafézinho
 3. ☐ meia taça
 4. ☐ copo comum – 200 ml
 5. ☐ outro _____
 8. ☐ Não se aplica
 9. ☐ Não sabe

TIPOCAFE

☐

20D. Desde que ficou grávida a sra. tem tomado chá (chá mate, chá preto, chá verde, chá amarelo, chá branco ou outro)?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 24D**
 9. ☐ Não sabe

CHA

☐

21D. Quantos dias por semana a sra. toma chá? __

8. ☐ Não se aplica
 9. ☐ Não sabe

DIASCHA

☐

22D. Quantas vezes por dia a sra. toma chá? __

88. ☐ Não se aplica
 99. ☐ Não sabe

VEZESCHA

☐ ☐

23D. Qual o tipo de vasilha em que a sra. costuma tomar chá?

1. ☐ xícara de chá
2. ☐ xícara de cafezinho
3. ☐ meia taça
4. ☐ copo comum – 200 ml
5. ☐ outro _____
8. ☐ Não se aplica
9. ☐ Não sabe

TIPOCHA

☐

24D. Desde que ficou grávida a sra. tem tomado refrigerantes (coca-cola ou Pepsi-cola)?

1. ☐ Sim
2. ☐ Não **Passe para a questão 28D**
9. ☐ Não sabe

REFRI

☐

25D. Quantos dias por semana a sra. toma refrigerante? _

8. ☐ Não se aplica
9. ☐ Não sabe

DIASREFRI

☐

26D. Quantas vezes por dia a sra. toma refrigerante? _ _

88. ☐ Não se aplica
99. ☐ Não sabe

VEZESREFRI

☐ ☐

27D. Qual o tipo de vasilha em que a sra. costuma tomar refrigerante?

1. ☐ xícara de chá
2. ☐ xícara de cafezinho
3. ☐ meia taça
4. ☐ copo comum – 200 ml
5. ☐ outro _____
8. ☐ Não se aplica
9. ☐ Não sabe

TIPOREFRI

☐

28D. Desde que ficou grávida a sra. tem tomado bebidas com chocolate?

1. ☐ Sim
2. ☐ Não **Passe para a questão 32D**
9. ☐ Não sabe

CHOCOLATE

☐

29D. Quantos dias por semana a sra. toma bebidas com chocolate? _

8. ☐ Não se aplica
9. ☐ Não sabe

DIASCHOCO

☐

30D. Quantas vezes por dia a sra. toma bebidas com chocolate? _ _

88. ☐ Não se aplica
99. ☐ Não sabe

VEZESCHOCO

☐ ☐

31D. Qual o tipo de vasilha em que a sra. costuma tomar bebidas com chocolate?

1. ☐ xícara de chá
2. ☐ xícara de cafezinho
3. ☐ meia taça
4. ☐ copo comum – 200 ml
5. ☐ outro _____
8. ☐ Não se aplica
9. ☐ Não sabe

TIPOCHOCO

☐

32D. Desde que ficou grávida a sra. tem comido alimentos com chocolate?

1. ☐ Sim
2. ☐ Não **Passe para a questão 1E**
9. ☐ Não sabe

ALCHOCO

☐

33D. Quantos dias por semana a sra. come chocolate? __

8. ☐ Não se aplica
9. ☐ Não sabe

DIASALCHO

☐

34D. Quantas vezes por dia a sra. come chocolate? __

88. ☐ Não se aplica
99. ☐ Não sabe

VEZESALCHO

☐ ☐

BLOCO E – DADOS DO COMPANHEIRO

Caso o companheiro seja a pessoa com maior renda ou não more junto com a gestante, passe para a questão 1F.

1E. Qual a idade do companheiro atual? __

88. ☐ Não se aplica – não tem companheiro atual
99. ☐ Não sabe

IDCOMPP

☐ ☐

2E. O seu companheiro sabe ler e escrever?

1. ☐ Sim
2. ☐ Não
8. ☐ Não se aplica
9. ☐ Não sabe

LERCOMPP

☐

3E. O seu companheiro frequenta ou frequentou escola?

1. ☐ Sim
2. ☐ Não **Passe para a questão 6E**
8. ☐ Não se aplica
9. ☐ Não sabe

ESCCOMPP

☐

4E. Qual o último curso que seu companheiro frequentou ou frequenta?

1. ☐ Alfabetização de jovens e adultos
2. ☐ Ensino fundamental ou 1o grau
3. ☐ Ensino médio ou 2o grau
4. ☐ Superior graduação incompleto **Passe para a questão 6E**
5. ☐ Superior graduação completo **Passe para a questão 6E**
8. ☐ Não se aplica
9. ☐ Não sabe

CURSOCOMPP

☐

5E. Qual a série que seu companheiro frequenta ou até que série estudou?

1. ☐ Primeira
2. ☐ Segunda
3. ☐ Terceira
4. ☐ Quarta
5. ☐ Quinta
6. ☐ Sexta
7. ☐ Sétima
8. ☐ Oitava
88. ☐ Não se aplica
99. ☐ Não sabe

SERIECOMPP

☐☐

6E. O seu companheiro está trabalhando no momento?

1. ☐ Sim
2. ☐ Não **Passe para a questão 1F**
8. ☐ Não se aplica
9. ☐ Não sabe

TRABCOMPP

☐

7E. Qual a ocupação atual (ou no que trabalha) o seu companheiro atual? (Descreva a ocupação. Caso seja aposentado, colocar a última atividade que exerceu).

-
88. ☐ Não se aplica
 99. ☐ Não sabe

OCUPCOMPP

☐☐

8E. Qual a relação de trabalho do seu companheiro?

1. ☐ Trabalha por conta própria
2. ☐ Assalariado ou empregado
3. ☐ Dono de empresa-empregador
4. ☐ Faz bico
8. ☐ Não se aplica
9. ☐ Não sabe

RELCOMPP

☐

BLOCO F – DADOS DA SAÚDE SEXUAL E REPRODUTIVA

1F. Com quantos parceiros do sexo masculino a sra. já teve relação sexual? __

99. ☐ Não sabe

NPARCEIROS

2F. Algum parceiro sexual masculino já disse à sra. que teve doença sexualmente transmissível (doença venérea)?

1. ☐ Sim

2. ☐ Não **Passe para a questão 4F**

9. ☐ Não sabe

DST

☐

3F. Caso o parceiro já tenha lhe dito, qual o nome da (s) doença(s)?

88. ☐ Não se aplica

99. ☐ Não sabe

NOMEDST

4F. Algum dos parceiros sexuais masculinos já lhe disse ser HIV positivo ou ter AIDS?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HIV

☐

5F. A sra. já foi chamada por algum serviço de saúde por ter tido relação sexual com alguma pessoa com doença sexualmente transmissível?

1. ☐ Sim

2. ☐ Não **Passe para a questão 7F**

9. ☐ Não sabe

SSDST

☐

6F. Caso tenha sido chamada, qual o nome da (s) doença(s)?

88. ☐ Não se aplica

99. ☐ Não sabe

NOMESSDST

7F. No ano anterior a essa gravidez, a sra. estava usando algum método contraceptivo?

1. ☐ Sim

2. ☐ Não **Passe para a questão 1G**

9. ☐ Não sabe

METODO

8F. Caso sim, qual o método contraceptivo que a sra. estava usando? (pode marcar mais de uma opção)?

1. ☐ Camisinha somente

2. ☐ Camisinha/ espermicida

3. ☐ Diafragma somente

4. ☐ Diafragma/ Espermicida

5. ☐ Espermicida somente

6. ☐ Esponja

7. ☐ DIU
 8. ☐ Pílulas orais
 9. ☐ Tabela/ritmo (calendário, temperatura)
 10. ☐ Coito interrompido
 11. ☐ Ducha vaginal
 12. ☐ Injetáveis
 13. ☐ Outros
 88. ☐ Não se aplica
 99. ☐ Não sabe

QUALMETODO

☐ ☐

BLOCO G – CARACTERÍSTICAS DA GESTAÇÃO ATUAL, DO PRÉ-NATAL E HIPERTENSÃO

Pergunte se a gestante dispõe do cartão da gestante e se está de posse do mesmo. Confirme as respostas no cartão

1G. A sra tem cartão da gestante?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica – não fez pré-natal
 9. ☐ Não sabe

CARTAOP

☐

2G. Qual a data da sua última menstruação (DD/MM/AAAA)?

__/__/____

99999999. ☐ Não sabe

DUMP

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

3G. Caso não saiba o dia, informar o mês e o ano (MM/AAAA)?

888888. ☐ Não se aplica

999999. ☐ Não sabe

MESANOUMP

☐ ☐ ☐ ☐ ☐ ☐

4G. Qual o seu peso antes de engravidar? ____, _ kg

9999. ☐ Não sabe

PESOANTESP

☐ ☐ ☐ ☐

5G. Qual a sua altura antes de engravidar? ____, _ cm

9999. ☐ Não sabe

ALTURAANTP

☐ ☐ ☐ ☐

6G. A sra está fazendo pré-natal?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 11G**
 9. ☐ Não sabe

PNP

☐

7G. Qual a data da primeira consulta pré-natal
 (DD/MM/AAAA)? __/__/____

88888888. ☐ Não se aplica

99999999. ☐ Não sabe

DT1CPNP

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

8G. Em que mês de gravidez a sra. iniciou as consultas de
 pré-natal? __

88. ☐ Não se aplica

99. ☐ Não sabe

MES1CPNP

☐ ☐

9G. Quantas consultas de pré-natal a sra. fez no 1º trimestre de gestação? __

88. ☐ Não se aplica

99. ☐ Não sabe

QTCPN1TP

☐ ☐

10G. Onde a sra. está realizando o pré-natal nesta gestação?

1. ☐ SUS (posto de saúde, hospital universitário)

2. ☐ Plano de Saúde

3. ☐ Particular

8. ☐ Não se aplica

9. ☐ Não sabe

LOCALPN

☐

11G. A sra. tem hipertensão (pressão alta) fora da gestação diagnosticada por médico ou enfermeiro?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HIPERTF

☐

12G. A sra. tem hipertensão (pressão alta) na gestação atual?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HIPERTG

☐

13G. A mãe da sra. tem hipertensão crônica (pressão alta)?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HIPERTM

☐

14G. A mãe da sra. teve hipertensão (pressão alta) na gestação?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HIPERTMG

☐

15G. O pai da sra. tem hipertensão crônica (pressão alta)?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HIPERTP

☐

16G. O(s) irmão(s) ou irmã(s) da sra. têm hipertensão crônica (pressão alta)?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

Passe para a questão 18G

HIPERTH

☐

17G. Caso sim, quantos irmãos ou irmãs têm hipertensão crônica (pressão alta)? _____ irmãos ou irmãs

8. ☐ Não se aplica

9. ☐ Não sabe

HIPERTHQ

☐

18G. A(s) sua(s) irmãs tiveram hipertensão (pressão alta) na gestação?

1. ☐ Sim

2. ☐ Não

Passe para o bloco H

9. ☐ Não sabe

HIPERTIG

☐

19G. Caso sim, quantas irmãs tiveram hipertensão (pressão alta) na gestação? _____ irmãs

8. ☐ Não se aplica

9. ☐ Não sabe

HIPERTIGQ

☐

BLOCO H – EXPOSIÇÃO A METAIS

1H. Alguém que vive na sua casa trabalha em fábrica de tintas, adubos, vidros, cerâmica, plástico, conserto de baterias ou galvanoplastia (revestimento de metais)?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

FABTINTA

☐

2H. Tem alguma fábrica ou indústria perto da casa da sra.?

1. ☐ Sim

2. ☐ Não

Passe para a questão 4H

9. ☐ Não sabe

FABRICA

☐

3H. Caso sim, qual é a distância da fábrica para a casa da sra.?
(estimativa por quarteirões) __

00. ☐ No mesmo quarteirão

88. ☐ Não se aplica

99. ☐ Não sabe

DISTFABRICA

☐ ☐

4H. A sra. reformou ou pintou a sua casa recentemente ?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

PINTOUCASA

☐

5H. A sra. tem alguma atividade de lazer/hobby como pintura, cerâmica, pesca ou tiro com arma de fogo?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

LAZER

☐

6H. A sra. come peixe, marisco ou camarão?

1. ☐ Sim

2. ☐ Não

Passe para a questão 8H

9. ☐ Não sabe

PEIXE

☐

7H.Caso sim, qual é a frequência?

1. ☐ Uma vez na semana
2. ☐ Mais de uma vez na semana
3. ☐ Uma vez ao mês
4. ☐ 2 a 3 vezes ao mês
5. ☐ Diariamente
8. ☐ Não se aplica
9. ☐ Não sabe

FREQPEIXE

☐ ☐

8H.A sra. consumiu peixe, camarão ou marisco nos últimos 2 dias?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

PEIXE2D

☐

9H.A sra. toma remédios, chás, energéticos, vitaminas?

1. ☐ Sim
2. ☐ Não **Passe para a questão 25H**
9. ☐ Não sabe

REM

☐

Caso tome, qual o nome do remédio e qual o mês de início e de término? Se continua usando, anotar que está em uso?

10H.Remédio

-
88. ☐ Não se aplica
 99. ☐ Não sabe

REM1

☐ ☐

11H. Mês da gestação de início__

00. ☐ Usa desde antes da gravidez
88. ☐ Não se aplica
99. ☐ Não sabe

IREM1

☐ ☐

12H. Mês da gestação de término _

10. ☐ Ainda usa
88. ☐ Não se aplica
99. ☐ Não sabe

TREM1

☐ ☐

13H.Remédio

-
88. ☐ Não se aplica
 99. ☐ Não sabe

REM2

☐ ☐

14H. Mês da gestação de Início _		
00. ☐ Usa desde antes da gravidez		
88. ☐ Não se aplica		
99. ☐ Não sabe		☐ ☐
15H. Mês da gestação de término _	IREM2	
10. ☐ Ainda usa		
88. ☐ Não se aplica		
99. ☐ Não sabe	TREM2	☐ ☐
16H. Remédio		
88. ☐ Não se aplica		
99. ☐ Não sabe	REM3	☐ ☐
17H. Mês da gestação de Início _		
00. ☐ Usa desde antes da gravidez		
88. ☐ Não se aplica		
99. ☐ Não sabe	IREM3	☐ ☐
18H. Mês da gestação de término _		
10. ☐ Ainda usa		
88. ☐ Não se aplica		
99. ☐ Não sabe	TREM3	☐ ☐
19H. Remédio		
88. ☐ Não se aplica		
99. ☐ Não sabe	REM4	☐ ☐
20H. Mês da gestação de Início _		
00. ☐ Usa desde antes da gravidez		
88. ☐ Não se aplica		
99. ☐ Não sabe	IREM4	☐ ☐
21H. Mês da gestação de término _		
10. ☐ Ainda usa		
88. ☐ Não se aplica		
99. ☐ Não sabe	TREM4	☐ ☐
22H. Remédio		
88. ☐ Não se aplica		
99. ☐ Não sabe	REM5	☐ ☐
23H. Mês da gestação de Início _		
00. ☐ Usa desde antes da gravidez		
88. ☐ Não se aplica		
99. ☐ Não sabe	IREM5	☐ ☐

24H. Mês da gestação de término _		
10. ☐ Ainda usa		
88. ☐ Não se aplica		
99. ☐ Não sabe		☐ ☐
	TREM5	
25H. A sra. masca chicletes?		
1. ☐ Sim		
2. ☐ Não Passe para a questão 27H		
9. ☐ Não sabe		☐
	CHICLETE	
26H. Caso sim, com que frequencia?		
1. ☐ Diariamente		
2. ☐ Ocasionalmente		
8. ☐ Não se aplica		
9. ☐ Não sabe		☐
	FREQCHIC	
27H. A sra. tem dentes com restaurações (obturações) de metal (amálgama)?		
1. ☐ Sim		
2. ☐ Não Passe para a questão 29H		
9. ☐ Não sabe		☐
	RESTAURA	
28H. Caso sim, há quanto tempo? __ anos		
88. ☐ Não se aplica		
99. ☐ Não sabe		☐ ☐
	TRESTAURA	
29H. A sra. usa tintura, tonalizante ou descolorante de cabelo?		
1. ☐ Sim		
2. ☐ Não Passe para a questão 31H		
9. ☐ Não sabe		☐
	TINTURA	
30H. Caso sim, com que frequencia?		
1. ☐ Diariamente		
2. ☐ Ocasionalmente		
8. ☐ Não se aplica		
9. ☐ Não sabe		☐
	FREQTINT	
31H. A sra. tem horta em casa?		
1. ☐ Sim		
2. ☐ Não Passe para o BLOCO I		
9. ☐ Não sabe		☐
	HORTA	
32H. Caso sim, a sra. come as verduras da sua horta?		
1. ☐ Sim		
2. ☐ Não		
8. ☐ Não se aplica		
9. ☐ Não sabe		☐
	COMEVERD	

33H. Qual a água que a sra. usa para irrigar a sua horta?

1. ☐ Água da torneira
2. ☐ Água do poço
3. ☐ Outra: _____
8. ☐ Não se aplica
9. ☐ Não sabe

AGUAHORTA

☐

BLOCO I – SAÚDE BUCAL

Agora vamos fazer algumas perguntas sobre sua saúde bucal (condições de seus dentes e gengivas) e sobre a utilização de serviços de odontologia.

11. Tente se lembrar como era a sua saúde bucal antes da gravidez. Antes desta gravidez, como a sra. considerava a sua saúde bucal?

1. ☐ Excelente
2. ☐ Muito boa
3. ☐ Boa
4. ☐ Razoável
5. ☐ Ruim
9. ☐ Não sabe

SBUCAL

☐

21. Desde que começou esta gravidez, a sra. percebeu alguma mudança na condição de suas gengivas (ficaram inflamadas (inchadas) ou passaram a sangrar durante a escovação ou mesmo sem tocar nelas)?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

GENGIVA

☐

31. Desde que começou esta gravidez a sra. percebeu alguma mudança na condição de seus dentes (apareceram novas cáries nos dentes ou as cáries antigas ficaram piores)?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

DENTE

☐

41. Antes dessa gestação, algum dentista alguma vez disse que a sra. precisaria fazer tratamento para a gengiva?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

DENTISTA

☐

BLOCO K – ESCORE DE BLOCK (GORDURAS)

Pense agora a respeito dos seus hábitos alimentares no último ano. Aproximadamente, com que frequência você consome cada um dos alimentos abaixo. (Se não consumir, marcar a primeira opção).

Alimentos	1 vez ou menos ao mês	2 a 3 vezes ao mês	1 a 2 vezes por semana	3 a 4 vezes por semana	5 ou mais vezes por semana		
1K. Hambúrguer, cheeseburger ou bauru	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG1	☐
2K. Bife ou carne assada	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG2	☐
3K. Frango frito	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG3	☐
4K. Salsicha, cachorro quente, salame	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG4	☐
5K. Presunto, mortadela e embutidos	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG5	☐
6K. Maionese comum	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG6	☐
7K. Margarina ou Manteiga	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG7	☐
8K. Ovos	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG8	☐
9K. Bacon ou linguiça	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG9	☐
10K. Queijos e requeijão	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG10	☐
11K. Leite integral	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG11	☐
12K. Batata frita	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG12	☐
13K. Salgadinhos de pacote, pipoca	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG13	☐
14K. Sorvetes	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG14	☐
15K. Bolachas, doces de padaria, bolos, pastéis, massas folhadas	0. ☐	1. ☐	2. ☐	3. ☐	4. ☐	BLOCKG15	☐

BLOCO L – EXAME CLÍNICO

1L. Peso (medido): ____, _Kg

2L. Percentual de gordura corporal: ____, _%

3L. Altura (medida): ____, _cm

4L. Pressão arterial sistólica 1ª. medida: ____ mmHg

5L. Pressão arterial sistólica 2ª. medida: ____ mmHg

6L. Pressão arterial diastólica 1ª. medida: ____ mmHg

7L. Pressão arterial diastólica 2ª. medida: ____ mmHg

8L. Pulso: ____ batimentos por minuto

PESO

PGORDURA

ALTURA

PAS1

PAS2

PAD1

PAD2

PULSO

MATERIAL BIOLÓGICO E EXAMES DA GESTANTE

9L. Foi feita coleta de sangue?

1. ☐ Sim2. ☐ Não

SANGUE

☐

10L. Foi feita coleta de urina?

1. ☐ Sim2. ☐ Não

URINA

☐

11L. Foi feita coleta de secreção vaginal?

1. ☐ Sim2. ☐ Não

SECVAG

☐

12L. Foi feito exame odontológico?

1. ☐ Sim2. ☐ Não

ODONTO

☐

13L. Foi feita ultrassonografia?

1. ☐ Sim2. ☐ Não

USGP

☐

ANEXO B – QUESTIONÁRIO PRÉ-NATAL AUTOAPLICADO

BLOCO U – CONSUMO DE DROGAS

1U. Você utilizou ou tem utilizado alguma droga (maconha, merla, cocaína, crack, outras) nos três meses antes ou durante a gestação atual?

1. ☐ Sim
 2. ☐ Não **Se você marcou esta opção, não precisa responder às demais questões**
 9. ☐ Não sabe

DROGAS ☐

Caso sim, qual a droga utilizada e quantidade?

Nos 3 meses antes da gravidez	Não usou 0	Diária mente 1	Semanal mente 2	Mensal mente 3	Rara mente 4	Não sabe 9
2U. Maconha	☐	☐	☐	☐	☐	☐
3U. Merla	☐	☐	☐	☐	☐	☐
4U. Cocaína	☐	☐	☐	☐	☐	☐
5U. Crack	☐	☐	☐	☐	☐	☐
6U. Outras	☐	☐	☐	☐	☐	☐
Nos 3 primeiros meses de gravidez	Não usou 0	Diária mente 1	Semanal mente 2	Mensal mente 3	Rara mente 4	Não sabe 9
7U. Maconha	☐	☐	☐	☐	☐	☐
8U. Merla	☐	☐	☐	☐	☐	☐
9U. Cocaína	☐	☐	☐	☐	☐	☐
10U. Crack	☐	☐	☐	☐	☐	☐
11U. Outras	☐	☐	☐	☐	☐	☐
Do 4º. mês em diante	Não usou 0	Diária mente 1	Semanal mente 2	Mensal mente 3	Rara mente 4	Não sabe 9
12U. Maconha	☐	☐	☐	☐	☐	☐
13U. Merla	☐	☐	☐	☐	☐	☐
14U. Cocaína	☐	☐	☐	☐	☐	☐
15U. Crack	☐	☐	☐	☐	☐	☐
16U. Outras	☐	☐	☐	☐	☐	☐

MACONHA ☐

MERLA ☐

COCAINA ☐

CRACK ☐

OUTRAS ☐

MACONHA1T ☐

MERLA1T ☐

COCAINA1T ☐

CRACK1T ☐

OUTRAS1T ☐

MACONHA4 ☐

MERLA4 ☐

COCAINA4 ☐

CRACK4 ☐

OUTRAS4 ☐

ANEXO C – QUESTIONÁRIO NASCIMENTO-MÃE



QUESTIONÁRIO DO NASCIMENTO - MÃE

BLOCO A – DADOS DE IDENTIFICAÇÃO

1A. Número de identificação: _____

1ª casela: 1 Ribeirão Preto

2 São Luís

2ª casela: 1 Pré-natal

2 Nascimento

3 1º ano

3ª casela: M. Avaliação no pré-natal

A. Avaliação no nascimento RN 1

B. Avaliação no nascimento RN 2

C. Avaliação no nascimento RN 3

D. Avaliação no nascimento RN 4

4ª e 5ª caselas: QM. Questionário da mãe

QC. Questionário do RN

SC. Saliva da criança

CO. Cordão umbilical

6ª à 9ª. caselas: número seqüencial para cada cidade

NUMERO

2A. Cidade:

1. ☐ Ribeirão Preto2. ☐ São LuísCIDADE

3A. Coorte

1. ☐ Iniciada no Pré-natal2. ☐ Iniciada no NascimentoCOORTE

4A. Data da Entrevista (DD/MM/AAAA): __/__/____

DATAENT

Entrevistador (a) : _____

5A. Hospital de Nascimento:

SÃO LUÍS

1. ☐ HU Materno- Infantil2. ☐ Benedito Leite3. ☐ Marly Sarney4. ☐ Santa Casa5. ☐ Maria do Amparo6. ☐ N Sra. da Penha7. ☐ Clínica São Marcos8. ☐ Clínica Luíza Coelho9. ☐ Hospital S Domingos10. ☐ Hospital Aliança11. ☐ Clínica São José

RIBEIRÃO PRETO

12. ☐ Hospital das Clínicas13. ☐ Hospital Ribeirânia14. ☐ Hospital São Lucas15. ☐ Hospital Santa Lydia16. ☐ Hospital Santa Casa17. ☐ Mater18. ☐ H. Sinhá Junqueira19. ☐ Hospital São PauloHOSPITAL

6A. Nome completo da mãe do RN (não abreviar):

NOMEMAE

1

7A. Data de nascimento da mãe do RN (DD/MM/AAAA):

__/__/____

DNMAE

8A. Idade da mãe do RN __

99. ☐ Não sabeIDADEMAE

9A. Qual a idade do pai do bebê? __

99. ☐ Não sabeIDADEPAI

BLOCO C – DADOS SOCIOECONÔMICOS E DEMOGRÁFICOS

1C. A sra. sabe ler e escrever?

1. ☐ Sim2. ☐ Não9. ☐ Não sabe

LERMAE

☐

2C. A sra. frequenta ou frequentou escola?

1. ☐ Sim2. ☐ Não **Passe para a questão 6C**9. ☐ Não sabe

ESCOLMAE

☐

3C. A sra. ainda estuda ?

1. ☐ Sim2. ☐ Não8. ☐ Não se aplica9. ☐ Não sabe

MAEESTUDA

☐

4C. Qual foi o último curso que a sra frequentou ou frequenta?

1. ☐ Alfabetização de jovens e adultos2. ☐ Ensino fundamental ou 1o grau3. ☐ Ensino médio ou 2o grau4. ☐ Superior graduação incompleto **Passe para a questão 6C**5. ☐ Superior graduação completo **Passe para a questão 6C**8. ☐ Não se aplica9. ☐ Não sabe

CURSOMAE

☐

5C. Até que série a sra. frequentou ou ainda frequenta?

1. ☐ Primeira2. ☐ Segunda3. ☐ Terceira4. ☐ Quarta5. ☐ Quinta6. ☐ Sexta7. ☐ Sétima8. ☐ Oitava88. ☐ Não se aplica99. ☐ Não sabe

SERIEMAE

☐☐

6C. Qual a cor da sua pele?

1. ☐ branca2. ☐ preta/negra3. ☐ parda/mulata/cabocla/morena4. ☐ amarelo/oriental5. ☐ indígena9. ☐ não sabe

CORMAE

☐

7C. Qual a situação conjugal atual da sra.?

1. ☐ Casada
2. ☐ União consensual (Mora junto)
3. ☐ Solteira
4. ☐ Separada/desquitada/divorciada
5. ☐ Viúva
9. ☐ Não sabe

SITCONMAE

☐

8C. Quantas pessoas vivem atualmente na casa onde a sra. mora? (Considere apenas as pessoas que estão morando na casa há pelo menos 3 meses, e que não são temporários, como um tio que está temporariamente vivendo com a sra. por menos de 3 meses ou visitantes). __

99. ☐ Não sabe

PESSOAS

☐ ☐

9C. A sra. mora atualmente com o marido ou companheiro?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

MORACOMP

☐

10C. A sra. mora atualmente com filhos (biológicos ou não)?

1. ☐ Sim
2. ☐ Não **Passe para a questão 12C**
9. ☐ Não sabe

MORAFILHO

☐

11C. Caso sim, com quantos filhos? __

88. ☐ Não se aplica
99. ☐ Não sabe

QTFILHOS

☐ ☐

12C. A sra. mora atualmente com outros familiares?

1. ☐ Sim
2. ☐ Não **Passe para a questão 14C**
9. ☐ Não sabe

MORAFAM

☐

13C. Caso sim, com quantos familiares? __

88. ☐ Não se aplica
99. ☐ Não sabe

QTFAM

☐ ☐

14C. A sra. mora atualmente com outras pessoas que não são familiares?

1. ☐ Sim
2. ☐ Não **Passe para a questão 16C**
9. ☐ Não sabe

MORANFAM

☐

15C. Caso sim, com quantos não familiares? __

88. ☐ Não se aplica
99. ☐ Não sabe

QTNFAM

☐ ☐

16C. A sra. tem alguma religião ou culto?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

TEMRELIG

☐

17C. Caso tenha alguma religião, qual é a sua religião?

1. ☐ Católica
2. ☐ Evangélica. Ex: Batista, Assembléia de Deus, Bethesda, Universal, Adventista, Testemunha de Jeová, Luterana.
3. ☐ Espírita/Kardecista
4. ☐ Umbanda/Candomblé
5. ☐ Judaica
6. ☐ Orientais. Ex: Budista
7. ☐ Outra, Qual? _____
88. ☐ Não se aplica
99. ☐ Não sabe

RELIGIAO

☐ ☐

18C. A sra. exerce alguma atividade remunerada dentro ou fora de casa?

1. ☐ Sim
2. ☐ Não **Passe para a questão 26C**
9. ☐ Não sabe

ATIVREM

☐

CASO NÃO TENHA ALGUMA ATIVIDADE REMUNERADA OU NÃO SAIBA, PASSE PARA A QUESTÃO 26C

19C. Qual a sua ocupação (o que faz atualmente no trabalho)?

88. ☐ Não se aplica
99. ☐ Não sabe

OCUPMAE

☐ ☐

20C. Qual a sua relação de trabalho?

1. ☐ Trabalha por conta própria
2. ☐ Assalariado ou empregado
3. ☐ Dono de empresa-empregador
4. ☐ Faz bico
8. ☐ Não se aplica
9. ☐ Não sabe

RELACAO

☐

21C. Quantos dias por semana a sra. trabalhava com remuneração durante a gestação? _

8. ☐ Não se aplica
9. ☐ Não sabe

DIATRAB

☐

22C. Nos dias de trabalho remunerado durante a gestação, quantas horas por dia a sra. trabalhava? __

88. ☐ Não se aplica
99. ☐ Não sabe

HORATRAB

☐ ☐

23C. Durante o seu trabalho, a sra. tinha que ficar em pé a maior parte do tempo?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

TRABPE

☐

24C. Durante o seu trabalho, a sra. tinha que levantar coisas pesadas?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

LEVPESO

☐

25C. Há quantas semanas atrás a sra. parou de trabalhar? __

00. ☐ menos de 1 semana
 88. ☐ Não se aplica
 99. ☐ Não sabe

PAROUTRAB

☐ ☐

26C. Na sua casa, quem faz o trabalho de casa para sua família?

1. ☐ A sra. faz todo trabalho
 2. ☐ A sra. faz parte do trabalho
 3. ☐ Outra pessoa
 9. ☐ Não sabe

TRABCASA

☐

27C. Quem é a pessoa da família com maior renda atualmente? (considerar chefe da família aquele de maior renda)

1. ☐ A entrevistada **Passe para a questão 36C**
 2. ☐ Companheiro
 3. ☐ Mãe
 4. ☐ Pai
 5. ☐ Avó
 6. ☐ Avô
 7. ☐ Madrasta
 8. ☐ Padrasto
 9. ☐ Tia
 10. ☐ Tio
 11. ☐ Irmã
 12. ☐ Irmão
 13. ☐ Outro _____
 99. ☐ Não sabe

CHEFE

☐ ☐

28C. Qual o sexo da pessoa da família com maior renda?

1. ☐ Masculino
 2. ☐ Feminino
 8. ☐ Não se aplica
 9. ☐ Não sabe

SEXOCHEFE

☐

29C. Qual a idade da pessoa da família com maior renda (anos completos)? __

88. ☐ Não se aplica
 99. ☐ Não sabe

IDCHEFE

☐☐

30C. Essa pessoa sabe ler e escrever?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

LERCHEFE

☐

31C. Essa pessoa frequenta ou frequentou escola?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 34C**
 8. ☐ Não se aplica
 9. ☐ Não sabe

ESCCHEFE

☐

32C. Qual foi o último curso que essa pessoa frequentou ou frequenta?

1. ☐ Alfabetização de jovens e adultos
 2. ☐ Ensino fundamental ou 1o grau
 3. ☐ Ensino médio ou 2o grau
 4. ☐ Superior graduação incompleto **Passe para a questão 34C**
 5. ☐ Superior graduação completo **Passe para a questão 34C**
 8. ☐ Não se aplica
 9. ☐ Não sabe

CURSOCHefe

☐

33C. Qual a série que essa pessoa frequenta ou até que série frequentou?

1. ☐ Primeira
 2. ☐ Segunda
 3. ☐ Terceira
 4. ☐ Quarta
 5. ☐ Quinta
 6. ☐ Sexta
 7. ☐ Sétima
 8. ☐ Oitava
 88. ☐ Não se aplica
 99. ☐ Não sabe

SERIECHefe

☐☐

34C. Qual a ocupação atual (ou no que trabalha) a pessoa com a maior renda da família? (Descreva a ocupação. Caso seja aposentado, colocar a última atividade que exerceu).

88. ☐ Não se aplica

99. ☐ Não sabe

OCUPCHEFE

☐ ☐

35C. Qual a relação de trabalho do chefe da família?

1. ☐ Trabalha por conta própria

2. ☐ Assalariado ou empregado

3. ☐ Dono de empresa-empregador

4. ☐ Faz bico

8. ☐ Não se aplica

9. ☐ Não sabe

RELCHEFE

☐

36C. No mês passado quanto ganharam as pessoas da família que trabalham?

1ª pessoa R\$ __. __. __.

2ª pessoa R\$ __. __. __.

3ª pessoa R\$ __. __. __.

4ª pessoa R\$ __. __. __.

5ª pessoa R\$ __. __. __.

A família tem outra renda? __. __. __.

Renda total R\$ __. __. __.

99999. ☐ Não sabe

REDAF

☐ ☐ ☐ ☐ ☐

Quanto itens abaixo a família possui? (circule a resposta)

	Quantidade de itens				
	0	1	2	3	4 ou mais

37C. Televisão em cores	0	1	2	3	4
-------------------------	---	---	---	---	---

TELEVISAO

☐

38C. Rádio	0	1	2	3	4
------------	---	---	---	---	---

RADIO

☐

39C. Banheiro	0	4	5	6	7
---------------	---	---	---	---	---

BANHEIRO

☐

40C. Automóvel	0	4	7	9	9
----------------	---	---	---	---	---

AUTOMOVEL

☐

41C. Empregada mensalista	0	3	4	4	4
---------------------------	---	---	---	---	---

EMPREGADA

☐

42C. Máquina de lavar	0	2	2	2	2
-----------------------	---	---	---	---	---

MAQLAVAR

☐

43C. Videocassete ou DVD	0	2	2	2	2
--------------------------	---	---	---	---	---

DVD

☐

44C. Geladeira	0	4	4	4	4
----------------	---	---	---	---	---

GELADEIRA

☐

45C. Freezer (aparelho independente ou parte da geladeira duplex)	0	2	2	2	2
---	---	---	---	---	---

FREEZER

☐

46C. Grau de Instrução da pessoa com maior renda

Analfabeto/Primário incompleto/ Até 3ª Série Fundamental	0
Primário completo/ Até 4ª Série Fundamental/Ginasial incompleto	1
Ginasial completo/ Fundamental completo/Colegial incompleto	2
Colegial completo/ Médio completo/Superior incompleto	4
Superior completo	8

INSTRUCAO

☐**BLOCO D – HÁBITOS DE VIDA****Agora vamos conversar um pouco sobre o consumo de bebida alcoólica.**

1D. Durante a gravidez, a sra. tomou cerveja?

1. ☐ Sim2. ☐ Não **Passe para a questão 14D**9. ☐ Não sabe

CERVEJA

☐

2D. Durante a gravidez, a sra. tomou cerveja nos três primeiros meses de gravidez?

1. ☐ Sim2. ☐ Não **Passe para a questão 6D**8. ☐ Não se aplica9. ☐ Não sabe

CERVEJA1T

☐

3D. Quantos dias por semana? _

8. ☐ Não se aplica9. ☐ Não sabe

DIACERV1T

☐

4D. Quanto tomava por dia (número de vasilhas) __

88. ☐ Não se aplica99. ☐ Não sabe

QTCERV1T

☐☐

5D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)2. ☐ Lata (350ml)3. ☐ Garrafa pequena (300ml) – long neck4. ☐ Garrafa (600-720ml)5. ☐ Outro8. ☐ Não se aplica9. ☐ Não sabe

TIPOCERV1T

☐

6D. Durante a gravidez a sra tomou cerveja dos 4 aos 6 meses de gravidez?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 10D**

8. ☐ Não se aplica

9. ☐ Não sabe

CERVEJA2T

☐

7D. Quantos dias por semana? _

8. ☐ Não se aplica

9. ☐ Não sabe

DIACERV2T

☐

8D. Quanto tomava por dia (número de vasilhas) _ _

88. ☐ Não se aplica

99. ☐ Não sabe

QTCERV2T

☐ ☐

9D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)
 2. ☐ Lata (350ml)
 3. ☐ Garrafa pequena (300ml) – long neck
 4. ☐ Garrafa (600-720ml)
 5. ☐ Outro
 8. ☐ Não se aplica
 9. ☐ Não sabe

TIPOCERV2T

☐

10D. Durante a gravidez a sra. tomou cerveja dos 7 meses ao final da gravidez?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 14D**

8. ☐ Não se aplica

9. ☐ Não sabe

CERVEJA3T

☐

11D. Quantos dias por semana? _

8. ☐ Não se aplica

9. ☐ Não sabe

DIACERV3T

☐

12D. Quanto tomava por dia (número de vasilhas) _ _

88. ☐ Não se aplica

99. ☐ Não sabe

QTCERV3T

☐ ☐

13D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)
 2. ☐ Lata (350ml)
 3. ☐ Garrafa pequena (300ml) – long neck
 4. ☐ Garrafa (600-720ml)
 5. ☐ Outro
 8. ☐ Não se aplica
 9. ☐ Não sabe

TIPOCERV3T

☐

14D. Durante a gravidez a sra tomou vinho?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 27D**
 9. ☐ Não sabe

VINHO

☐

15D. Durante a gravidez a sra tomou vinho nos três primeiros meses de gravidez?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 19D**
 8. ☐ Não se aplica
 9. ☐ Não sabe

VINHO1T

☐

16D. Quantos dias por semana? __

8. ☐ Não se aplica
 9. ☐ Não sabe

DIAVINHO1T

☐

17D. Quanto tomava por dia (número de vasilhas) __ __

88. ☐ Não se aplica
 99. ☐ Não sabe

QTVINHO1T

☐ ☐

18D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)
 2. ☐ Cálice ou taça (400 ml)
 3. ☐ Garrafa pequena (300ml)
 4. ☐ Garrafa (600-720ml)
 5. ☐ Outro
 8. ☐ Não se aplica
 9. ☐ Não sabe

TIPOVINHO1T

☐

19D. Durante a gravidez a sra. tomou vinho dos 4 aos 6 meses de gravidez?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 23D**
 8. ☐ Não se aplica
 9. ☐ Não sabe

VINHO2T

☐

20D. Quantos dias por semana? __

8. ☐ Não se aplica
 9. ☐ Não sabe

DIAVINHO2T

☐

21D. Quanto tomava por dia (número de vasilhas) __ __

88. ☐ Não se aplica
 99. ☐ Não sabe

QTVINHO2T

☐ ☐

22D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)
2. ☐ Cálice ou taça (400 ml)
3. ☐ Garrafa pequena (300ml)
4. ☐ Garrafa (600-720ml)
5. ☐ Outro
8. ☐ Não se aplica
9. ☐ Não sabe

TIPOVINHO2T

☐

23D. Durante a gravidez a sra. tomou vinho dos 7 meses ao final da gravidez?

1. ☐ Sim
2. ☐ Não **Passe para a questão 27D**
8. ☐ Não se aplica
9. ☐ Não sabe

VINHO3T

☐

24D. Quantos dias por semana? _

8. ☐ Não se aplica
9. ☐ Não sabe

DIAVINHO3T

☐

25D. Quanto tomava por dia (número de vasilhas) __

88. ☐ Não se aplica
99. ☐ Não sabe

QTVINHO3T

☐ ☐

26D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)
2. ☐ Cálice ou taça (400 ml)
3. ☐ Garrafa pequena (300ml)
4. ☐ Garrafa (600-720ml)
5. ☐ Outro
8. ☐ Não se aplica
9. ☐ Não sabe

TIPOVINHO3T

☐

27D. Durante a gravidez a sra. tomou algum outro tipo de bebida como uísque, vodka, gim, rum?

1. ☐ Sim
2. ☐ Não **Passe para a questão 40D**
9. ☐ Não sabe

DEST

☐

28D. Durante a gravidez a sra. Tomou algum outro tipo de bebida como uísque, vodka, gim, rum nos primeiros meses de gravidez?

1. ☐ Sim
2. ☐ Não **Passe para a questão 32D**
8. ☐ Não se aplica
9. ☐ Não sabe

DEST1T

☐

29D. Quantos dias por semana? _

8. ☐ Não se aplica
9. ☐ Não sabe

DIADEST1T

☐

30D. Quanto tomava por dia (número de vasilhas) __

88. ☐ Não se aplica

99. ☐ Não sabe

QTDEST1T

☐ ☐

31D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)

2. ☐ Cálice, taça (400 ml)

3. ☐ Martelo, copo de pinga (100ml)

4. ☐ Lata (350ml) retirar

5. ☐ Garrafa pequena (300ml)

6. ☐ Garrafa (600-720ml)

7. ☐ Outro

8. ☐ Não se aplica

9. ☐ Não sabe

TIPODEST1T

☐

32D. Durante a gravidez a sra tomou algum outro tipo de bebida como uísque, vodka, gim, rum dos 4 aos 6 meses de gravidez?

1. ☐ Sim

2. ☐ Não **Passe para a questão 36D**

8. ☐ Não se aplica

9. ☐ Não sabe

DEST2T

☐

33D. Quantos dias por semana? __

8. ☐ Não se aplica

9. ☐ Não sabe

DIATEST2T

☐

34D. Quanto tomava por dia (número de vasilhas) __

88. ☐ Não se aplica

99. ☐ Não sabe

QTDEST2T

☐ ☐

35D. Qual o tipo de vasilha?

1. ☐ Copo comum (200ml)

2. ☐ Cálice, taça (400 ml)

3. ☐ Martelo, copo de pinga (100ml)

4. ☐ Lata (350ml)

5. ☐ Garrafa pequena (300ml)

6. ☐ Garrafa (600-720ml)

7. ☐ Outro

8. ☐ Não se aplica

9. ☐ Não sabe

TIPODEST2T

☐

36D. Durante a gravidez a sra tomou algum outro tipo de bebida como uísque, vodka, gim, rum dos 7 meses ao final da gravidez?

1. ☐ Sim

2. ☐ Não **Passe para a questão 40D**

8. ☐ Não se aplica

9. ☐ Não sabe

DEST3T

☐

37D. Quantos dias por semana? __		
8. ☐ Não se aplica		
9. ☐ Não sabe	DIATEST3T	☐
38D. Quanto tomava por dia (número de vasilhas) __		
88. ☐ Não se aplica		
99. ☐ Não sabe	QTDEST3T	☐ ☐
39D. Qual o tipo de vasilha?		
1. ☐ Copo comum (200ml)		
2. ☐ Cálice, taça (400 ml)		
3. ☐ Martelo, copo de pinga (100ml)		
4. ☐ Lata (350ml)		
5. ☐ Garrafa pequena (300ml)		
6. ☐ Garrafa (600-720ml)		
7. ☐ Outro		
8. ☐ Não se aplica		
9. ☐ Não sabe	TIPODEST3T	☐

Agora vamos conversar um pouco sobre o hábito de fumar.

40D. A sra. tem ou teve o hábito de fumar cigarros?		
1. ☐ Sim		
2. ☐ Não Passe para a questão 51D		
9. ☐ Não sabe	HABITOFUMO	☐
41D. Com que idade a sra. começou a fumar cigarros? __		
88. ☐ Não se aplica		
99. ☐ Não sabe	IDAEFUMO	☐ ☐
42D. Se a sra. parou de fumar, com que idade parou? __		
88. ☐ Não se aplica ou ainda fuma		
99. ☐ Não sabe	IDAEPAROU	☐ ☐
43D. Se a sra. parou, quantos cigarros por dia em média a sra. costumava fumar? __		
88. ☐ Não se aplica ou ainda fuma		
99. ☐ Não sabe	NCIGPAROU	☐ ☐
44D. No período de 6 meses antes desta gravidez a sra. fumava?		
1. ☐ Sim		
2. ☐ Não		
8. ☐ Não se aplica		
9. ☐ Não sabe	FUMOANTES	☐
45D. A sra. fumou durante esta gravidez?		
1. ☐ Sim		
2. ☐ Não Passe para questão 51D		
8. ☐ Não se aplica		
9. ☐ Não sabe	FUMOGRAV	☐

46D. A sra. fumou do 1º ao 3º mês de gestação?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

FUMO1T

☐

47D. A sra. fumou do 4º ao 6º mês de gestação?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

FUMO2T

☐

48D. A sra. fumou do 7º mês de gestação até o final?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

FUMO3T

☐

49D. Durante a gravidez a sra. fumava todos os dias?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

FUMODIA

☐

50D. Quantos cigarros a sra. fumava por dia? ____

88. ☐ Não se aplica
 99. ☐ Não sabe

QTFUMO

☐ ☐

51D. A sra. convive em casa com outras pessoas que fumam?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 54D**
 9. ☐ Não sabe

FUMOCASA

☐

52D. Quantas pessoas que residem com a sra. fumam? _

8. ☐ Não se aplica
 9. ☐ Não sabe

QTFUMCASA

☐

53D. Estas pessoas fumam perto da sra. em sua casa?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

FUMOPERTO

☐

54D. E no trabalho, as pessoas fumam perto da sra.?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

FUMOTRAB

☐

55D. Caso a sra. tenha ficado próxima a pessoas que fumam, quantas horas por dia fica perto de fumantes?

1. ☐ Menos de 1 hora por dia

HORASFUMO

☐

2. ☐ Mais de 1 hora por dia
 8. ☐ Não se aplica
 9. ☐ Não sabe

Agora vamos conversar um pouco sobre o consumo de café.

56D. A sra. costumava tomar café uma vez ou mais por semana durante a gravidez?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 1E**
 9. ☐ Não sabe

CAFE1T

☐

57D. A sra. costumava tomar café uma vez ou mais por semana nos 3 primeiros meses da gravidez?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 61D**

8. ☐ Não se aplica
 9. ☐ Não sabe

CAFE1T

☐

58D. Quantos dias por semana a sra. tomava café? __

8. ☐ Não se aplica
 9. ☐ Não sabe

DIASCAFE1T

☐

59D. Quantas vezes por dia a sra. tomava café? __

88. ☐ Não se aplica

99. ☐ Não sabe

VEZESCAFE1T

☐ ☐

60D. Qual o tipo de vasilha em que a sra. costumava tomar café?

1. ☐ xícara de chá
 2. ☐ xícara de cafezinho
 3. ☐ meia taça
 4. ☐ copo comum – 200 ml
 5. ☐ outro _____

8. ☐ Não se aplica
 9. ☐ Não sabe

QTCAFE1T

☐

61D. A sra. costumava tomar café uma vez ou mais por semana dos 4 aos 6 meses da gravidez?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 65D**

8. ☐ Não se aplica
 9. ☐ Não sabe

CAFE2T

☐

62D. Quantos dias por semana a sra. tomava café? __

8. ☐ Não se aplica
 9. ☐ Não sabe

DIASCAFE2T

☐

63D. Quantas vezes por dia a sra. tomava café? __

88. ☐ Não se aplica

99. ☐ Não sabe

VEZESCAFE2T

☐ ☐

64D. Qual o tipo de vasilha em que a sra. costumava tomar café?

1. ☐ xícara de chá
2. ☐ xícara de cafezinho
3. ☐ meia taça
4. ☐ copo comum - 200 ml
5. ☐ outro _____
8. ☐ Não se aplica
9. ☐ Não sabe

QTCAFE2T

☐

65D. A sra. costumava tomar café uma vez ou mais por semana dos 7 meses ao final da gravidez?

1. ☐ Sim
2. ☐ Não **Passe para a questão 1E**
8. ☐ Não se aplica
9. ☐ Não sabe

CAFE3T

☐

66D. Quantos dias por semana a sra. tomava café? _

8. ☐ Não se aplica
9. ☐ Não sabe

DIASCAFE3T

☐

67D. Quantas vezes por dia a sra. tomava café? _ _

88. ☐ Não se aplica
99. ☐ Não sabe

VEZESCAFE3T

☐ ☐

68D. Qual o tipo de vasilha em que a sra. costumava tomar café?

1. ☐ xícara de chá
2. ☐ xícara de cafezinho
3. ☐ meia taça
4. ☐ copo comum - 200 ml
5. ☐ outro
8. ☐ Não se aplica
9. ☐ Não sabe

QTCAFE3T

☐

BLOCO E – DADOS DO COMPANHEIRO

Caso o companheiro seja a pessoa com maior renda ou não more junto com a mãe do RN, passe para a questão 1F.

1E. Qual a idade do companheiro atual? _ _

88. ☐ Não se aplica – não tem companheiro atual
99. ☐ Não sabe

IDCOMP

☐ ☐

2E. O seu companheiro sabe ler e escrever?

1. ☐ Sim
2. ☐ Não
8. ☐ Não se aplica
9. ☐ Não sabe

LERCOMP

☐

3E. O seu companheiro frequenta ou frequentou escola?

1. ☐ Sim

2. ☐ Não **Passe para a questão 6E**

8. ☐ Não se aplica

9. ☐ Não sabe

ESCCOMP

☐

4E. Qual o último curso que seu companheiro frequentou ou frequenta?

1. ☐ Alfabetização de jovens e adultos

2. ☐ Ensino fundamental ou 1o grau

3. ☐ Ensino médio ou 2o grau

4. ☐ Superior graduação incompleto **Passe para a questão 6E**

5. ☐ Superior graduação completo **Passe para a questão 6E**

8. ☐ Não se aplica

9. ☐ Não sabe

CURSOCOMP

☐

5E. Qual a série que seu companheiro frequenta ou até que série estudou?

1. ☐ Primeira

2. ☐ Segunda

3. ☐ Terceira

4. ☐ Quarta

5. ☐ Quinta

6. ☐ Sexta

7. ☐ Sétima

8. ☐ Oitava

88. ☐ Não se aplica

99. ☐ Não sabe

SERIECOMP

☐ ☐

6E. O seu companheiro está trabalhando no momento?

1. ☐ Sim

2. ☐ Não **Passe para a questão 1F**

8. ☐ Não se aplica

9. ☐ Não sabe

TRABCOMP

☐

7E. Qual a ocupação atual (ou no que trabalha) o seu companheiro atual? (Descreva a ocupação. Caso seja aposentado, colocar a última atividade que exerceu).

88. ☐ Não se aplica

99. ☐ Não sabe

OCUPCOMP

☐ ☐

8E. Qual a relação de trabalho do seu companheiro?

1. ☐ Trabalha por conta própria

2. ☐ Assalariado ou empregado

3. ☐ Dono de empresa-empregador

4. ☐ Faz bico

8. ☐ Não se aplica

9. ☐ Não sabe

RELCOMP

☐

BLOCO F – DADOS DA SAÚDE SEXUAL E REPRODUTIVA

1F. Em que mês de gestação a sra. soube que estava grávida?

—

9. ☐ Não sabe

MESGRAV

☐

2F. Que idade a sra. tinha quando menstruou pela primeira vez? __

99. ☐ Não sabe

IDMENARCA

☐

3F. Qual era sua idade quando a sra. teve sua primeira relação sexual? __

99. ☐ Não sabe

IDRELSEX

☐

4F. Que idade a sra. tinha quando engravidou pela primeira vez? __

99. ☐ Não sabe

IDGRAV

☐

5F. Quantas vezes a sra. engravidou? (incluindo a gravidez atual) Contar todas as gestações até as que não chegaram no final, inclusive abortos. __

99. ☐ Não sabe

GESTA

☐

6F. A gravidez atual foi planejada?

1. ☐ Sim2. ☐ Não9. ☐ Não sabe

PLANGEST

☐

7F. Quantos filhos nasceram vivos (incluindo o atual)? __

99. ☐ Não sabe

NASCVIVO

☐

8F. Qual a idade da sra. no início da gestação anterior? __

88. ☐ Não se aplica99. ☐ Não sabe

IDULTGEST

☐

9F. Quantos partos a sra. teve (incluindo o atual)? __

99. ☐ Não sabe

PARTOS

☐

10F. Quantos partos foram cesáreas (incluindo o atual, caso tenha sido cesárea)? _

9. ☐ Não sabe

PARTOCESA

☐

11F. O parto anterior da sra. foi:

1. ☐ Vaginal/Vaginal com fórceps2. ☐ Cesárea8. ☐ Não se aplica9. ☐ Não sabe

TIPARTOANT

☐

12F. Quantos partos anteriores da sra. foram antes do tempo (premáturos) incluindo o atual? _

9. ☐ não sabe

PARTOANTPT

☐

13F. A sra. teve algum aborto?

1. ☐ Sim2. ☐ Não **Passe para questão 15F**9. ☐ Não sabe

ABORTO

☐

14F. Caso sim, quantos abortos foram? _

8. ☐ Não se aplica9. ☐ Não sabe

QTABORTO

☐

15F. A sra teve algum filho que nasceu morto (incluir atual)?

1. ☐ Sim

2. ☐ Não **Passe para a questão 17F**

9. ☐ Não sabe

NASCORTO

☐

16F. Caso sim, quantos foram (incluir atual)? _

8. ☐ Não se aplica

9. ☐ Não sabe

QTMORTO

☐

17F. Quantos filhos vivos a sra. tem atualmente? _ _

99. ☐ Não sabe

NFILHOS

☐ ☐

BLOCO G - MORBIDADES

Agora vamos conversar com a sra. sobre os problemas que teve durante a gravidez.

1G. A sra. teve hipertensão (pressão alta) fora da gestação diagnosticada por médico ou enfermeiro?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HIPERT

☐

2G. A sra. teve hipertensão (pressão alta) na gestação atual?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HIPERTGEST

☐

3G. A sra. teve antes da gestação nível elevado de açúcar no sangue (diabetes) diagnosticado por médico ou enfermeiro?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

DIABETES

☐

4G. A sra. teve nível elevado de açúcar no sangue (diabetes) diagnosticado por médico ou enfermeiro durante a gestação?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

DIABGEST

☐

5G. A sra. teve herpes durante a gestação diagnosticado por médico, enfermeiro ou dentista?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

HERPESGEST

☐

6G. A sra. teve sarampo durante a gestação diagnosticado por médico, enfermeiro ou dentista?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

SARAMPGEST

☐

7G. A sra. teve catapora durante a gestação diagnosticado por médico, enfermeiro ou dentista?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

CATAPGEST

☐

8G. A sra. teve rubéola durante a gestação diagnosticado por médico, enfermeiro ou dentista?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

RUBGEST

☐

9G. A sra. teve algum episódio de febre alta (temperatura acima de 38°) que durou mais de 24 horas durante esta gestação diagnosticada por médico, enfermeiro ou dentista?

1. ☐ Sim

2. ☐ Não **Passe para a questão 11G**

9. ☐ Não sabe

FEBREGEST

☐

10G. Caso sim, quantas vezes? __

88. ☐ Não se aplica

99. ☐ Não sabe

QTFEBRE

☐ ☐

11G. A sra. teve anemia antes da gestação diagnosticado por médico ou enfermeiro?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

ANEMIA

☐

12G. A sra. teve anemia diagnosticado por médico ou enfermeiro durante a gestação?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

ANEMIAGEST

☐

13G. A sra. teve toxoplasmose antes da gestação diagnosticado por médico ou enfermeiro?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

TOXO

☐

14G. A sra. teve toxoplasmose diagnosticado por médico ou enfermeiro durante a gestação?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

TOXOGEST

☐

15G. A sra. teve sífilis antes da gestação diagnosticado por médico ou enfermeiro?

1. ☐ Sim

2. ☐ Não

9. ☐ Não sabe

SIFILIS

☐

16G. A sra. teve sífilis durante a gestação diagnosticado por médico ou enfermeiro?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

SIFILISGEST

☐

17G. A sra. teve alguma infecção urinária/ durante a gestação atual diagnosticada por médico ou enfermeiro?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

INFECURI

☐

18G. A sra. teve algum corrimento vaginal durante a gestação atual?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

CORRIMENTO

☐

19G. A sra. sofreu alguma queda ou acidente durante a gestação?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

QUEDA

☐

20G. A sra. teve algum sangramento vaginal nos últimos 3 meses da gestação atual?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

SANGVAGINA

☐

21G. A sra. foi internada alguma vez durante a gestação atual seja por qualquer motivo?

1. ☐ Sim
2. ☐ Não **Passe para a questão 23G**
9. ☐ Não sabe

HOSP

☐

22G. Qual foi o motivo da internação?

88. ☐ Não se aplica
98. ☐ Não sabe

CAUSAHOSP

☐ ☐

23G. A sra. teve ameaça de aborto na gestação atual?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

AMABORTO

☐

24G. A sra. teve ameaça de parto prematuro (antes do tempo) na gestação atual?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

AMPT

☐

25G. A sra. teve outra doença durante a gestação atual?

1. ☐ Sim

2. ☐ Não **Passe para a questão 27G**

9. ☐ Não sabe

OUTRAD

☐

26G. Qual doença? _____

88. ☐ Não se aplica

99. ☐ Não sabe

DOENCAGEST

☐ ☐

27G. A sra. usou algum remédio durante a gravidez?

1. ☐ Sim

2. ☐ Não **Passe para o bloco H**

9. ☐ Não sabe

REMGEST

☐

Caso tenha utilizando, qual o nome do remédio e qual o mês da gestação a sra. Estava quando começou e quando parou (mês da gestação)? Se continua usando, anotar que está em uso?

28G. Remédio

88. ☐ Não se aplica

99. ☐ Não sabe

NOMEREM1

☐ ☐

29G. Mês de início _

0. ☐ Usa desde antes da gravidez

88. ☐ Não se aplica

99. ☐ Não sabe

MESIREM1

☐ ☐

30G. Mês de término _

10. ☐ Ainda usa

88. ☐ Não se aplica

99. ☐ Não sabe

MESTREM1

☐ ☐

31G. Remédio

88. ☐ Não se aplica

99. ☐ Não sabe

NOMEREM2

☐ ☐

32G. Mês de Início _

0. ☐ Usa desde antes da gravidez

88. ☐ Não se aplica

99. ☐ Não sabe

MESIREM2

☐ ☐

33G. Mês de término _

10. ☐ Ainda usa

88. ☐ Não se aplica

9.9 ☐ Não sabe

MESTREM2

☐ ☐

34G. Remédio

88. ☐ Não se aplica99. ☐ Não sabe

NOMEREM3

☐☐

35G. Mês de início _

0. ☐ Usa desde antes da gravidez88. ☐ Não se aplica99. ☐ Não sabe

MESIREM3

☐☐

36G. Mês de término _

10. ☐ Ainda usa88. ☐ Não se aplica99. ☐ Não sabe

MESTREM3

☐☐

37G. Remédio

88. ☐ Não se aplica99. ☐ Não sabe

NOMEREM4

☐☐

38G. Mês de início _

0. ☐ Usa desde antes da gravidez88. ☐ Não se aplica99. ☐ Não sabe

MESIREM4

☐☐

39G. Mês de término _

10. ☐ Ainda usa88. ☐ Não se aplica99. ☐ Não sabe

MESTREM4

☐☐

40G. Remédio

88. ☐ Não se aplica99. ☐ Não sabe

NOMEREM5

☐☐

41G. Mês de início _

0. ☐ Usa desde antes da gravidez88. ☐ Não se aplica99. ☐ Não sabe

MESIREM5

☐☐

42G. Mês de término _

10. ☐ Ainda usa88. ☐ Não se aplica99. ☐ Não sabe

MESTREM5

☐☐

BLOCO H – CARACTERÍSTICAS DA GESTAÇÃO ATUAL E DO PRÉ-NATAL

Pergunte se a puérpera dispõe do cartão da gestante e se está de posse do mesmo. Confirme as respostas no cartão

1H. A sra tem cartão da gestante?

1. ☐ Sim

2. ☐ Não

8. ☐ Não se aplica – não fez pré-natal

9. ☐ Não sabe

CARTAO

☐

2H. Qual a data da sua última menstruação (DD/MM/AA)?

___/___/___

99999999. ☐ Não sabe

DUM

☐☐☐☐☐☐☐☐

3H. Caso não saiba a data, informar o mês/ano: ___/___

888888. ☐ Não se aplica

999999. ☐ Não sabe

MESANOUM

☐☐☐☐☐☐

4H. Qual o seu peso antes de engravidar? ____, __ kg

9999. ☐ Não sabe

PESOANTES

☐☐☐☐

5H. Qual a sua altura antes de engravidar? ____, __ cm

9999. ☐ Não sabe

ALTURAANT

☐☐☐☐

6H. A sra fez pré-natal?

1. ☐ Sim

2. ☐ Não **Passe para questão 39H**

9. ☐ Não sabe

PN

☐

7H. Qual a data da primeira consulta pré-natal (DD/MM/AA)?

___/___/___

88888888. ☐ Não se aplica

99999999. ☐ Não sabe

DT1CPN

☐☐☐☐☐☐☐☐

8H. Em que mês de gravidez a sra. iniciou as consultas de pré-natal? __

88. ☐ Não se aplica

99. ☐ Não sabe

MES1CPN

☐☐

9H. Quantas consultas de pré-natal a sra. fez no 1º trimestre de a gestação? __

88. ☐ Não se aplica

99. ☐ Não sabe

QTCPN1T

☐☐

10H. Quantas consultas de pré-natal a sra. fez no 2º trimestre de a gestação? __

88. ☐ Não se aplica

99. ☐ Não sabe

QTCPN2T

☐☐

11H. Quantas consultas de pré-natal a sra. fez no 3º trimestre de a gestação? __

88. ☐ Não se aplica

99. ☐ Não sabe

QTCPN3T

☐☐

12H. O médico ou enfermeiro encaminhou a senhora para fazer tratamento com o dentista durante esta gravidez?

1. ☐ Sim

2. ☐ Não

8. ☐ Não se aplica

9. ☐ Não sabe

TRATDEN

☐

13H. Quantas consultas médicas a sra. fez no pré-natal? __

88. ☐ Não se aplica

99. ☐ Não sabe

QTCPNMED

☐☐

14H. Quantas consultas com enfermeira(o) a sra. fez no pré-natal? __

88. ☐ Não se aplica

99. ☐ Não sabe

QTCPNENF

☐☐

15H. Quantos exames de radiografia (incluindo radiografias dos dentes) foram feitos durante a gestação atual? __

88. ☐ Não se aplica

99. ☐ Não sabe

QTEXRADI

☐☐

16H. Em que local a sra. fez o pré-natal?

1. ☐ SUS

3. ☐ Plano de saúde/ seguro saúde

4. ☐ Particular

8. ☐ Não se aplica

9. ☐ Não sabe

LOCALNP

☐

17H. Qual a data da última consulta pré-natal (DD/MM/AA)?

__/__/__

88888888. ☐ Não se aplica

99999999. ☐ Não sabe

DTUCPN

☐☐☐☐☐☐☐☐☐☐

18H. Caso não saiba a data, informar o mês de gravidez aproximado: __

8. ☐ Não se aplica

9. ☐ Não sabe

MESUCPN

☐

19H. Quantos exames de ultrassonografia foram feitos durante a gestação atual? __

8. ☐ Não se aplica

9. ☐ Não sabe

NUSPN

☐

Durante as consultas de pré-natal o médico ou enfermeira alguma vez:

20H. Solicitou exame de sangue?

1. ☐ Sim

2. ☐ Não

8. ☐ Não se aplica

9. ☐ Não sabe

EXSANGUE

☐

21H. Solicitou exame de urina?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXURINA

☐

22H. Perguntou a data da última menstruação?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

PDUM

☐

23H. Verificou o seu peso?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXPESO

☐

24H. Mediu a sua barriga?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXBARRIGA

☐

25H. Receitou Cálcio?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

RECCALCIO

☐

26H. Mediu a sua pressão?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXPA

☐

27H. Fez exame ginecológico?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXGINEC

☐

28H. Receitou remédio para anemia?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

RECANEMIA

☐

29H. Receitou vitamina?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

RECVIT

☐

30H. Orientou sobre amamentação?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

OAMAMENT

☐

31H. O médico perguntou se a sra. estava usando algum medicamento?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

OMEDIC

☐

32H. Orientou sobre o risco do uso de remédios sem orientação médica durante a gravidez?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

OREM

☐

33H. Orientou sobre como evitar toxoplasmose durante a gravidez? (lavar muito bem frutas e verduras, não comer carne mal passada, evitar contato com gatos, não manipular terra, lavar muito bem as mãos antes das refeições).

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

OTOXO

☐

34H. Examinou o seu seio?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXSEIO

☐

35H. Fez exame de prevenção de câncer de colo de útero?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXCOLOUT

☐

36H. Fez exame de sangue para sífilis no pré-natal?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXSIFILIS

☐

37H. Fez exame de sangue para saber o tipo de sangue?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXTIPOSANG

☐

38H. Ofereceu exame de sangue para HIV no pré-natal?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EXHIV

☐

39H. Você tomou vacina contra o tétano?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 41H**
 3. ☐ Já estava vacinada antes da gravidez. **Passe para a questão 41H**

9. ☐ Não sabe

VACTET

☐

40H. Quantas doses de antitetânica a sra. recebeu? _

8. ☐ Não se aplica
 9. ☐ Não sabe

QTDOSSETET

☐

41H. Durante a gestação atual, a sra. recebeu atendimento de auxiliar de enfermagem?

1. ☐ Sim
 2. ☐ Não
 9. ☐ Não sabe

AUXENFPN

☐

42H. Durante a gestação atual, a sra. recebeu visita do agente de saúde?

1. ☐ Sim
 2. ☐ Não
 9. ☐ Não sabe

AGSAUDEPN

☐

43H. Durante a gestação atual, a sra. recebeu atendimento de parteira leiga?

1. ☐ Sim
 2. ☐ Não
 9. ☐ Não sabe

PARTPN

☐

44H. Durante a gestação atual, a sra. recebeu atendimento do programa de Saúde da Família (PSF)?

1. ☐ Sim
 2. ☐ Não
 9. ☐ Não sabe

PSFPN

☐

45H. Qual o seu peso ao final da gravidez? __ __ __, __ kg

9999. ☐ Não sabe

PESOFINAL

☐ ☐ ☐ ☐

46H. A sra. fez algum tratamento para engravidar?

1. ☐ Sim
 2. ☐ Não **Passe para a questão 1I**
 9. ☐ Não sabe

TRATGRAV

☐

47H. Tomou algum medicamento para induzir a ovulação?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

OVULA

☐

48H. Fez inseminação artificial?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

INSEMINA

☐

49H. Fez fertilização in vitro (bebê de proveta)?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

FERTILIZA

☐

50H. Fez injeção de espermatozoides?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

INJECAO

☐

BLOCO I - CARACTERÍSTICAS DO PARTO E DO NASCIMENTO

1I. Qual foi o tipo de parto?

1. ☐ Normal
 2. ☐ Cesárea **Passe para a questão 4I**
 3. ☐ Fórceps **Passe para a questão 4I**
 4. ☐ Vácuo extração **Passe para a questão 4I**

TIPARTO

☐

9. ☐ Não sabe

2I. Se normal, a sra. fez episiotomia (corte, pique)?

1. ☐ Sim
 2. ☐ Não
 8. ☐ Não se aplica
 9. ☐ Não sabe

EPISIO

☐

3I. Quantas horas decorreram entre a internação e o parto normal? __

88. ☐ Não se aplica
 99. ☐ Não sabe

HORASPARTO

☐ ☐

4I. Foi feita anestesia nas costas?

1. ☐ Sim
 2. ☐ Não
 9. ☐ Não sabe

ANESTESIA

☐

51. Caso cesárea, qual foi o motivo para fazer a cesárea?

1. ☐ sofrimento fetal (batidas do coração do bebê diminuiu / ou o bebê fez cocô dentro da barriga da mãe)
2. ☐ desproporção feto-pélvica (bacia pequena/bebê grande)
3. ☐ distócia de apresentação (bebê sentado/ posição errada)
4. ☐ hemorragia materna (teve sangramento)
5. ☐ parada de progressão (parou trabalho de parto/ pararam as dores)
6. ☐ eclâmpsia, pré-eclâmpsia (pressão alta)
7. ☐ pós-maturidade (passou do tempo)
8. ☐ morte fetal (o bebê morreu)
9. ☐ diabetes materna (açúcar no sangue)
10. ☐ cesáreas anteriores (já fez outra cesárea antes)
11. ☐ laqueadura (para ligar trompas)
12. ☐ mãe pediu (cesárea porque a mãe queria)
13. ☐ médico quis (na hora o médico resolveu fazer cesárea)
14. ☐ cesárea programada (médico marcou durante gravidez)
15. ☐ cirurgias ginecológicas anteriores (Miomectomia, plástica perineal)
16. ☐ outro _____

88. ☐ não se aplica

99. ☐ não sabe

MOTCESA

☐ ☐

61. Quantas horas decorreram entre a internação e a cesárea?

— —

88. ☐ Não se aplica

99. ☐ Não sabe

HORASCESA

☐ ☐

71. A sra. ligou as trompas?

1. ☐ Sim

2. ☐ Não **Passe para a questão 91**

9. ☐ Não sabe

LAQUEADURA

☐

81. Qual o motivo pelo qual a sra. ligou as trompas?

1. ☐ Já fez muita cesárea

2. ☐ Por problemas de saúde. Qual? _____

3. ☐ Questões financeiras

4. ☐ Já tinha o número de filhos que desejava

5. ☐ Outros _____

88. ☐ Não se aplica

99. ☐ Não sabe

MOTLAQ

☐ ☐

9l. O que a sra. sentiu que a fez vir para o hospital?

1. ☐ Sangramento vaginal
2. ☐ Perdeu líquido (água) vaginal
3. ☐ Sentiu contração ou dor ou cólica/ barriga endurecida
4. ☐ Febre/ infecção/infecção urinária
5. ☐ O bebê parou de mexer/ diminuíram movimentos
6. ☐ A vinda foi agendada para esta data
7. ☐ O médico encaminhou
8. ☐ Outro _____
99. ☐ Não sabe

MOTHOSP

☐ ☐

10l. Caso tenha feito cesárea, quando a sra. internou já sabia que iria fazer cesárea?

1. ☐ Sim
2. ☐ Não
8. ☐ Não se aplica
9. ☐ Não sabe

SABIACESA

☐

11l. A sra. passou por outros serviços (maternidades) antes de vir para esse hospital?

1. ☐ Sim
2. ☐ Não **Passe para a questão 13l**
9. ☐ Não sabe

PEREGRINOU

☐

12l. Caso sim, por quantos serviços passou? _

8. ☐ Não se aplica
9. ☐ Não sabe

QTPEREG

☐

13l. Quando a sra. foi hospitalizada estava sentindo as dores do trabalho de parto?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

DORPARTO

☐

14l. O médico precisou romper a bolsa?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

ROMPBOLSA

☐

15l. Foi preciso colocar soro ou outro remédio para começar o trabalho de parto ou para ajudar o bebê a nascer?

1. ☐ Sim
2. ☐ Não **Passe para a questão 18l**
9. ☐ Não sabe

SORONASC

☐

16l. Qual a medicação utilizada?

1. ☐ Vaginal
2. ☐ Soro (endovenosa)
8. ☐ Não se aplica
9. ☐ Não sabe

TIPOMED

☐

17l. Por que foi necessário ajudar o bebê a nascer?

1. ☐ Passou do tempo de nascer
2. ☐ Pressão alta
3. ☐ Rompeu a bolsa
4. ☐ Incompatibilidade sanguínea (sangue não combina)
5. ☐ O bebê estava morto
6. ☐ O médico indicou
7. ☐ O trabalho de parto parou
8. ☐ Outra razão. Qual? _____
88. ☐ Não se aplica
99. ☐ Não sabe

MOTAJUDA

☐ ☐

18l. Quem atendeu ao parto?

1. ☐ Médico
2. ☐ Enfermeira
3. ☐ Auxiliar de enfermagem
4. ☐ Parteira leiga
5. ☐ Outro
9. ☐ Não sabe

QUEMPARTO

☐

19l. O parto foi realizado pelo mesmo médico que fez o pré-natal?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

MEDICO

☐

20l. Qual a categoria de atendimento ao parto?

1. ☐ SUS
2. ☐ Plano de saúde/ seguro saúde
3. ☐ Particular
9. ☐ Não sabe

CATP

☐

21I. Na hora do nascimento, quem atendeu o RN na sala de parto?

1. ☐ Médico obstetra
2. ☐ Médico pediatra/neonatologista
3. ☐ Anestesista
2. ☐ Enfermeira
3. ☐ Auxiliar de enfermagem
4. ☐ Parteira leiga
5. ☐ Outro _____
99. ☐ Não sabe

ATENDRN

☐ ☐

22I. O pediatra falou com a sra na sala de parto antes ou depois que o bebê nasceu?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

PEDIATRA

☐

23I. Número de filhos nascidos no parto: _

9. ☐ Não sabe

FETOS

☐

BLOCO J – EXPOSIÇÃO AO CITOMEGALOVÍRUS (CMV)

1J. Na sua casa morou ou está morando alguma criança?

1. ☐ Sim
2. ☐ Não **Passe para a questão 8J**
9. ☐ Não sabe

MORACRI

☐

2J. Caso sim, ela tem até três anos de idade?

1. ☐ Sim
2. ☐ Não **Passe para a questão 8J**
8. ☐ Não se aplica
9. ☐ Não sabe

CRIATE3

☐

3J. Caso tenha até 3 anos de idade, ela morou com a sra. durante a gestação?

1. ☐ Sim
2. ☐ Não
8. ☐ Não se aplica
9. ☐ Não sabe

MOROUGEST

☐

4J. Caso a criança não tenha morado com a sra. durante sua gestação, a criança morou com a sra. nos 12 meses anteriores?

1. ☐ Sim
2. ☐ Não
8. ☐ Não se aplica
9. ☐ Não sabe

MOROU12M

☐

5J. Essa(s) criança(s) frequentam creches ou escolas?

1. ☐ Sim

2. ☐ Não **Passe para a questão 8J**

8. ☐ Não se aplica

9. ☐ Não sabe

CRECHE

☐

6J. Quantos dias da semana a(s) criança(s) frequenta(m) a creche ou escola? _

8. ☐ Não se aplica

9. ☐ Não sabe

DIASCRECHE

☐

7J. Quantas horas por dia da semana a(s) criança(s) frequenta(m) a creche ou escola? __

88. ☐ Não se aplica

99. ☐ Não sabe

HORACRECHE

☐ ☐

8J. A sra. realizou algum cuidado direto com crianças de 3 anos ou menores (como trocar as fraldas, dar banhos, alimentar o bebê, etc.)? Incluindo crianças de sua família ou para amigos de seus filhos, ou no seu trabalho durante o ano que antecedeu ou atualmente na sua gravidez?

1. ☐ Sim

2. ☐ Não

8. ☐ Não se aplica

9. ☐ Não sabe

CUIDACRI

☐

9J. Caso a sra. tenha realizado algum cuidado direto com crianças de 3 anos ou menores, qual a frequência de cuidados com essa criança?

1. ☐ Menos de 12 vezes no ano

2. ☐ Uma vez por mês

3. ☐ Duas ou mais vezes no mês

4. ☐ Uma vez por semana

5. ☐ Mais de uma vez por semana

8. ☐ Não se aplica

9. ☐ Não sabe

QTCUIDA

☐

BLOCO K – DADOS DO PRONTUÁRIO

1K. Há registro de administração de ocitocina durante o trabalho de parto?

1. ☐ Sim

2. ☐ Não

OCITOCINA

☐

2K. Horário de início do medicamento (indução): __: __

8888. ☐ Não se aplica 9999. ☐ Não informado

HORAIMED

☐ ☐ ☐ ☐

3K. Horário do término do medicamento (indução): __: __

8888. ☐ Não se aplica 9999. ☐ Não informado

HORATMED

☐ ☐ ☐ ☐

4K. Caso o parto tenha sido cesárea, anotar a indicação da cesárea do prontuário _____

88. ☐ Não se aplica 99. ☐ Não informado

INDICACESA

☐ ☐

5K. Registro da idade gestacional avaliada pela Ultrassonografia (anotar o primeiro ultrassom) __ semanas

99. ☐ Não informado

IDGESTUS

☐ ☐

6K. Data da Ultrassonografia __/__/____

99999999. ☐ Não informado

DATAUS

☐ ☐ ☐ ☐ ☐ ☐ ☐ ☐

ANEXO D – QUESTIONÁRIO NASCIMENTO-RN



QUESTIONÁRIO DO NASCIMENTO - RN

Generated by Foxit PDF Creator © Foxit Software
http://www.foxitsoftware.com For evaluation only.

BLOCO A – DADOS DE IDENTIFICAÇÃO

1A. Número de identificação: _ _ _ _ _

1ª casela: 1 Ribeirão Preto

2 São Luís

2ª casela: 1 Pré-natal

2 Nascimento

3 1º ano

3ª casela: M. Avaliação no pré-natal

A. Avaliação no nascimento RN 1

B. Avaliação no nascimento RN 2

C. Avaliação no nascimento RN 3

D. Avaliação no nascimento RN 4

4ª e 5ª caselas: QM. Questionário da mãe

QC. Questionário do RN

SC. Saliva da criança

CO. Cordão umbilical

6ª à 9ª. caselas: número seqüencial para cada cidade

NUMERO

2A. Cidade:

1. ☐ Ribeirão Preto

2. ☐ São Luís

CIDADE

3A. Data da Entrevista (DD/MM/AAAA):

_ _ / _ _ / _ _ _ _

DATAENT

4A. Entrevistador (a) :

5A. Data do nascimento do RN (DD/MM/AAAA):

_ _ / _ _ / _ _ _ _

DATANASC

6A. Caso tenha nascido mais de um filho, indique qual a sequência de nascimento desse RN?

1. ☐ 1º

2. ☐ 2º

3. ☐ 3º

4. ☐ 4º

5. ☐ 5º

8. ☐ Não se aplica

9. ☐ Não sabe

ORDEMNASC

7A. Qual o sexo do recém-nascido?

1. ☐ Masculino

2. ☐ Feminino

9. ☐ Não sabe

SEXO

8A. Como nasceu o filho da sra.?

1. ☐ Vivo

2. ☐ Morto

9. ☐ Não sabe

NASC

9A. Qual foi a apresentação do bebê na hora do parto?

1. ☐ Cefálica
 2. ☐ Pélvica
 3. ☐ Transversa
 9. ☐ Não sabe

APRESENT

☐

Caso o RN tenha nascido morto, passe para o Bloco B

10A. Qual nome a sra. pretender dar ao seu bebê?

NOME

11A. O RN apresentou algum problema de saúde?
 (perguntar para a mãe e anotar do prontuário)

1. ☐ Sim
 2. ☐ Não **Passe para a questão 17A**
 8. ☐ Não se aplica
 9. ☐ Não sabe

PROBLEMA

☐

Caso tenha apresentado problema, nos diga qual foi.

12A.Problema

-
88. ☐ Não se aplica
 99. ☐ Não sabe

PROB1

☐ ☐

13A.Problema

-
88. ☐ Não se aplica
 99. ☐ Não sabe

PROB2

☐ ☐

14A.Problema

-
88. ☐ Não se aplica
 99. ☐ Não sabe

PROB3

☐ ☐

15A.Problema

-
88. ☐ Não se aplica
 99. ☐ Não sabe

PROB4

☐ ☐

16A.Problema

-
88. ☐ Não se aplica
 99. ☐ Não sabe

PROB5

☐ ☐

17A. Após o nascimento, quando saiu da sala de parto, o RN foi para:

1. ☐ Berço ao lado da mãe
2. ☐ Berçário
3. ☐ Cama da mãe
4. ☐ UTI Neonatal
5. ☐ Outro _____
8. ☐ Não se aplica
9. ☐ Não sabe

LOCALRN

☐

BLOCO B – DADOS DO PRONTUÁRIO

1B. Horário de nascimento do RN: __:__:__

HORANASC

☐ ☐ ☐ ☐

2B. Apgar 1º minuto: __

99. ☐ Não avaliado

APGAR1

☐ ☐

3B. Apgar 5º minuto: __

99. ☐ Não avaliado

APGAR5

☐ ☐

4B. Em caso de nascido morto:

1. ☐ Antes do parto
2. ☐ No momento do parto
8. ☐ Não se aplica
9. ☐ Não sabe

TIPONASC

☐

5B. O RN apresentou algum defeito congênito (malformação)?

1. ☐ Sim
2. ☐ Não Passe para a questão 7B
9. ☐ Não sabe

DEFCONG

☐

6B. Qual foi o tipo de defeito apresentado pelo RN?

1. ☐ Gastrosquise
2. ☐ Onfalocele
3. ☐ Defeitos do tubo neural - meningomieloceles
4. ☐ Defeitos do tubo neural - encefalocele
5. ☐ Trato genito-urinário - unilateral
5. ☐ Trato genito-urinário - bilateral
6. ☐ Cardíaco. Qual? _____
7. ☐ Outro _____
88. ☐ Não se aplica
99. ☐ Não sabe

TIPODEFCONG

☐ ☐

ANEXO F – DOCUMENTOS DE APROVAÇÃO DO COMITÊ DE ÉTICA E PESQUISA DE RIBEIRÃO PRETO E SÃO LUÍS



HOSPITAL DAS CLÍNICAS DA FACULDADE DE MEDICINA
DE RIBEIRÃO PRETO DA UNIVERSIDADE DE SÃO PAULO

www.hcrp.fmrp.usp.br



Ribeirão Preto, 13 de novembro de 2008

Ofício nº 4116/2008
CEP/RCC

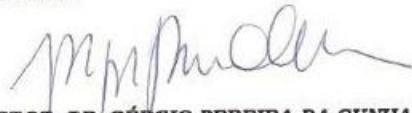
Prezados Professores,

O trabalho intitulado **“FATORES ETIOLÓGICOS DO NASCIMENTO PRÉ-TERMO E CONSEQUÊNCIAS DOS FATORES PERINATAIS NA SAÚDE DA CRIANÇA: COORTES DE NASCIMENTOS EM DUAS CIDADES BRASILEIRAS”**, foi analisado pelo Comitê de Ética em Pesquisa, em sua 276ª Reunião Ordinária realizada em 10/11/2008, e enquadrado na categoria: **APROVADO, bem como o Termo de Consentimento Livre e Esclarecido.**

Este Comitê segue integralmente a Conferência Internacional de Harmonização de Boas Práticas Clínicas (IGH-GCP), bem como a Resolução nº 196/96 CNS/MS.

Lembramos que devem ser apresentados a este CEP, o Relatório Parcial e o Relatório Final da pesquisa.

Atenciosamente.


PROF. DR. SÉRGIO PEREIRA DA CUNHA
Coordenador do Comitê de Ética em Pesquisa
do HCRP e da FMRP-USP

Ilustríssimos Senhores
PROF. DR. MARCO ANTONIO BARBIERI
PROFª. DRª. HELOISA BETTIOL
Depto. de Puericultura e Pediatria

Comitê de Ética em Pesquisa HCRP e FMRP-USP- Campus Universitário
FWA-00002733; IRB-00002186 e Registro SISNEP/CONEP nº 4
Fone (16) 3602-2228 • E-mail : cep@hcrp.fmrp.usp.br
Monte Alegre 14048-900 Ribeirão Preto SP

**UNIVERSIDADE FEDERAL DO MARANHÃO**

Fundação Instituída nos termos da Lei 5.152 de 21/10/1966

São Luís – Maranhão

RESOLUÇÃO Nº 721-CONSEPE, de 16 de dezembro de 2009.**Homologa Projetos de Pesquisa**

O Vice-Reitor da Universidade Federal do Maranhão, na qualidade de **PRESIDENTE EM EXERCÍCIO DO CONSELHO DE ENSINO, PESQUISA E EXTENSÃO**, no uso de suas atribuições estatutárias e regimentais;

Consideração o que decidiu referido Conselho em sessão desta data;

RESOLVE:**Art. 1º**

HOMOLOGAR os Projetos de Pesquisa, adiante relacionados:

- I. **Processo 9792/2007-63.** Projeto Comportamento de misturas e determinação de parâmetro de solubilidade em petróleos brasileiros, do Departamento de Tecnologia Química;
- II. **Processo 5264/2008-17.** Projeto Equações diferenciais e aplicações, do Departamento de Matemática;
- III. **Processo 6390/2008-93.** Projeto Avaliação da atividade antifúngica de folhas da espécie vegetal *Piper aduncum* (*Piperaceae*) – pimenta de macaco, desenvolvido pelo Departamento de Química e Departamento de Patologia;
- IV. **Processo 6392/2008-80.** Projeto Simulações computacionais em nanotubos de carbono, do Departamento de Física;
- V. **Processo 6393/2008-24.** Projeto Atividades de extratos e de frações parcialmente purificadas de folhas de *Ricinus communis* e de *Calotropis procera* em larvas e fêmeas ingurgitadas de *Boophilus microplus*, do Departamento de Química;
- VI. **Processo 6396/2008-65.** Projeto Síntese, caracterização e reatividade de tetraaminas de rutênio com ligantes tiois de interesse biológico em posição trans a ligantes N-heterocíclicos, do Departamento de Química;
- VII. **Processo 7338/2008-57.** Projeto Efeitos do tanino condensado de acácia negra (*Acacia mearnsii*) e acácia magium (*acacia magium*) sobre helmintos de caprinos infectados naturalmente, do Curso de Zootecnia/CCAA;
- VIII. **Processo 3377/2009-68.** Projeto Modelagem do processamento mecânico a quente de aços inoxidáveis utilizados em implantes ortopédicos, do Departamento de Física;



UNIVERSIDADE FEDERAL DO MARANHÃO

Fundação Instituída nos termos da Lei 5.152 de 21/10/1966

São Luís – Maranhão

2

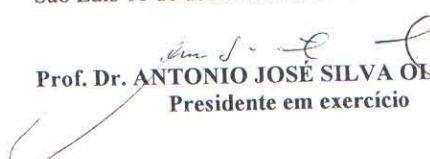
- IX. **Processo: 4051/2009-56.** Projeto Caracterização de haplótipos em pacientes com anemia falciforme no Estado do Maranhão, do Departamento de Farmácia;
- X. **Processo 4175/2009-38.** Projeto Recuperação das áreas degradadas de recarga e descarga do Aquífero Barreiras da Bacia do Rio Igarauá, do Departamento de Geociências;
- XI. **Processo 4548/2009-78.** Projeto Influência da amamentação natural no desenvolvimento de hábitos de sucção não-nutritivos, do Departamento de Odontologia I;
- XII. **Processo 4694/2009-09.** Projeto Fatores etiológicos do nascimento pré-termo e consequências dos fatores perinatais na saúde da criança: coortes de nascimento em duas cidades brasileiras, do Departamento de Saúde Pública;
- XIII. **Processo 4792/2009-33.** Projeto Projeção do amaciamento entre deformações no forjamento a quente e avaliação da resistência à corrosão do aço inoxidável ISSO 5832-9 utilizado em implantes ortopédicos, do Departamento de Física;
- XIV. **Processo 5167/2009-03.** A participação da mulher na agricultura familiar e na organização da comunidade negra rural quilombola do Só Assim/Agravila Só Assim, Município de Alcântara/MA, do Departamento de Geociências;
- XV. **Processo 5168/2009-57.** Educação ambiental no Campus do Bacanga, São Luís/MA, do Departamento de Geociências;
- XVI. **Processo 5189/2009-76.** Avaliação de cultivares de mandioca (*manihot esculenta crantz*) do Curso de Zootecnia/CCAA;
- XVII. **Processo 5193/2009-37.** Estudo da composição e distribuição sazonal das espécies vesicantes do gênero *paederus* (*Coleoptera, staphylinidae*) no município de Chapadinha, Maranhão, Brasil, do Curso de Ciências Biológicas/CCAA;
- XVIII. **Processo 5195/2009-24.** Potencial eólico de Chapadinha, do Curso de Agronomia/CCAA;
- XIX. **Processo 5200/2009-03.** Soluções para interconexão de fontes renováveis eólica e solar com redes CA de energia, do Departamento de Engenharia de Eletricidade;
- XX. **Processo 5204/2009-88.** Indicadores educacionais do ensino médio da região do Baixo Parnaíba, do Curso de Agronomia/CCAA;
- XXI. **Processo 5205/2009-21.** Projeto Estudo das características morfofisiológicas do capim-mombaça (*Panicum Maximum*) submetido a doses crescentes de nitrogênio, do Curso de Zootecnia/CCAA;
- XXII. **Processo 5209/2009-06.** Projeto Estudo do capim-elefante roxo (*Pennisetum Purpureum*) submetido à adubação orgânica e mineral, do Curso de Zootecnia/CCAA;
- XXIII. **Processo 5211/2009-80.** Estudo do *stylosanthes* campo grande em função de fontes de fósforo associado a aplicação de gesso e/ou calcário, do Curso de Zootecnia/CCAA;

**UNIVERSIDADE FEDERAL DO MARANHÃO**

Fundação Instituída nos termos da Lei 5.152 de 21/10/1966
São Luís – Maranhão

3

XXIV. **Processo 5413/2009-20** Projeto Disseminação ideológica: normas e usos da propaganda político-partidária no Brasil, do Departamento de Comunicação Social.
Dê-se ciência. Publique-se. Cumpra-se.
São Luís 16 de dezembro de 2009.


Prof. Dr. ANTONIO JOSÉ SILVA OLIVEIRA
Presidente em exercício



UNIVERSIDADE FEDERAL DO MARANHÃO
HOSPITAL UNIVERSITÁRIO
COMITÊ ÉTICA EM PESQUISA



PARECER CONSUBSTANCIADO

Parecer Nº223/2009

Pesquisador (a) Responsável: Antônio Augusto Moura da Silva

Equipe executora: Antônio Augusto Moura da Silva, Marco Antonio Barbieri, Heloisa Bettioli, Fernando Lamy Filho, Liberata Campos Coimbra, Maria Teresa Seabra S.B. e Alves, Raimundo Antonio da Silva, Valdinar Sousa Ribeiro, Vania Maria de Farias Aragão, Wellington da Silva Mendes, Zeni Carvalho Lamy, Mari Ada Conceição Saraiva, Alcione Miranda dos Santos, Arlene de Jesus Mendes Caldas, Cecília Claudia Costa Ribeiro, Silma Regina P. Martins, Flávia Raquel F. Nascimento, Marília da Glória Martins, Virginia P.L. Ferriani, Marisa Márcia M. Pinhata, Jacqueline P. Monteiro José S. Camelo Junior, Carlos Eduardo, Martinelli Júnior, Sonir Roberto R. Antonini e Aparecida Yulie Yamamoto

Tipo de Pesquisa: Projeto Temático

Registro do CEP: 350/08 Processo 4771/2008-30

Instituição onde será desenvolvido: Hospital Universitário, Maternidade Marly Sarney, Clínica São Marcos, Maternidade Benedito Leite, Maternidade Maria do Amparo, Santa Casa de Misericórdia do Maranhão, Maternidade Nazira Assub, Clínica São José e Clínica Luiza Coelho.

Grupo: III

Situação: APROVADO

O Comitê de Ética em Pesquisa do Hospital Universitário da Universidade Federal do Maranhão analisou na sessão do dia 20.03.08 o processo Nº. 4771/2008-30, referente ao projeto de pesquisa: **"Fatores etiológicos do nascimento pré-termo e consequências dos fatores perinatais na saúde de criança: coortes de nascimento em duas cidades brasileiras"**, tendo como pesquisadora responsável Antônio Augusto Moura da Silva, cujo objetivo geral é **"Investigar novos fatores na etiologia da prematuridade, utilizando-se abordagem integrada e colaborativa em duas cidades brasileiras numa coorte de conveniência, iniciada no pré-natal"**.

Tendo apresentado pendências na época de sua primeira avaliação, veio em tempo hábil supri-las adequada e satisfatoriamente de acordo com as exigências das Resoluções que regem esse Comitê. Assim, mediante a importância social e científica que o projeto apresenta a sua aplicabilidade e conformidade com os requisitos éticos, somos de parecer favorável à



UNIVERSIDADE FEDERAL DO MARANHÃO
HOSPITAL UNIVERSITÁRIO
COMITÊ ÉTICA EM PESQUISA



realização do projeto classificando-o como **APROVADO**, pois o mesmo atende aos requisitos fundamentais da Resolução 196/96 e suas complementares do Conselho Nacional de Saúde.

Solicita-se à pesquisadora o envio a este CEP, relatório parciais sempre quando houver alguma alteração no projeto, bem como o relatório final gravado em CD ROM.

São Luís, 08 de abril de 2009.

João Inácio L. de Souza
Prof. Dr. João Inácio Lima de Souza

Coordenador do Comitê de Ética em Pesquisa

Hospital Universitário da UFMA

Ethica homini habitat est

APÊNDICE

Quadro 3. Exemplos da configuração das variáveis do estudo

Bloco de estudos	Nome da variável	Tipo de variável	Valores possíveis	Autor/ Referência
Determinantes Sociodemográficos	Classificação socioeconômica- Abep	Categórica Ordinal	Classe A/B; Classe C; Classe D/E	Adhikari et al., 2019; Fergus et al., 2013; Khatibi t al., 2019 Ding, L., Yin, X., Wen, G. et al., 2024; Araújo et al, 2023
	Chave	Numérica discreta		
	Cidade	Categórica Nominal	Ribeirão Preto= 1; São Luís=2	
	Indivíduo com maior renda na família	Categórica Nominal	A entrevistada; Companheiro; Mãe; Pai; Avó; Avô; Madrasta; Padrasto; Tia; Tio; Irmã; Irmão; Outro; Não sabe	Adhikari et al., 2019; Fergus et al., 2013; Khatibi t al., 2019
	Cor da mãe	Categórica Nominal	Branca; Não Branca	Gao et al., 2019; Lacey et al., 2020; Koivu e Sairanen, 2020; Araújo et al, 2023
	Idade da mãe	Numérica discreta	Número natural	Gao et al., 2019; Khatibi et al., 2019; Lacey et al., 2020; Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024; Sun Q, Zou X, et al 2022; Araújo et al, 2023
	Cor da criança declarada pela mãe/responsável	Categórica Nominal	branca; preta/negra; parda/mulata/cabocla/morena; amarelo/oriental; indígena	Adhikari et al., 2019
	Peso do pai	Numérica discreta	Qualquer número real	Adhikari et al., 2019
	Altura do pai	Numérica discreta	Qualquer número real	Adhikari et al., 2019

	Faz atividade remunerada fora de casa	Categórica Nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Bazo, 2020
	Número de dias de trabalho por semana	Numérica discreta	Qualquer número real	Adhikari et al., 2019; Fergus et al., 2013; Bazo, 2020
	Ocupação da mãe	Categórica Nominal	Qualificação não-manual; Manual semi-especializada; Manual não-qualificada; Desempregado(a)	Adhikari et al., 2019; Fergus et al., 2013; Bazo, 2020
	Trabalha em pé	Categórica Nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Bazo, 2020
	Levanta peso	Categórica Nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Bazo, 2020
	Ocupação do Chefe da Casa	Categórica Nominal	Qualificação não-manual; Manual semi-especializada; Manual não-qualificada; Desempregado(a)	Adhikari et al. 2019; Fergus et al., 2013;
	Frequentou escola	Categórica nominal	Não; Sim	Adhikari et al., 2019; Koivu e Sairanen, 2020
	Sabe ler	Categórica nominal	Não; Sim	Adhikari et al., 2019; Koivu e Sairanen, 2020
	Situação conjugal da puérpera	Categórica nominal	Casada; União consensual (Mora junto); Solteira; Separada/desquitada/divorciada; Viúva	Adhikari et al., 2019
	Mora com companheiro	Categórica nominal	Não; Sim	Adhikari et al., 2019
	Último curso que fez	Categórica nominal	Alfabetização de jovens e adultos; Ensino fundamental ou 1º grau ;Ensino médio ou 2º grau; Superior graduação incompleto; Superior graduação completo; Não se aplica; Não sabe	Adhikari et al., 2019 Koivu e Sairanen, 2020
	Renda per capita	Numérica contínua	Qualquer número real	Adhikari et al., 2019; Fergus et al., 2013;
	Pessoa que realiza atividades domésticas	Categórica nominal	A sra. faz todo o trabalho; A sra. faz parte do trabalho; Não sabe	Adhikari et al., 2019

Determinantes Comportamentais	Tomou bebida alcoólica na gestação	Categórica nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Gao et al., 2019; Khatibi t al., 2019; Bazo, 2020
	Tem religião	Categórica nominal	Não; Sim	Adhikari et al., 2019
	Usou drogas ilícitas na gestação	Categórica nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Gao et al., 2019; Khatibi t al., 2019; Lacey et al., 2020; Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024
	Usou cocaína na gestação	Categórica nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Gao et al., 2019; Khatibi t al., 2019; Lacey et al., 2020; Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024
	Usou crack na gestação	Categórica nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Gao et al., 2019; Khatibi t al., 2019; Lacey et al., 2020; Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024
	Usou maconha na gestação	Categórica nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Gao et al., 2019; Khatibi t al., 2019; Lacey et al., 2020; Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024
	Usou merla na gestação	Categórica nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Gao et al., 2019; Khatibi t al., 2019;

				Lacey et al., 2020; Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024
	Fumou antes da gravidez	Categórica nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Gao et al., 2019; Khatibi t al., 2019; Lacey et al., 2020; Bazo, 2020; Koivu e Sairanen, 2020 Ding, L., Yin, X., Wen, G. et al., 2024
	Fumou durante a gravidez	Categórica Nominal	Não; Sim;	Adhikari et al., 2019; Fergus et al., 2013; Gao et al., 2019; Khatibi t al., 2019; Lacey et al., 2020; Bazo, 2020; Koivu e Sairanen, 2020 Ding, L., Yin, X., Wen, G. et al., 2024
	Duração de caminhada (mínimo 10 min/dia)	Numérica discreta	Número natural	Khatibi t al., 2019
	Dias da semana de prática de atividades moderadas (mínimo 10min/dia)	Numérica discreta	Número natural	Khatibi t al., 2019
Saúde materna, pré-natal e parto	Idade Gestacional em semanas	Numérica Contínua	Qualquer número real	Alleman et al., 2013; Fergus et al., 2013;
	Idade da menarca	Numérica discreta	Número natural	Alleman et al., 2013; Fergus et al., 2013;
	Número de gestações	Numérica discreta	Número natural	Adhikari et al., 2019; Fergus et al., 2013; Khatibi t al., 2019; Lacey et al., 2020 Koivu e Sairanen, 2020

	Idade Gestacional em dias	Numérica Contínua	Qualquer número real	Alleman et al., 2013; Fergus et al., 2013;
	Peso antes de engravidar	Numérica contínua	Qualquer número real	Adhikari et al., 2019; Fergus et al., 2013; Khatibi t al., 2019; Lacey et al., 2020 Koivu e Sairanen, 2020
	Peso ao final da gravidez	Numérica contínua	Qualquer número real	Adhikari et al., 2019; Fergus et al., 2013; Khatibi t al., 2019; Lacey et al., 2020 Koivu e Sairanen, 2020
	Ganho de peso durante a gestação	Numérica contínua	Qualquer número real	Adhikari et al., 2019; Fergus et al., 2013; Khatibi t al., 2019; Lacey et al., 2020 Ding, L., Yin, X., Wen, G. et al., 2024 Koivu e Sairanen, 2020
	Número de comorbidades	Numérica discreta	Número natural entre 0 e 10	Fergus et al., 2013; Gao et al., 2019; Bazo, 2020
	Comorbidades que acometeram a mãe durante a gestação (hipertensão, diabetes, toxoplasmose sífilis, herpes, sarampo, catapora, rubéola, anemia, ou infecção urinária)	Categórica Nominal	Nenhuma; Uma ou mais	Fergus et al., 2013; Gao et al., 2019; Bazo, 2020 Koivu e Sairanen, 2020 Ding, L., Yin, X., Wen, G. et al., 2024
	Mês de início da consulta pré-natal	Numérica discreta	Número natural entre 0 e 9	Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024; Araújo et al, 2023
	Quantidade consulta pré-natal no 1º trimestre	Numérica discreta	Número natural	Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024;

				Araújo et al, 2023
	Quantidade consulta pré-natal no 2º trimestre	Numérica discreta	Número natural	Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024; Araújo et al, 2023
	Consultas pré-natal	Numérica discreta	Número natural	Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024
	Quantidade consulta com enfermeiro	Numérica discreta	Número natural	Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024; Araújo et al, 2023
	Quantidade consulta com médico	Numérica discreta	Número natural	Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024; Araújo et al, 2023
	Médico ou enfermeira encaminhou para tratamento dentário durante a gravidez	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Bazo, 2020; Ding, L., Yin, X., Wen, G. et al., 2024; Araújo et al, 2023
	Anemia antes da gestação	Categórica Nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Khatibi t al., 2019; Lacey et al., 2020 Ding, L., Yin, X., Wen, G. et al., 2024 Koivu e Sairanen, 2020
	Anemia durante a gestação	Categórica Nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Khatibi t al., 2019; Lacey et al., 2020 Ding, L., Yin, X., Wen, G. et al., 2024 Koivu e Sairanen, 2020
	Diabetes antes ou durante a gestação	Categórica Nominal	Não; Sim	Fergus et al., 2013; Khatibi t al., 2019; Bazo, 2020 Koivu e Sairanen, 2020

	Hipertensão na gestação atual	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Gao et al., 2019; Khatibi et al., 2019; Bazo, 2020 Koivu e Sairanen, 2020
	Herpes durante a gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Bazo, 2020
	Infecção urinária durante a gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Lacey et al., 2020; Bazo, 2020
	Rubéola durante a gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Bazo, 2020
	Sarampo durante a gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Bazo, 2020
	Sífilis antes da gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Bazo, 2020; Koivu e Sairanen, 2020
	Sífilis durante a gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Bazo, 2020; Koivu e Sairanen, 2020
	Toxoplasmose antes da gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Bazo, 2020
	Toxoplasmose durante a gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Bazo, 2020
	Catapora durante a gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Fergus et al., 2013; Bazo, 2020
	Episódio de febre alta durante a gestação	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Bazo, 2020
	Tomou vacina contra tétano	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Bazo, 2020
	Queda durante a gestação	Categórica Nominal	Não; Sim	Fergus et al., 2013; Bazo, 2020; Koivu e Sairanen, 2020
	Apresentou sangramento vaginal durante a gestação	Categórica Nominal	Não; Sim	Yu QY et al., 2024

	Ameaça parto prematuro na gestação atual	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Yu QY et al., 2024
	Número de parto prematuro anterior	Numérica discreta	Número natural	Yu QY et al., 2024
	Ameaça de aborto na gestação atual	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Alleman et al., 2013; Fergus et al., 2013; Koivu e Sairanen, 2020
	Motivo ida ao hospital	Categórica Nominal	Sangramento vaginal; Perdeu líquido (água) vaginal; Sentiu contração ou dor ou cólica/ barriga endurecida; Febre/ infecção/infecção urinária; Bebê parou de mexer; A vinda foi agendada para esta data; O médico encaminhou; Outro; Não sabe	Fergus et al., 2013; Bazo, 2020
	Hospitalização durante a gestação	Categórica Nominal	Não; Sim	Adhikari et al., 2019; Fergus et al., 2013; Bazo, 2020
	Gravidez planejada	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Adhikari et al., 2019;
Saúde do recém-nascido	Sexo do RN	Categoria Nominal	Masculino; Feminino	Ding, L., Yin, X., Wen, G. et al., 2024
	Problema de saúde	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Khatibi et al., 2019; Ding, L., Yin, X., Wen, G. et al., 2024
	Defeito congênito	Categórica Nominal	Sim; Não; Não se aplica; Não sabe	Khatibi et al., 2019; Ding, L., Yin, X., Wen, G. et al., 2024
	Tipo de defeito congênito	Categórica Nominal	Gastrosquise; Onfalocele; Defeitos do tubo neural- mielomeningocele; Defeitos do tubo neural-encefalocele; Trato genito-urinário-unilateral; Trato genito-urinário- bilateral; Cardíaco; Outro	Khatibi et al., 2019; Ding, L., Yin, X., Wen, G. et al., 2024
	Nascimento pré-termo extremo	Categórica Nominal	Sim; Não	Gao et al., 2019;

Fonte: Instrumentos de coleta de dados – Coortes BRISA (2010)