

***CATCH-UP* ESTATURAL EM CRIANÇAS NOS
PRIMEIROS MIL DIAS DE VIDA:
UM ESTUDO DE PREDIÇÃO USANDO APRENDIZAGEM DE
MÁQUINA**

SÃO LUÍS, MA

2025

CIBELLY DE FÁTIMA VIEIRA FERREIRA

***CATCH-UP* ESTATURAL EM CRIANÇAS NOS PRIMEIROS MIL DIAS DE VIDA:
UM ESTUDO DE PREDIÇÃO USANDO APRENDIZAGEM DE MÁQUINA**

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

Orientador: Prof. Dr. Bruno Feres de Souza

Coorientador: Profa. Dra. Erika Bárbara Abreu
Fonseca Thomaz

SÃO LUÍS, MA

2025

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

Ferreira, Cibelly de Fátima Vieira.

Catch-up estatural em crianças nos primeiros mil dias de vida: : um estudo de predição usando Aprendizagem de Máquina / Cibelly de Fátima Vieira Ferreira. - 2025.
172 f.

Orientador(a): Bruno Feres de Souza Erika Barbara Abreu Fonseca Thomaz.

Dissertação (Mestrado) - Programa de Pós-graduação em Saúde Coletiva/ccbs, Universidade Federal do Maranhão, São Luís - Ma, 2025.

1. Saúde da Criança. 2. Antropometria. 3. Aprendizado de Máquina Supervisionado. 4. Aprendizagem Por Probabilidade. 5. Crescimento Corporal. I. Erika Barbara Abreu Fonseca Thomaz, Bruno Feres de Souza. II. Título.

***CATCH-UP* ESTATURAL EM CRIANÇAS NOS PRIMEIROS MIL DIAS DE VIDA:
UM ESTUDO DE PREDIÇÃO USANDO APRENDIZAGEM DE MÁQUINA**

Cibelly de Fátima Vieira Ferreira

Defesa de Dissertação de Mestrado apresentada em 28 de fevereiro de 2025 à banca examinadora constituída dos seguintes membros:

Banca Examinadora:

Prof. Dr. Bruno Feres de Souza
Orientador
Universidade Federal do Maranhão

Prof^a. Dr^a. Erika Barbara Abreu Fonseca Thomaz
Coorientadora
Universidade Federal do Maranhão

Prof. Dr. Thiago Augusto Hernandes Rocha
Examinador Externo
Organização Panamericana de Saúde, OPAS

Prof^a. Dr^a. Poliana Cristina de Almeida Fonseca Viola
Examinador Interno
Universidade Federal do Maranhão

Prof. Dr. Josenildo Costa da Silva
Examinador Externo (Suplente)
Instituto Federal do Maranhão

Dedico este trabalho a Deus e à minha família.

AGRADECIMENTOS

Agradeço a Deus por permitir viver e dar vida aos projetos que Ele tem para mim, e por me sustentar em meio às adversidades; às minhas mães Rosa e Zilar, meu pai Edelcy, minhas irmãs Izabelly, Edelcyany e Rosalcy, por todo amor, apoio e incentivo que nunca me faltaram, por acreditarem em mim, mesmo quando achei que não conseguiria. A Sirius, *in memoriam*, e Tito. A Eduardo Roger, meu amado, por toda força e paciência.

Aos meus amigos Alan Araujo, Lucas Daylor, Gabriel Angelo, Barbara Kaynny e Marcileia Almeida e aos outros amigos, que tive o privilégio de conhecer e conviver na jornada da vida. À toda a turma do Programa de Pós-Graduação em Saúde Coletiva (PPGSC), à qual criei importantes memórias afetivas e, em especial, a Rayane Sobrinho e Lilian Cavalcante.

Aos meus queridos professores, Prof. Dr. Bruno Feres e Profa. Dra. Erika Thomaz, os quais tive a honra de ser orientada. Obrigada por serem verdadeiros presentes divinos que trazidos pelo PPGSC, pelo grande apoio, parceria e compreensão frente a tantas mudanças e readaptações que vivi durante o mestrado.

Aos grupos de pesquisa pelas parcerias, a todos os professores do PPGSC pelos ensinamentos repassados e à banca examinadora pela dedicação de tempo e disponibilidade para a leitura deste trabalho.

À Universidade Federal do Maranhão (UFMA), ao PPGSC e sua Secretaria, bem como à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES, finance code 001) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, processo Nº 441882/2023-4), por oportunizarem a possibilidade de cursar este mestrado tão desejado.

“Se você nunca tentar, você nunca saberá”.
(Coldplay)

LISTA DE FIGURAS

Figura 1. Exemplos de problemas de classificação e de regressão.....	29
Figura 2. Exemplos de Subajuste, Ajuste adequado e Sobreajuste, respectivamente.....	30
Figura 3. Gráfico de Comprimento para Idade de 0 a 2 anos.....	36
Figura 4. Fluxograma dos estudos de coortes envolvidos, seus respectivos inícios, fases e quantitativo de participantes.....	37
Figura 5. Modelos teóricos do estudo com validações interna e externa, respectivamente.....	42
Figura 6. Matriz de confusão para problemas de duas classes.....	47
Figura 7. Modelos teóricos do estudo com validações interna e externa, respectivamente.....	58
Figura 8. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados completa.	67
Figura 9. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados SL.....	70
Figura 10. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados RP.....	73
Figura 11. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados SL como teste e RP como treino.....	76
Figura 12. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados RP como teste e SL como treino.....	79

LISTA DE TABELAS

Tabela 1. Condições sociodemográficas, características físicas, comportamentais e de saúde dos pais, dados do nascimento e do primeiro ano de vida das crianças da amostra, em ambas as cidades, n=659.....	60
Tabela 2. Resultado comparativo entre os algoritmos, considerando os desempenhos com e sem o selecionador de atributos Boruta, na base de dados conjunta.....	66
Tabela 3. Resultado comparativo entre os algoritmos, considerando SL.....	69
Tabela 4. Resultado comparativo entre os algoritmos, considerando RP.....	72
Tabela 5. Resultado comparativo entre os algoritmos, considerando treino em SL e teste em RP.....	75
Tabela 6. Resultado comparativo entre os algoritmos, considerando treino em RP e teste em SL.....	78

LISTA DE QUADROS

Quadro 1. Configuração das variáveis do estudo.....	39
Quadro 2. Hiperparâmetros dos algoritmos usados e suas respectivas descrições.....	46

LISTA DE ABREVIATURAS E SIGLAS

AGP	-	glicoproteína ácida alfa-1
AM	-	Aprendizagem de Máquina
APS	-	Atenção Primária à Saúde
AUC	-	<i>Area Under Curve</i>
BPN	-	Baixo Peso ao Nascer
BRISA	-	Estudo de Coortes Brasileiras de Nascimento de Ribeirão Preto e São Luís
CIUR	-	Crescimento Intrauterino Restrito
DOHaD	-	<i>Developmental Origins of Health and Disease</i>
DP	-	desvios-padrão
IDE	-	ambiente de desenvolvimento integrado
IG	-	Idade gestacional
IL1	-	Interleucina-1
IMC	-	Índice de Massa Corporal
IPAQ	-	questionário internacional de atividade física
k-NN	-	<i>K-Nearest Neighbors</i>
LDA	-	<i>Linear Discriminant Analysis</i>
MS	-	Ministério da Saúde

NASC	-	Nascimento
OMS	-	Organização Mundial da Saúde
PCR	-	Proteína C reativa
RLR	-	<i>Regularized Logistic Regression</i>
PN	-	Pré-natal
PT	-	Pré Termo
RN	-	Recém-Nascidos
RP	-	Ribeirão Preto
SL	-	São Luís
SVM	-	<i>Support Vector Machine</i>
TCLE	-	Termo de Consentimento Livre e Esclarecido
XGBoost	-	<i>Extreme Gradient Boosting</i>

FERREIRA, Cibelly de Fátima Vieira. *Catch-up* Estatural em Crianças nos Primeiros Mil Dias de Vida: um Estudo de Predição usando Aprendizagem de Máquina, 2025, Dissertação (Mestrado em Saúde Coletiva) - Programa de Pós-graduação em Saúde Coletiva, Universidade Federal do Maranhão, São Luís, 165.

RESUMO

Introdução: Apesar dos estudos de fatores associados, pouco se sabe sobre predições de *catch-up* estatural nos primeiros mil dias de vida, especialmente por Aprendizagem de Máquina (AM). **Objetivo:** Predizer a não ocorrência de *catch-up* estatural até 24 meses, de crianças que nasceram com comprimento abaixo do esperado para idade, por meio de técnicas de AM, a partir de dados sociodemográficos, do pré-natal, nascimento, hábitos, características dos pais e primeiro ano da criança. **Métodos:** Utilizaram-se os dados de pré-natal e nascimento *Brazilian Ribeirão Preto and São Luís cohorts* (BRISA), de duas cidades brasileiras. 659 crianças classificadas em baixa ou muito baixa estatura para idade ao nascer, segundo gráfico INTERGROWTH-21st, não gemelares, foram selecionadas e avaliadas. Aspectos hormonais, marcadores químicos de crescimento ou outras variáveis de difícil acesso/coleta não foram incluídos. No programa R studio, os algoritmos *Linear Discriminant Analysis* (LDA), *Regularized Logistic Regression* (RLR), *Support Vector Machine* (SVM) e *Extreme Gradient Boosting* (XGBoost) foram empregados para construir modelos de predição, com e sem o selecionador de atributos Boruta, em 5 diferentes cenários. **Resultados:** Neste estudo, a especificidade foi uma importante medida de desempenho. No cenário com junção de ambas as cidades para treino e teste (validação cruzada), os modelos em LDA e RLR com Boruta tiveram os melhores desempenhos em sensibilidade (72,8% e 72,7%), porém sem ganhos expressivos para acurácia $\cong$ 68 e 69%, especificidade $\cong$ 64% e 65%, precisão $\cong$ 68%, F1 score $\cong$ 70% e AUC $\cong$ 75%, cada. Entre todos, o cenário de São Luís, utilizando o LDA com Boruta teve o melhor desempenho de modo geral: acurácia de 70,9%, sensibilidade de 75%, precisão de 70,5%, F1 Score de 72,5% e AUC de 74,4%, porém com especificidade de 66,5%. A maior especificidade foi em RP, como treino, e SL, como teste (variando de 78,3% a 87,3%), no qual o LDA sem Boruta teve o maior valor obtido. Todos os modelos testados foram estatisticamente superiores à classe majoritária (até 51,5%). Das variáveis, o peso ao nascer, altura da mãe, estatura para idade ao nascer, peso para idade ao nascer, fumo tiveram grande contribuição para os modelos; grau de prematuridade apareceu

em todos os cenários. **Conclusão:** Variáveis de fácil acesso relativas aos primeiros mil dias de vida foram capazes de prever, com acurácia, o não *catch-up* estatural de crianças nascidas com déficit de comprimento até 24 meses. Intervenções precoces se tornam mais estratégicas, podendo causar impacto positivo ao longo da vida dessas crianças.

Palavras-chave: Saúde da Criança. Antropometria. Aprendizagem de Máquina Supervisionada. Aprendizagem por Probabilidade. Crescimento Corporal.

FERREIRA, Cibelly de Fátima Vieira. *Catch-up* Estatural em Crianças nos Primeiros Mil Dias de Vida: um Estudo de Predição usando Aprendizagem de Máquina, 2025, Dissertação (Mestrado em Saúde Coletiva) - Programa de Pós-graduação em Saúde Coletiva, Universidade Federal do Maranhão, São Luís, 165.

ABSTRACT

Introduction: Despite studies of associated factors, little is known about predictions of height catch-up in the first thousand days of life, especially through Machine Learning (ML).

Objective: To predict the non-occurrence of height catch-up up to 24 months of age, in children who were born with a length below the expected for their age, through ML techniques, based on sociodemographic, prenatal, birth, habits, parental characteristics and child's first year data. **Methods:** Prenatal and birth data from the Brazilian Ribeirão Preto and São Luís cohorts (BRISA) from two Brazilian cities were used. 659 children classified as short or very short for their age at birth, according to the INTERGROWTH-21st chart, non-twins, were selected and evaluated. Hormonal aspects, chemical growth markers or other variables that were difficult to access/collect were not included. In the R studio program, the Linear Discriminant Analysis (LDA), Regularized Logistic Regression (RLR), Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) algorithms were used to build prediction models, with and without the Boruta feature selector, in 5 different scenarios.

Results: In this study, specificity was an important measure of performance. In the scenario with both cities combined for training and testing (cross-validation), the models in LDA and RLR with Boruta had the best performances in sensitivity (72.8% and 72.7%), but without significant gains in accuracy $\cong$ 68 and 69%, specificity $\cong$ 64% and 65%, precision $\cong$ 68%, F1 score $\cong$ 70% and AUC $\cong$ 75%, each. Among all, the São Luís scenario, using LDA with Boruta, had the best performance in general: accuracy of 70.9%, sensitivity of 75%, precision of 70.5%, F1 Score of 72.5% and AUC of 74.4%, however specificity of 66.5%. The highest specificity was in RP, as training, and SL, as testing (ranging from 78.3% to 87.3%), in which LDA without Boruta had the highest value obtained. All models tested were statistically superior to the majority class (up to 51.5%). Of the variables, birth weight, maternal height, height for age at birth, weight for age at birth, and smoking had a great contribution to the

models; degree of prematurity appeared in all scenarios. **Conclusion:** Easily accessible variables related to the first thousand days of life were able to accurately predict the lack of height catch-up of children born with length deficit up to 24 months. Early interventions become more strategic and can have a positive impact throughout the lives of these children.

Keywords: Child Health. Anthropometry. Supervised Machine Learning. Probability Learning. Growth.

SUMÁRIO

1 INTRODUÇÃO.....	16
2 OBJETIVOS.....	19
2.1 Geral.....	19
2.2 Específicos.....	19
3 REFERENCIAL TEÓRICO.....	19
3.1 O Crescimento nos Primeiros Anos de Vida.....	20
3.2 Problemas no crescimento estatural.....	21
3.3 O catch-up estatural.....	22
3.4 Fatores relacionados ao crescimento e catch-up estatural.....	24
3.5 Aprendizagem de Máquina.....	28
4 MÉTODOS.....	33
4.1 Desenho do estudo.....	33
4.2 Local do estudo.....	33
4.3 Amostra do estudo BRISA.....	34
4.4 Amostra deste estudo e critérios de inclusão e exclusão.....	35
4.5 Instrumentos de coleta de dados na coorte BRISA.....	38
4.6 Variáveis do Estudo.....	38
4.7 Processamento e Análise dos Dados.....	42
4.8 Aspectos Éticos e Legais.....	48
5 RESULTADOS.....	49
6 CONSIDERAÇÕES FINAIS.....	83
REFERÊNCIAS.....	84
APÊNDICE A - Configuração das variáveis do estudo.....	93
ANEXOS.....	95

1 INTRODUÇÃO

Os estudos em epidemiologia do ciclo vital deram origem a um ramo da epidemiologia conhecido como Origens Desenvolvimentistas da Saúde e Doença ou *Developmental Origins of Health and Disease* (DOHaD) (Barker, 2007; Wadhwa et al., 2009). Segundo essa teoria, a exposição a determinados fatores durante a fase fetal e/ou nos primeiros anos de vida - período considerado crítico para a programação fisiológica - pode influenciar o desenvolvimento de doenças e agravos não transmissíveis, gerando repercussões profundas ao longo de todo o ciclo vital (Fleming et al., 2018; Juan; Yang, 2021).

Nesse sentido, o crescimento infantil é um relevante indicador de saúde da criança, favorecido ou dificultado por diferentes interações nos níveis genético, ambiental, econômico, social e atitudinal, que compõem o contexto de vida de cada sujeito (Sharma; Shastri; Sharma, 2016; Juan; Yang, 2021). Estudos de coortes de nascimento são fundamentais para estudar temas propostas pelo DOHaD, pois acompanham os indivíduos prospectivamente desde o nascimento (ou gestação) até a idade adulta, investigando possíveis influências em diferentes períodos da vida (Oken et al, 2023).

Uma das formas mais simples e comuns de acompanhar o crescimento é o registro periódico de aferições antropométricas, como peso e comprimento/altura, em gráficos internacionalmente aceitos, que permitem uma fácil identificação de desvios do crescimento (Villar, 2014; Brasil, 2021). No Brasil, essa avaliação contínua é feita preferencialmente por profissionais da Atenção Primária à Saúde (APS) via Caderneta da Criança, formulada pelo Ministério da Saúde, no Sistema Único de Saúde (SUS) (Palombo et al., 2014; Brasil, 2021).

Para indivíduos que nasceram pequenos para idade, seja devido ao baixo peso em geral ou, especificamente, devido à prematuridade, ou crescimento intrauterino restrito (CIUR), é possível observar o fenômeno do *catch-up* como uma propriedade do crescimento humano pelo qual a criança retorna à sua trajetória genética depois de um período de atraso no crescimento (Ong et al., 2000). Trata-se de um crescimento acelerado, a fim de recuperar um déficit prévio (Silveira et al, 2024), permitindo que esses indivíduos consigam nivelar o seu crescimento aos padrões esperados (Rodrigues et al., 2018).

No entanto, indivíduos nos quais não se observa essa “recuperação espontânea” requerem investigação. Por ser um parâmetro mais estável ao longo do tempo do que outros, como peso ou IMC, a estatura envolvendo idade representa o efeito cumulativo de situações

adversas sobre o crescimento da criança, sendo um relevante indicador de saúde da criança. É considerado como o índice mais sensível para aferir a qualidade de vida da população infantil (Sisvan, 2004). O catch-up estatural nos primeiros anos de vida está fortemente associado ao estado de saúde na infância e desenvolvimento cognitivo (Freitas et al., 2012).

Medidas antropométricas, quando registradas por ocasião do nascimento, refletem o desenvolvimento intrauterino (Brasil, 2012) e, ao decorrer dos anos de acompanhamento, podem indicar anormalidades e problemas graves, como a desnutrição crônica (Leroy; Frontillo, 2020; Soliman et al., 2021; Roro; Deressa; Lindtjorn, 2022). Déficits no crescimento em estatura estão associados a uma perpetuação de menores índices de altura, aumento da morbimortalidade, impactos neurocognitivos, motores, de autoestima, diminuição dos rendimentos na vida adulta, dentre outros efeitos negativos (Martinez et al., 2020; Soliman et al., 2021; Nguyen et al., 2021; Sartika et al., 2021).

Por isso, identificar dentro do período dos primeiros mil dias de vida (em média, 270 dias da gestação e 365 dos primeiros dois anos), podem tornar as intervenções ainda mais estratégicas por serem muito precoces, causando impacto positivo ao longo da vida dessas crianças. Para auxiliar nessa predição, a Aprendizagem de Máquina (AM) surgiu como uma opção viável para a análise de dados, ajudando a encontrar relações não triviais entre variáveis preditoras e o desfecho de interesse (Mello; Ponti, 2018). Sucintamente, AM é um campo da Inteligência Artificial que estuda métodos capazes de extrair conceitos a partir de amostras de dados. Assim, pode-se extrapolar o conhecimento aprendido para situações não vistas previamente.

De fato, alguns trabalhos utilizando AM para prever desfechos antropométricos nos primeiros anos de vida foram desenvolvidos recentemente, a saber: Sun et al., 2016; Kuhle et al., 2018; Li et al., 2020; Harrison et al., 2020; Lu et al., 2020; Saw et al., 2021; Tao et al., 2021; Adeeba et al., 2022; Bai et al., 2022; Khan et al., 2022; Islam Pollob et al., 2022; Bekele, 2022; Bitew, Sparks, Nyarko, 2022; Arayeshgari et al., 2023. Entretanto, de modo geral, percebe-se que tais estudos focaram em predições envolvendo o peso da criança como medida primária, em detrimento de outras grandezas antropométricas potencialmente úteis para avaliar o crescimento infantil, como a estatura.

Portanto, considerando a relevância desse tema, este estudo buscou realizar uma análise preditiva do crescimento de crianças que nasceram com comprimento abaixo do esperado, em coortes brasileiras (*Brazilian Ribeirão Preto and São Luís Cohort Studies - BRISA*), de realidades socioeconômicas contrastantes. Além disso, este trabalho se destaca por empregar a classificação de curvas INTERGROWTH-21st, adotada mundialmente e presente

na Caderneta da Criança, bem como utilizar dados do pré-natal, nascimento e pós-nascimento, de fácil acesso. As descobertas encontradas poderão ser usadas por profissionais do SUS que gerenciam o cuidado infantil. Apesar dos estudos acerca de fatores associados, não há estudos sobre predições acerca do não alcance da estatura esperada nos primeiros mil dias, especialmente por meio de Aprendizagem de Máquina (AM), com diferentes famílias de algoritmos. Dessa forma, objetivou-se construir e validar modelos de AM para prever a **não** ocorrência do *catch-up* estatural, contribuindo para a identificação precoce de crianças em risco e auxiliar na implementação de estratégias de intervenção.

2 OBJETIVOS

2.1 Geral

Predizer a não ocorrência de *catch-up* estatural até 24 meses de idade em crianças de duas coortes brasileiras, que nasceram com comprimento abaixo do esperado para idade, por meio de técnicas de Aprendizagem de Máquina (AM), a partir de determinantes sociodemográficos, dados de pré-natal, nascimento e primeiro ano da criança.

2.2 Específicos

1. Aplicar técnicas de AM para a predição de não *catch-up* estatural de crianças de até 24 meses de idade, utilizando dados sociodemográficos, de pré-natal, nascimento e primeiro ano da criança, considerando desfecho binário;
2. Comparar diferentes tipos de algoritmos de AM para identificar aqueles com melhores desempenhos;
3. Avaliar diferentes cenários de predição, considerando os dados de coortes de São Luís e Ribeirão Preto e;
4. Identificar os principais preditores relacionados à estatura das crianças da amostra.

3 REFERENCIAL TEÓRICO

Este capítulo apresenta o referencial teórico que fundamenta a temática estudada. Inicialmente, na Seção 3.1, discute-se o crescimento infantil nos primeiros anos de vida, com ênfase no contexto epigenético e nas oportunidades oferecidas pelo período crítico dos primeiros mil dias. Em seguida, na Seção 3.2, são abordados os problemas de crescimento e suas implicações ao longo do ciclo vital. Na sequência, nas Seções 3.3 e 3.4, são vistos o que é o *catch-up* estatural, a motivação para estudá-lo e os fatores relacionados a esse fenômeno. A Seção 3.5 trata da Aprendizagem de Máquina e estudos sobre o tema envolvendo essa metodologia, considerando a complexidade do processo de crescimento e sua aplicação na área da saúde. Por fim, a Seção 3.6 aborda sobre a Política Nacional de Atenção Integral à Saúde da Criança, na qual este trabalho se insere.

3.1 O Crescimento nos Primeiros Anos de Vida

O crescimento humano nos primeiros anos de vida é caracterizado por mudanças rápidas de tamanho, forma e maturação de órgãos e sistemas (Cameron; Schell, 2021). É um processo complexo e contínuo, mas não linear (Richmond; Rogol, 2014). Sofre influência de fatores intrínsecos e extrínsecos, que alteram a expressão de alguns genes, resultando em remodelação estrutural e funcional (Villares, 2016; Sharma; Shastri; Sharma, 2016).

Essas modificações epigenéticas, ou seja, as alterações de como o ambiente e o estilo de vida podem influenciar a ativação ou silenciamento de genes, podem ser transmitidas entre gerações (Noro; Gon, 2015). Dessa forma, desempenham um papel de reprogramação e, ao mesmo tempo, representam uma janela de oportunidade para adquirir hábitos saudáveis, diminuindo a ocorrência de doenças futuras (Villares, 2016; Karakochuk et al., 2017; Fleming et al., 2018; Juan; Yang, 2021).

Nos últimos anos, houve um aumento de pesquisas relacionadas à programação precoce, definida como o processo pelo qual a exposição a uma série de fatores externos, nas fases iniciais do desenvolvimento, produz mudanças fisiológicas permanentes que podem interferir em desfechos de saúde futuros (González-Muniesa et al., 2017; Perez-Muñoz et al., 2022). Nesse contexto, para os primeiros mil dias de vida, que compreendem os cerca de 270 dias da gestação com os 365 dias de cada ano de vida, sugere-se que a promoção de saúde se estenda ao início da vida intrauterina, aproveitando o período de plasticidade dos primeiros mil dias (Black; Trude; Kowalski, 2021).

Estudos de coorte sugerem que o crescimento nos primeiros mil dias de vida pode ser visto como um *continuum* entre o período fetal e a primeira infância (Roro; Deressa; Lindtjorn, 2022). Esse fenômeno ocorre em padrões distintos quanto ao ritmo e momento, dependendo do sexo e da faixa etária. A literatura aponta que a fase de maior crescimento humano ocorre justamente no primeiro ano de vida. Após adquirir comprimento e peso durante a gestação, a criança continua crescendo de maneira acelerada durante os dois primeiros anos de vida, mas esse ritmo desacelera gradualmente até que, na puberdade, acontece o estirão puberal (Richmond; Rogol, 2014; SBP, 2017).

Embora exista uma questão genética, o crescimento infantil é fortemente marcado por fatores ambientais (Romani; Lira, 2004), como na fase fetal, no qual o tamanho ao nascer é determinado mais pela nutrição materna e por fatores intra uterinos e placentários do que pela composição genética.

3.2 Problemas no crescimento estatural

Condições adversas de crescimento no início da vida, como CIUR, baixo peso ao nascer (BPN) e pré-termo (PT) (Sharma, Shastri E Sharma, 2016; Sartika et al., 2021) são fatores de risco para um posterior falha no crescimento das crianças. Definida como a taxa de crescimento fetal abaixo do esperado diante do potencial de crescimento, o CIUR tem etiologia variada, mas ocorre especialmente devido a problemas relacionadas à mãe, que podem ser identificadas nas consultas de pré-natal (Zohdi et al., 2012; Sharma; Shastri; Sharma, 2016).

Essa limitação patológica representa um grande risco de morbimortalidade para os bebês, podendo causar hipóxia e sofrimento fetal (Ryckman; Berberich; Dagle, 2015), deficiências no desenvolvimento neurológico (Miller; Huppi; Mallard, 2016), distúrbios do apetite e riscos metabólicos para doenças crônicas (Nardoza et al., 2017). Não é incomum as crianças continuarem com tamanho corporal reduzido quando comparadas às sem CIUR, apesar de boa parte delas conseguir recuperá-lo nos primeiros dois anos de vida (Batista et al., 2012).

Por sua vez, o BPN é a condição de nascer com peso inferior a 2500 gramas e, em geral, contribui para o déficit de crescimento e desenvolvimento pós-natal, dificultando a amamentação dessas crianças e tornando-as mais vulneráveis à ocorrência de doenças frequentes e prolongadas com sequelas importantes e até morte (Rahman et al., 2016).

A prematuridade, ocorrida em partos com menos de 37 semanas completas, está associada a resultados adversos de curto e longo prazo, sendo a principal causa de morbidade e mortalidade perinatal em países desenvolvidos e em desenvolvimento, produzindo aumentos dos custos públicos de saúde (Jiang et al., 2018). Essa condição, bem como baixo peso ao nascer, diarreia e cobertura vacinal básica completa foram associadas ao atraso no crescimento infantil (Sartika et al., 2021).

Ao longo do desenvolvimento, crianças podem apresentar falhas no crescimento linear (estatura para idade). Esse retardo ocorre quando o indivíduo não atinge o seu potencial de crescimento, calculado pela altura média das alturas dos pais e ajustada ao sexo da criança (Richmond; Rogol, 2014), ficando, muito baixas para sua idade (Leroy; Frontillo, 2019). A baixa estatura é definida como uma altura 2 desvios-padrão (DP) ou mais abaixo esperado para sexo e idade nos padrões de crescimento infantil da Organização Mundial da Saúde (OMS) (Brasil, 2012).

Déficits no crescimento em estatura estão associados a uma perpetuação de menores índices de altura, aumento da morbimortalidade, menor rendimento escolar no futuro, redução da produtividade econômica e, no caso das mulheres, descendentes com menor peso ao nascer, dentre outros efeitos negativos (Brasil, 2012; Martinez et al., 2020; Soliman et al., 2021; Nguyen et al., 2021).

3.3 O *catch-up* estatural

O *catch-up* estatural é a “recuperação” espontânea do crescimento de crianças, especialmente em relação ao peso e à altura. A estatura é um indicador mais estável ao longo do tempo do que outros parâmetros, como peso ou IMC. Em indivíduos que nasceram com comprimento abaixo do esperado, pelas razões anteriormente abordadas, é possível observar o fenômeno do *catch-up*.

Para Ong et al., 2000, o *catch-up* é uma propriedade do crescimento humano pelo qual a criança retorna à sua trajetória genética depois de um período de atraso no crescimento. A expressão, adotada por Prader et al. (1963), refere-se a um crescimento rápido, como um ganho superior a 0,67 escore Z em relação à média da população de referência (Ong et al., 2000), permitindo que esses indivíduos consigam nivelar o seu crescimento aos padrões esperados (Rodrigues et al., 2018).

Geralmente, ocorre primeiro no perímetro cefálico, depois no comprimento e por fim no peso (Rugolo, 2005). Embora o crescimento de recuperação possa ser expresso em termos

de velocidade de alteração de altura, a mudança na pontuação do desvio padrão da altura é mais apropriada (Wit; Boersma, 2002). É importante acompanhar o crescimento da criança, pois um crescimento acelerado ou excessivo pode estar associado a doenças metabólicas, cardiovasculares e obesidade na vida adulta (Rossi, 2010).

A maior variação nas taxas de ganho de peso é vista nos primeiros 1-2 anos de vida, quando os bebês podem apresentar uma taxa significativa de crescimento de recuperação. Essas taxas de crescimento variáveis geralmente compensam o RCU ou o aumento do crescimento fetal e, aos dois anos de idade, o crescimento geralmente segue a trajetória genética (Ong et al, 2000).

Para facilitar esse acompanhamento infantil, existem curvas e tabelas específicas, elaboradas a partir de diferentes populações (Kale et al., 2018; Hocquette et al., 2021; Brasil, 2021). Esses gráficos são obtidos a partir do cálculo entre a idade da criança e variáveis como o peso, a altura e o perímetro da cabeça, e permitem observar os indicadores de crescimento, desenvolvimento e estado nutricional infantis. Neles os percentis são mostrados como linhas curvas. O ideal é que o ponto de encontro das linhas esteja sempre perto da linha central da curva, que se refere à média da população.

A expectativa quanto ao crescimento de recém-nascidos prematuros é que a maioria apresente *catch-up*, atingindo seu canal de crescimento entre os percentis de normalidade nas curvas de referência até os 2-3 anos de idade (Vila; Macedo; Melo, 2017). Exemplos são as curvas de crescimento americanas de Alexander et al. (1996) e Oken et al. (2003), de Kramer (2001), da OMS (2006), de Fenton (2013) e Villar (2014).

Atualmente, para crianças menores de 5 anos, a curva da OMS de 2006 tem sido mais comumente utilizada. Padrões internacionais de crescimento foram descritos para estabelecer o crescimento normal esperado em condições ambientais ideais a fim de avaliar crianças em qualquer lugar, independente de etnia, status socioeconômico e tipo de alimentação. Para isso, foi realizado um estudo de base populacional com gestantes de diferentes continentes - África, Américas, Ásia e Europa. Participaram mulheres saudáveis, com estado nutricional e assistência pré-natal adequados, residentes e parturientes em áreas geográficas urbanas (OMS, 2006).

Posteriormente, o Consórcio Internacional de Crescimento Fetal e Neonatal para o século XXI ou INTERGROWTH-21st, uma rede de pesquisadores, clínicos e instituições, complementou a curva da OMS, visando a prescrever padrões de crescimento para o período fetal, neonatal e pós-natal de recém nascidos (RN) pré-termo, estendendo-se aos dois anos de

idade. Participaram do estudo, RNs do Brasil, China, Índia, Itália, Quênia, Oman, Reino Unido e Estados Unidos, que, apesar de terem alguns fatores de risco para CIUR, não apresentavam malformações congênitas ou evidências ultrassonográficas de danos antes do nascimento (Villar et al., 2014; Villar et al., 2016).

Na prática, um relevante instrumento de acompanhamento infantil brasileiro é a Caderneta da Criança, elaborada pelo Ministério da Saúde (MS) (Palombo et al., 2014). Os gráficos recomendados pela OMS foram formulados pelo estudo INTERGROWTH-21st para crianças de 0 a 2 anos e pré-termos. Eles abrangem índices como peso por idade, comprimento/estatura por idade, peso por estatura/comprimento e índice de massa corporal (IMC) por idade. Para cada um dos índices, estabeleceram-se valores de referência de escore Z para sexo e faixa etária (Brasil, 2021).

A Caderneta se constitui como uma ferramenta essencial para garantir a integralidade na assistência à criança e, desde 2005, é disponibilizada gratuitamente no SUS (Brasil, 2005). Dessa forma, o acompanhamento e registro periódico na Caderneta adquirem um papel vital na promoção do desenvolvimento adequado e na prevenção de possíveis atrasos (Palombo et al., 2014). Um padrão normal de crescimento geralmente sugere boa saúde geral, enquanto o crescimento mais lento que o normal aumenta a possibilidade de uma doença subaguda ou crônica subjacente, incluindo uma causa endocrinológica (Richmond; Rogol, 2014).

Existe uma orientação feita pelo MS quanto à forma de aferir o comprimento/altura, desfecho de interesse deste trabalho. Menores de 2 anos devem ser avaliados deitados (comprimento) e crianças com 2 anos ou mais devem ser medidas de pé (altura). Tal particularidade deve ser observada, pois existe uma diferença de 0,7cm entre a estatura da criança medida deitada e em pé. Assim, se uma criança de 2 ou mais anos tiver sua estatura aferida deitada, o valor encontrado deve ser diminuído de 0,7 antes de ser registrado no gráfico recomendado para 2 a 5 anos. Similarmente, se um menor de 2 anos for medido de pé, o valor encontrado deve ser acrescido de 0,7 antes de ser registrado no gráfico de 0 a 2 anos (Brasil, 2021).

3.4 Fatores relacionados ao crescimento e *catch-up* estatural

O crescimento humano possui uma natureza multifatorial, cujos mecanismos que o explicam estão relacionados a características genéticas partilhadas, efeitos epigenéticos, programação de alterações metabólicas (Onis; Branca, 2016) e fatores socioculturais, num

processo cíclico iniciado no útero (Millward, 2017). Alguns fatores que influenciam o crescimento e, consequentemente, o *catch-up* são mencionados.

Durante a fase intrauterina, um dos principais objetivos da consulta de pré-natal é acompanhar os indicadores de crescimento, que possibilitam identificar situações atípicas e evitar complicações futuras (Lu et al., 2020). Isso porque o início precoce das consultas pré-natais permite detectar fatores de risco e diagnosticar morbidades ocorridas na gestação, por meio da realização de exames físicos, imaginológicos e laboratoriais (Kale et al., 2018). O MS recomenda que minimamente sejam realizadas seis consultas de pré-natal, sendo uma no primeiro trimestre da gravidez, duas no segundo e três no terceiro (Brasil, 2013).

A ultrassonografia é o exame mais comumente utilizado para a triagem do crescimento fetal e avaliação de condições fisiológicas e patológicas feita no pré-natal (Oghli et al., 2021). Por meio dela, algumas informações podem ser obtidas, como a idade gestacional (IG) e o sexo, além de parâmetros da biometria fetal, como a circunferência cefálica, o diâmetro biparietal, a circunferência abdominal e o comprimento do fêmur, definidos com base na anatomia do feto. Medições ocorridas em série possibilitam a comparação com as curvas e valores de referência (Kale et al., 2018; Lu et al., 2020, Oghli et al., 2021).

O primeiro parâmetro a ser determinado no início da gravidez é a IG, que pode ser estimada pela data da última menstruação ou, de forma mais assertiva, por ultrassonografias até a 20ª semanas (Who, 2006; Silva et al., 2014; Ryckman; Berberich; Dagle, 2015; Kale et al., 2018). A IG é um forte preditor do resultado neonatal, importante para auxiliar na identificação de risco para parto prematuro, além de tomadas de decisão nos cuidados clínicos e pesquisas de saúde perinatal (Ryckman; Berberich; Dagle, 2015).

Roro, Deressa e Lindtjorn (2022) investigaram a influência de fatores intrauterinos no peso ao nascer e no crescimento linear infantil e apontaram como variáveis preditoras a biometria fetal (medida por ultrassonografia), sexo da criança, IG ao nascer, idade materna, escolaridade materna, ocupação materna, paridade, altura materna, situação econômica da família, tamanho da família e idade da criança. O estudo revelou que o crescimento linear foi influenciado pelo crescimento fetal, duração da gravidez, altura materna e tamanho da família.

Segundo o Manual de Gestação de Alto Risco (Brasil, 2022), algumas características podem trazer risco aumentado para desfechos desfavoráveis em saúde, inclusive com potencial de óbito materno-fetal. Mulheres menores de 15 e maiores de 40 anos (>35 anos já indica um risco intermediário), baixo peso no início da gestação (IMC <18), situação conjugal

insegura e insuficiência de apoio familiar, baixa escolaridade (<5 anos de estudo), altura menor que 1,45 m, uso de tabaco, álcool outras drogas, gestantes negras (cor de pele preta ou parda), alterações no crescimento intrauterino (restrição de crescimento fetal e macrosomia), entre outros, devem ser consideradas nos critérios de cuidado da APS com apoio de equipe multiprofissional ou apoio de ambulatório pré-natal de alto risco.

A nutrição do feto é um aspecto fundamental para o seu crescimento. Ela é influenciada pela função da placenta, que leva nutrientes e oxigênio da mãe para o bebê, e pode resultar em bebês com peso menor do que o esperado, quando disfuncionais (Sharma, Shastri E Sharma, 2016). No entanto, outros fatores de risco maternos podem interferir no crescimento fetal, como doenças crônicas, anemia, infecções e hábitos nocivos. Qualquer incompatibilidade entre o fornecimento de nutrientes pela placenta e a demanda do feto também leva à CIUR, em qualquer fase da gravidez (Yu et al., 2013; Nardoza et al., 2017; Kale et al., 2018; Tao et al., 2021).

Outros fatores de risco associados a resultados desfavoráveis no nascimento já foram descritos na literatura. Entre eles estão o tabagismo materno, idade materna jovem, distúrbios hipertensivos, pré-eclâmpsia, diabetes mellitus pré-gestacional (Rotem et al., 2020, Roro; Deressa; Lindtjorn, 2022), nascimentos múltiplos, uso de drogas ilícitas, trabalho exaustivo, baixo índice de massa corporal, ganho de peso insuficiente e histórico de partos prematuros (Wadhwa et al., 2001; Silva et al., 2014).

Para Yu et al. (2013), o baixo peso pré-gestacional aumenta o risco de ter um bebê com BPN, assim como o sobrepeso/obesidade pré-gestacional aumenta o risco para bebês com macrosomia (peso ao nascer > 4.000 g em um lactente a termo). Monali et al. (2021) encontraram a multiparidade, IMC pré-gravidez e a proteína A plasmática, uma enzima relacionada à placenta e tecidos adjacentes, que libera fator de crescimento semelhante à insulina, como preditoras de risco para nascimento de macrosômicos.

Mães com baixa estatura e especialmente adolescentes têm maior probabilidade de ter bebês com BPN e, posteriormente, maior probabilidade de apresentar falhas de crescimento durante a infância (Christian et al., 2013). Os bebês com BPN têm maior vulnerabilidade a infecções, como diarreia e infecções respiratórias inferiores, e pelo aumento do risco de complicações, incluindo apneia do sono, icterícia, anemia, doenças pulmonares crônicas, fadiga e perda de apetite em comparação com crianças com peso normal ao nascer (Rahman et al., 2016).

O comprimento ao nascer também é indicado como preditor de retardo de crescimento (Brasil, 2012; Ryckman; Berberich; Dagle, 2015). O estudo de Harrison et al. (2020) o mostrou como sendo um preditor significativamente importante do crescimento, enfatizando a importância das intervenções nos primeiros meses de vida para prevenir o atraso dos ganhos em altura subsequente e suas consequências.

Ademais, infecções agudas graves, (particularmente aquelas que envolvem o trato gastrointestinal, mas também do respiratório inferior), e crônicas prejudicam o crescimento linear (Wong et al., 2016; Millward, 2017). Os efeitos combinados e interativos das infecções, dos fatores ambientais e da desnutrição como possíveis determinantes do atraso no crescimento em crianças têm sido considerados de grande importância. Estas interações reforçam-se mutuamente através da infecção que exacerba qualquer desnutrição, devido à supressão do apetite, redução da ingestão de alimentos e má absorção de nutrientes (Thomson et al., 2014).

A desnutrição, por sua vez, reduz os sistemas de defesa imunitária, agravando assim a influência adversa das infecções. A infecção e a inflamação associada têm influências inibitórias diretas nos processos anabólicos em todo o organismo, incluindo a placa de crescimento, são adicionais à inibição do crescimento através da desnutrição (Millward, 2017).

Uma das formas de combater a fome, pobreza e desnutrição ocorre por programas de transferência de renda. No Brasil, instituído pela Lei nº 10.836, de 9 de janeiro de 2004, o **Bolsa Família** é um desses programas do Governo Federal. Distribuído a partir de condicionalidades, foi criado com o intuito de apoiar as famílias mais pobres e garantir o direito aos serviços sociais básicos, especialmente relacionados à saúde e educação, mediante o repasse mensal direto de valores pré-estabelecidos (Lima, 2018), que podem ter impacto significativo no crescimento infantil ao proporcionar, entre outros, uma alimentação mais adequada.

Para Neves, Ferreira e Welch (2021), o déficit estatural, indicativo de desnutrição crônica em estudos epidemiológicos, foi a condição mais prevalente entre crianças quilombolas do nordeste brasileiro menores de 5 anos. Os resultados do estudo evidenciaram a centralidade do contexto social da criança, que afeta o crescimento linear. Condições maternas, estimulação familiar, padrão cultural e racial, nível educacional e socioeconômico da família também estão relacionados à etiologia, desenvolvimento e manutenção do crescimento (Zebral, 2014; Silva et al., 2014; Karakochuk et al., 2017; Martinez, 2020).

A análise multivariada de Akombi et al. (2017) revelou que os fatores de risco significativos mais consistentes para o atraso no crescimento e o atraso no crescimento grave entre crianças de 0-23 meses e 0-59 meses são: sexo da criança (masculino), tamanho percebido da mãe ao nascer (pequeno e médio), índice de riqueza familiar (famílias pobres e mais pobres), duração da amamentação (mais de 12 meses), zona geopolítica (Nordeste, Noroeste, Centro-Norte) e crianças que foram relatadas como tendo tido diarreia nas 2 semanas anteriores ao inquérito aplicado.

A amamentação exerce uma influência significativa no crescimento e desenvolvimento infantil até os dois anos de idade. O leite materno fornece todos os nutrientes essenciais, como proteínas, gorduras, vitaminas e minerais, necessários para o desenvolvimento saudável da criança. Além disso, contém anticorpos que fortalecem o sistema imunológico, protegendo contra infecções e doenças comuns na infância (Brasil, 2022).

Estudos científicos corroboram esses benefícios. Santos et al (2019) em seu estudo, em Recife-PE, destacaram a importância do aleitamento materno exclusivo nos primeiros seis meses de vida, seguido da introdução de alimentos complementares adequados, mantendo o aleitamento até os dois anos ou mais. Além dos benefícios nutricionais e imunológicos, Spyrides et al (2005) revelaram o efeito da duração da amamentação predominante sobre o perfil antropométrico de crianças, indicando estar associada a padrões de crescimento saudáveis durante o primeiro ano. Outro estudo transversal retrospectivo analisou prontuários de 401 crianças e encontrou que aquelas que não foram amamentadas exclusivamente até os seis meses apresentaram maior prevalência de doenças nos primeiros dois anos de vida, reforçando a associação entre a duração do aleitamento materno e a redução de doenças prevalentes na infância (Nass et al, 2022).

3.5 Aprendizagem de Máquina

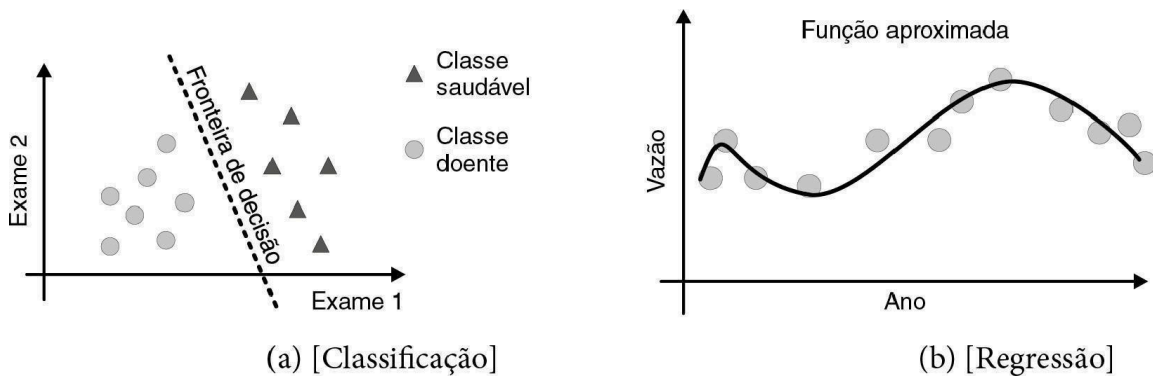
A Aprendizagem de Máquina (AM), do inglês *Machine Learning*, pode ser considerada um campo de pesquisa que busca compreender e reproduzir mecanismos naturais de aprendizagem, a fim de desenvolver sistemas aptos a lidar com problemas reais (Mello; Ponti, 2018). Para isso, a AM estuda abordagens que podem aumentar seu desempenho por meio da experiência, com base em modelos aprendidos a partir dos dados.

Dentre as estratégias para a construção de modelos de AM, destaca-se a Aprendizagem Indutiva, segundo a qual um conceito geral é inferido a partir de observações

particulares. Assim, um padrão pode ser automaticamente descoberto analisando-se um conjunto representativo de dados. Essa abordagem tem sido utilizada para a concepção de algoritmos de AM capazes de tratar diversos tipos de tarefas. Tais tarefas são agrupadas em duas grandes categorias principais: predição e descrição (Faceli et al., 2021). Na primeira, o objetivo é encontrar um mapeamento adequado nos dados disponíveis que possa ser utilizado para prever um valor que caracterize uma nova observação. Já na segunda, pretende-se explorar ou descrever um conjunto de dados.

Em estudos de predição, algoritmos de AM supervisionada são comumente considerados (Mello; Ponti, 2018). Neles, a aprendizagem ocorre a partir de conjuntos de dados rotulados, em que as observações são descritas por variáveis preditoras (ou independentes) e os respectivos desfechos (ou variável dependente). O algoritmo então aprende um modelo capaz de explicar o fenômeno em estudo. Se os desfechos pertencerem a um conjunto discreto de valores, tem-se um problema de classificação. Porém, se pertencerem a um conjunto infinito e ordenado de valores, tem-se um problema de regressão. A Figura 1 ilustra esses dois tipos de problemas.

Figura 1. Exemplos de problemas de classificação e de regressão.

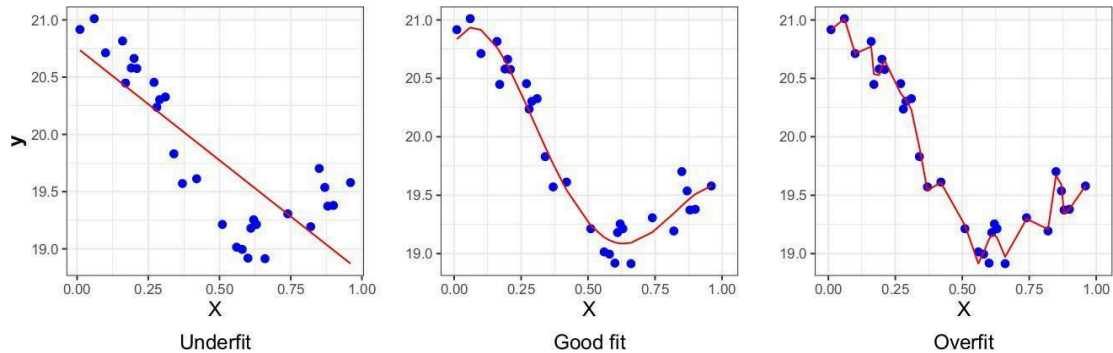


Fonte: Faceli et al. (2021)

Neste trabalho, algoritmos de classificação são empregados para prever a ocorrência ou não de *catch-up* estatural adequado, com base em fatores sociodemográficos, pré-natais, perinatais e pós-natais. A escolha do algoritmo, no entanto, não é trivial, pois diferentes abordagens de AM possuem vieses indutivos distintos, o que afeta a forma como os modelos generalizam os dados (Badillo et al., 2020). Por exemplo, um algoritmo pode gerar um modelo muito especializado aos dados disponíveis. Nesse caso, quando da aplicação do

modelo a dados não vistos, o desempenho preditivo tende a ser baixo, caracterizando o chamado sobreajuste (*overfitting*). Outro caso problemático ocorre quando o algoritmo não ajusta bem nem os dados vistos, em um subajuste (*underfitting*). A Figura 2 ilustra ambas as situações e destaca ao centro o cenário ideal.

Figura 2. Exemplos de Subajuste, Ajuste adequado e Sobreajuste, respectivamente



O exemplo ilustra um caso de regressão simples. Os pontos azuis são os dados e as linhas vermelhas, os ajustes do modelo. Fonte: Badillo et al (2020)

Portanto, para determinar os algoritmos mais adequados para a tarefa preditiva deste estudo, é necessário a realização de uma análise comparativa. Entre as abordagens consideradas, foram selecionados métodos bem conhecidos da comunidade de AM e de uso já estabelecido na área da Saúde (Badillo et al., 2020; Lo Vercio et al., 2020; Sarker, 2021), a saber: *Linear Discriminant Analysis*, *Regularized Logistic Regression*, *Support Vector Machines* e *eXtreme Gradient Boosting*. Eles abrangem diferentes famílias de classificadores, que permitem explorar os padrões nos dados de acordo com cada perspectiva única. A seguir, uma breve explicação sobre cada algoritmo é fornecida

O *Linear Discriminant Analysis* (LDA) é uma técnica de classificação que busca encontrar a combinação linear das variáveis preditoras que melhor separa as classes de um conjunto de dados. Ele opera projetando os dados originais em um espaço de menor dimensão, definido de acordo com dois critérios: a maximização da distância entre as médias das classes e a minimização da variância dentro de cada classe. A fronteira de decisão resultante é então aquela formada pelos eixos que otimizam a separação das classes. Devido à sua simplicidade, o LDA é frequentemente utilizado como referência para avaliar o desempenho de abordagens mais sofisticadas. Maiores detalhes sobre esse método podem ser encontrados no trabalho de Sarker (2021).

O *Regularized Logistic Regression* (RLR) é uma extensão da regressão logística tradicional, utilizada para modelar a relação entre uma variável dependente binária (como a ocorrência ou não de *catch-up* estatural) e um conjunto de variáveis independentes (por

exemplo, fatores pré-natais). Seu objetivo é ajustar uma função logística que estima a probabilidade de uma observação pertencer a uma determinada classe. Em conjuntos de dados de alta dimensionalidade, a regressão logística convencional pode sofrer com o sobreajuste, o que motivou a incorporação de técnicas de regularização, como L1 e L2. Essas técnicas adicionam uma penalização à função de custo do modelo, reduzindo o peso de certas variáveis ou eliminando-as completamente, o que melhora a generalização e a interpretabilidade do modelo. Lo Vercio et al., 2020 apresentam mais informação sobre o método.

As *Support Vector Machines* (SVM) são algoritmos de AM baseados na Teoria do Aprendizado Estatístico. Elas implementam o princípio da minimização estrutural, que busca aumentar a capacidade de generalização do modelo mantendo sua complexidade sob controle. Para isso, três aspectos são considerados. O primeiro é a construção do hiperplano separador ótimo, que maximiza a distância entre observações de classes distintas, o que tende a melhorar o resultado de classificação para dados novos. O segundo são os vetores de suporte, que correspondem às observações mais próximas ao hiperplano. Focando apenas nesses pontos nesses pontos mais relevantes, o modelo se torna mais eficiente. Por fim, o terceiro aspecto é o uso das funções de *kernel*, que mapeiam os dados para um espaço de maior dimensão onde os dados são melhor separados. Isso permite a construção de fronteiras de decisão não lineares e capazes de capturar a complexidade dos dados. Badillo et al. (2020) fornece boa base para o entendimento das SVMs.

O *eXtreme Gradient Boosting* (XGBoost) é, diferentemente das abordagens anteriores, baseado no conceito de combinação de classificadores. Ele utiliza como base as árvores de decisão, que correspondem a modelos de classificação estruturados de forma hierárquica, onde seus nós internos representam testes sobre variáveis do problema e os ramos indicam possíveis caminhos a serem seguidos até a predição. A depender da complexidade dos dados, árvores de decisão podem sofrer de sobreajuste, o que limita seu poder de generalização. Para lidar com o problema, o XGBoost emprega a técnica de *boosting*, na qual modelos fracos são combinados sequencialmente, com cada novo modelo focado em corrigir os erros cometidos pelo anterior. Ao final do processo, os modelos são agregados de forma ponderada, atribuindo maior peso àqueles com melhor desempenho, resultando em um classificador final mais preciso. Vide Sarker (2021) e referências ali contidas para um melhor entendimento da técnica.

3.6 Política Nacional de Atenção Integral à Saúde da Criança

Este trabalho está integrado ao Sistema Único de Saúde (SUS) por meio da Política Nacional de Atenção Integral à Saúde da Criança (PNAISC), instituída pelo Ministério da Saúde (MS). Com o objetivo de promover e proteger a saúde da criança e o aleitamento materno, a política abrange os cuidados com a criança desde a gestação até os 9 anos de idade, com especial atenção à primeira infância e às populações de maior vulnerabilidade (Brasil, 2015).

A PNAISC se estrutura em sete eixos estratégicos, com a finalidade de orientar e qualificar as ações e serviços de saúde da criança no território nacional, considerando os determinantes sociais e condicionantes, visando à efetivação de medidas que permitam o nascimento e o pleno desenvolvimento na infância, com condições dignas, bem como a redução das vulnerabilidades e riscos para o adoecimento e outros agravos, a prevenção das doenças crônicas na vida adulta e da morte prematura de crianças (Brasil, 2015).

Entre os eixos temáticos, este trabalho se insere no eixo III, que trata da promoção e acompanhamento do crescimento e do desenvolvimento integral, e IV acerca da atenção integral a crianças com agravos prevalentes na infância e com doenças crônicas. Trata-se, portanto, de um estudo englobando a puericultura, que em toda a sua história, sempre representou um conjunto de ações voltadas para o cuidado da criança, como a manutenção da saúde, prevenção de doenças e o acompanhamento do desenvolvimento da criança (Albernaz; Couto, 2022).

Nos serviços da APS, os profissionais que realizam o pré-natal frequentemente são os que seguirão acompanhando a família durante a puericultura da criança, seguindo um dos princípios mais importantes da Atenção Básica: a longitudinalidade. Sendo assim, o vínculo entre a equipe de saúde e a família do recém-nascido para o acompanhamento da criança deve preferencialmente se iniciar pelo menos desde o pré-natal. Essa possibilidade de acompanhar famílias ao longo do tempo mantém os profissionais da APS em uma situação privilegiada no reconhecimento de situações que necessitam ser mais bem entendidas e acompanhadas (Brasil, 2012), especialmente no cuidado infantil.

4 MÉTODOS

Neste capítulo, apresenta-se a abordagem metodológica expandida utilizada na condução da pesquisa. A Seção 4.1 se refere ao desenho do estudo. Por se tratar de uma pesquisa cujos dados foram obtidos a partir das coortes BRISA, as Seções 4.2 e 4.3 detalham o local onde o mesmo foi realizado e a amostra considerada. A Seção 4.4 discorre sobre a amostra do estudo, bem como dos critérios de inclusão e exclusão. Na Seção 4.5, são apresentados os instrumentos utilizados na coleta de dados das coortes. A Seção 4.6 versa sobre as variáveis independentes e o desfecho. A Seção 4.7 explica o processamento dos dados, detalhando as métricas e estratégia empregadas, assim como as ferramentas aplicadas para as análises estatísticas subsequentes. Por fim, na Seção 4.8, os aspectos éticos e legais da pesquisa são abordados.

4.1 Desenho do estudo

Trata-se de um estudo prospectivo, vinculado às coortes BRISA “Estudo de Coortes Brasileiras de Pré-Natal e Nascimento de Ribeirão Preto e São Luís” (Silva et al., 2014). Para este recorte, estudamos o fenômeno da não ocorrência de *catch-up* estatural a partir da quantificação de dados preditores dos primeiros mil dias de vida. Por meio da aplicação de modelos preditivos de AM, buscou-se identificar padrões capazes de apontar o desfecho conhecido, bem como avaliar os modelos propostos. Para isso, utilizaram-se os dados das coortes de ambas as cidades brasileiras.

4.2 Local do estudo

A coleta de dados ocorreu em São Luís (SL) e Ribeirão Preto (RP), iniciada em 2010 e seguida em 2011 a 2013. SL é a capital do estado do Maranhão, localizada no nordeste brasileiro. Em 2010, a cidade possuía o Índice de Desenvolvimento Humano Municipal (IDH-M) de 0,768, ocupando a 249ª posição no Brasil e sua população era de 1.014.837 habitantes. Dentre sua população, 35,27% eram vulneráveis à pobreza. A mortalidade infantil era de 16,02 por mil nascidos vivos e o percentual de crianças extremamente pobres de 7,34%. Por outro lado, RP encontra-se em uma região desenvolvida e industrializada. No ano de 2010, o seu IDH-M foi 0,800, considerado muito elevado, e sua população era de 604.682 habitantes. Dentre sua população, 9,46% eram vulneráveis à pobreza. A mortalidade infantil

era de 9,58 por mil nascidos vivos e o percentual de crianças extremamente pobres de 1,22% (IBGE, 2010; UNDP, 2010).

4.3 Amostra do estudo BRISA

Para entender como a amostra do presente trabalho foi composta, vale descrever a amostra dos estudos BRISA. As coortes prospectivas BRISA foram pensadas para desenvolver dois grandes objetivos: investigar os fatores etiológicos da prematuridade e suas consequências no início da vida; e estudar o crescimento, desenvolvimento, amamentação e infecções, acesso, uso e financiamento de serviços de saúde no primeiro ano de vida, bem como fatores de risco precoces para doenças crônicas (Barbieri et al, 2008; Silva et al., 2014).

Dessa forma, em ambas as cidades brasileiras, houve duas coortes: a primeira iniciada no pré-natal e a segunda, no momento do nascimento. Iniciadas em 2010, as coortes de pré-natal foram compostas por uma amostra de conveniência de gestantes recrutadas em serviços de saúde pré-natal públicos e privados, devido à falta de cadastros de gestantes ou de mulheres em acompanhamento pré-natal, que dificultou a obtenção de uma amostra aleatória representativa. Foram convidadas 1500 gestantes (SL) e 1400 (RP), que estivessem entre 22^a e a 25^a semana de IG (Silva et al., 2014). 1447 gestantes, em SL, e 1369, em RP, participaram da primeira coleta, no período de fevereiro de 2010 a junho de 2011. Das gestantes que participaram do início da coorte, 1382 e 1363 mães em SL e PR, respectivamente, foram entrevistadas novamente na ocasião do nascimento.

A segunda coorte, com início no nascimento, foi composta por uma amostra probabilística. Em SL, dez hospitais públicos e privados entre janeiro e dezembro de 2010, incluíram cerca de 97% de todos os nascimentos da capital. Ao todo, foram 21.401 nascimentos, sendo escolhidos um terço por sorteio (7.133). Destes, 5.475 nascimentos foram de mulheres residentes no município por pelo menos três meses e, portanto, elegíveis para o estudo. A amostra foi composta por 5.236 puérperas e, depois da exclusão de 70 natimortos, foi constituída por 5.166 nascimentos. Em RP, foram incluídos todos os nascimentos de residentes no município, totalizando 8103. 309 tiveram alta hospitalar precoce e 42 natimortos, resultando em 7.752 nascidos vivos (únicos e múltiplos).

De março de 2011 a março de 2013, em ambas as cidades, houve nova coleta de dados das coortes iniciadas no pré-natal e no nascimento, quando as crianças tinham entre 13 e 35 meses. Todas as mães e crianças participantes foram convidadas para nova avaliação. No entanto, em razão da impossibilidade de contato, por mudanças de telefone ou endereço,

recusas e não comparecimento, houve perdas. Detalhes das amostras podem ser encontrados em Moraes et al, 2017, Ribeiro et al., 2019, D’êça Junior et al, 2019 e Silva et al, 2019. Para este estudo, considerou-se a junção dos conjuntos de dados nas duas cidades. As bases de dados não estão disponíveis publicamente.

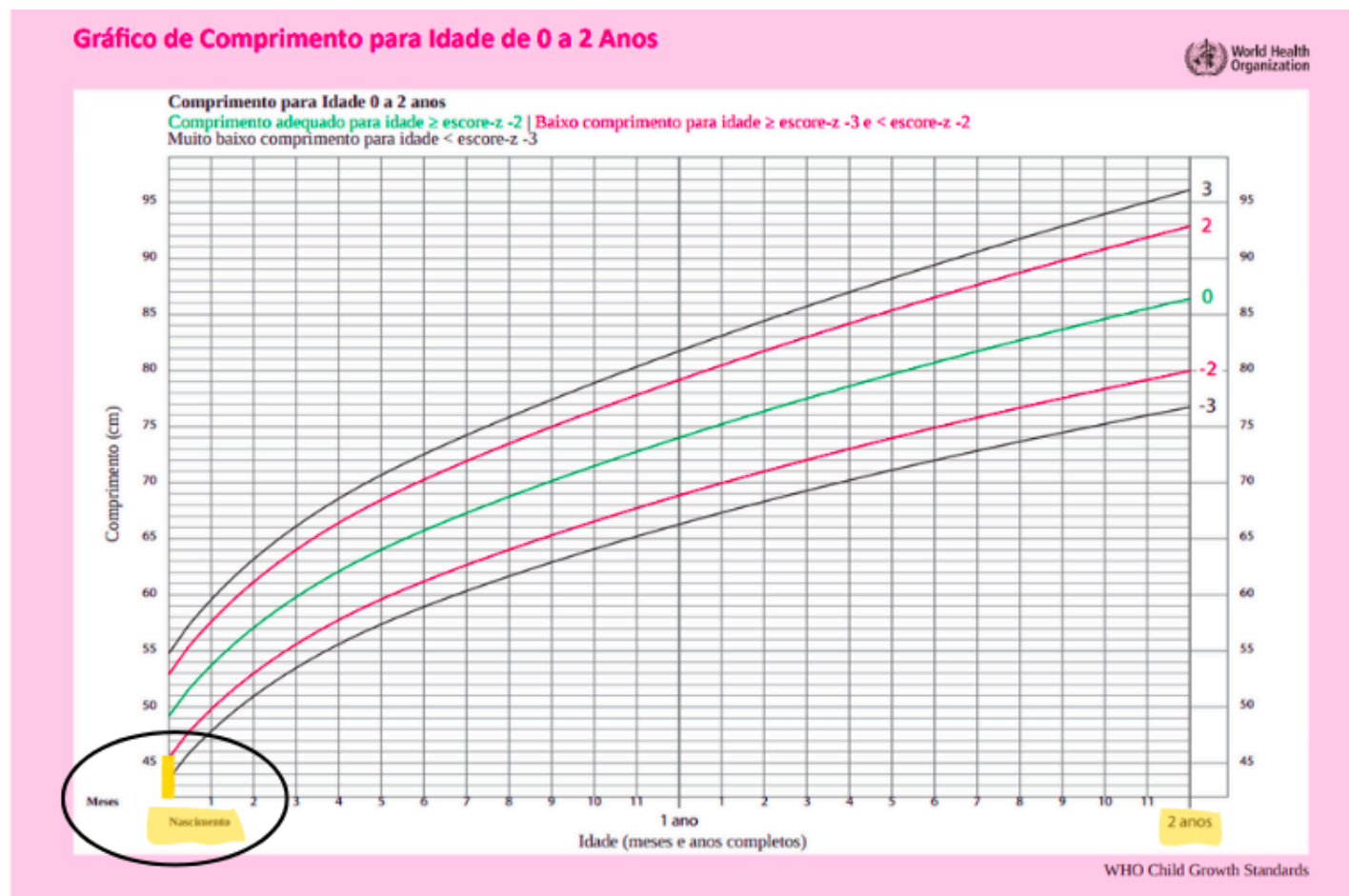
4.4 Amostra deste estudo e critérios de inclusão e exclusão

Para atender ao objetivo deste estudo, foi necessário aplicar critérios de inclusão e exclusão. Primeiramente, foram incluídas as crianças com até 24 meses na fase de seguimento pós-nascimento; e que nasceram com estatura baixa ou muito baixa, de acordo com o gráfico INTERGROWTH-21st, de comprimento para idade de 0 a 2 anos (excluindo aqueles que nasceram com estatura dentro da faixa considerada normal). Dessa forma, as crianças elegíveis apresentaram com escore Z de comprimento <-3 ou ≥ -3 e <-2 , respectivamente (Villar et al, 2014). A Figura 3 ilustra o gráfico usado, presente na caderneta de saúde da criança, distribuída pelo Ministério da Saúde, com a classificação conforme os valores em escore Z de desvio padrão.

Foram excluídos casos de gêmeos, a fim de evitar duplicidade de informações que pudessem comprometer a independência das observações e influenciar os resultados do estudo, bem como pelo fato de já terem sido excluídos nas coortes de pré-natal.

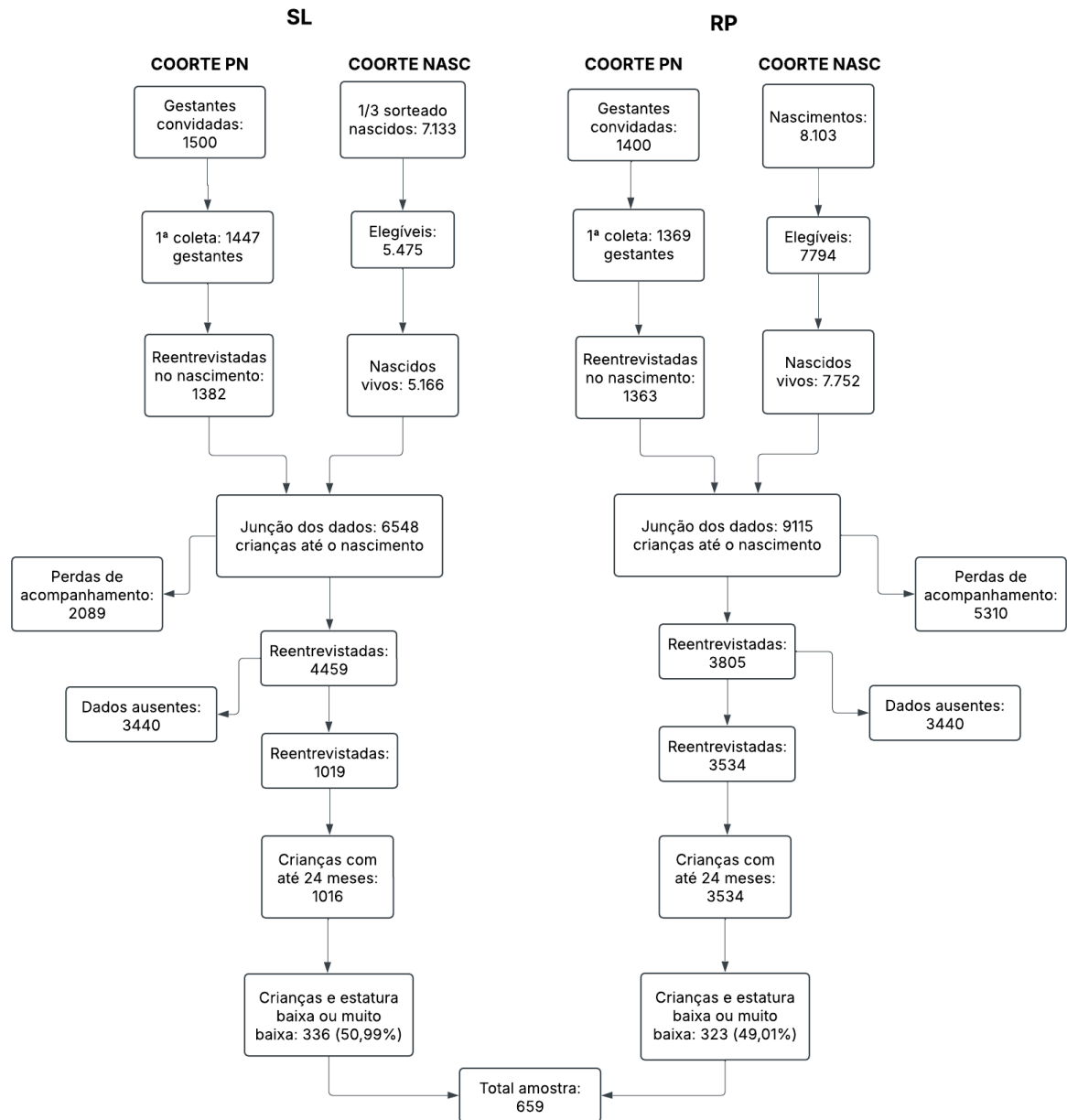
No entanto, houve perdas de acompanhamento de algumas crianças devido à mudança de endereço ou telefone, impossibilitando a coleta das informações necessárias para determinar a ocorrência ou não do *catch-up* estatural até os 24 meses de idade. Além disso, houveram dados faltantes, decorrentes da não resposta a algumas variáveis-chave, como comprimento ao nascer e/ou até 24 meses. Entre perdas de seguimento e dados ausentes, o total da amostra foi de 659 participantes, conforme observado no fluxograma da Figura 4.

Figura 3. Gráfico de Comprimento para Idade de 0 a 2 anos.



Em destaque, o perfil de crianças investigadas neste estudo: nascidas com baixo ou muito baixo comprimento para idade. Fonte: Caderneta da Criança - Menina (BRASIL, 2021).

Figura 4. Fluxograma dos estudos de coortes envolvidos, seus respectivos inícios, fases e quantitativo de participantes.



Legenda: SL=São Luís; RP=Ribeirão Preto; PN=Pré-Natal; Nasc=Nascimento. Fonte: Elaborado pelos autores (2025)

4.5 Instrumentos de coleta de dados na coorte BRISA

Os dados foram obtidos por meio de instrumentos padronizados. Durante o pré-natal, utilizou-se o questionário do pré-natal aplicado por entrevistador (Silva et al., 2014), contendo dados socioeconômicos e demográficos, de hábitos de vida da gestante, características da

gestação atual e da assistência do pré-natal, atividade física pelo questionário internacional de atividade física (IPAQ-versão curta), exame clínico, entre outros.

Na etapa de nascimento, as mães foram entrevistadas até 48 horas pós-parto (Chagas et al., 2017), por meio do questionário do nascimento, sobre aspectos socioeconômicos e demográficos, hábitos de vida, saúde sexual e reprodutiva, morbidades da mãe, características da gestação, do pré-natal, do parto, do recém-nascido e dados do prontuário. Para o seguimento, a coleta ocorreu por meio do questionário de entrevista do primeiro ano, contendo informações sobre hábitos de vida da mãe e exame clínico dos pais, além de alimentação, saúde, gastos com saúde, dados socioeconômicos e exame clínico da criança.

As avaliações antropométricas foram realizadas utilizando-se fita métrica (para circunferências e perímetros), antropômetro portátil e balança digital tipo Monitor de Composição Corporal modelo Tanita®, para aferir o peso (Marques, 2020). Todos os dados foram tabulados e armazenados no programa STATA.

4.6 Variáveis do Estudo

Para a análise deste estudo, as variáveis preditoras utilizadas foram escolhidas manualmente, considerando a plausibilidade teórica do problema e variáveis disponibilizadas no banco. Ao todo, 40 variáveis foram selecionadas e divididas em cinco blocos: preditores/características sociodemográficos; preditores/características físicas, comportamentais e de saúde dos pais; preditores do pré-natal; do nascimento; e 1º ano de vida. Algumas variáveis sofreram transformações, foram categorizadas ou dicotomizadas para melhor análise dos dados. Isso pode ser observado no Quadro 1 - Configuração das variáveis do estudo.

O desfecho deste estudo (não *catch-up* estatural) foi criado a partir da diferença entre o escore Z das crianças até 24 meses e o escore Z das crianças ao nascer. Crianças com uma mudança no escore Z de comprimento para idade $> 0,67$, são definidas como as que apresentaram *catch-up* estatural (Ong et al, 2000). Posteriormente, essa variável foi categorizada em “Sim”, para quem teve o *catch-up* e “Não”, em casos negativos, determinando a variável final de desfecho.

Quadro 1. Configuração das variáveis do estudo.

Bloco	Nome da variável	Tipo de variável	Valores possíveis
Preditores/ características sociodemográficos	Idade da mãe	Categórica Ordinal	20 a 39 anos; < 20 anos e $\geq$ 40 anos
	Idade do pai	Categórica Ordinal	20 a 39 anos; < 20 anos e $\geq$ 40 anos
	Escolaridade Materna	Categórica Ordinal	Até 8 anos ou acima de 8anos
	Ocupação materna	Categórica Nominal	Desempregada, Manual não qualificada, Manual especializada/semi-especializada, Qualificada não manual
	Com companheiro	Categórica Nominal	Sim ou não
	Cor da mãe	Categórica Nominal	Branca, Não Branca
	Classificação econômica ABEP	Categórica Ordinal	A/B, C, D/E
	Núcleo familiar	Categórica Ordinal	Até 4 pessoas, >5
	Bolsa família	Categórica Nominal	Sim ou Não
	Altura da mãe	Numérica Contínua	Qualquer número real
	Altura do pai	Numérica Contínua	Qualquer número real
	Peso do pai	Numérica Contínua	Qualquer número real
	Peso da mãe	Numérica Contínua	Qualquer número real

Preditores/ Características físicas, comportamentais e de saúde dos pais	Peso antes de engravidar	Numérica Contínua	Qualquer número real
	Uso de álcool na gestação	Categórica Nominal	Sim ou Não
	Uso de drogas ilícitas na gestação	Categórica Nominal	Sim ou Não
	Tabagismo materno	Categórica Nominal	Sim ou Não
	IPAQ (atividade física)	Categórica Nominal	Ativa ou Inativa
	Nº de Comorbidades	Categórica Ordinal	0 a 6
	Comorbidades	Categórica Nominal	Sim ou Não
	Diabetes materna	Categórica Nominal	Sim ou Não
	Hipertensão materna	Categórica Nominal	Sim ou Não
Preditores do Pré-Natal	Início consultas de pré-natal no 1º trimestre	Categórica Nominal	Sim ou Não
	Realizou mais de 6 consultas de pré-natal	Categórica Nominal	Sim ou Não
	Nº de ultrassonografias	Categórica Ordinal	Até 2, >3
	Local do pré-natal	Categórica Nominal	Sistema Público ou Privado
	Sexo do recém-nascido	Categórica Nominal	Masculino ou Feminino
	Tipo de parto	Categórica Nominal	Normal ou Cesárea
	Peso para idade ao nascer	Categórica Ordinal	Muito baixo peso para a idade, Baixo peso para a idade, Peso adequado para a idade, Peso elevado para a idade

Preditores do Nascimento	Estatura para idade ao nascer	Categórica Ordinal	Muito baixa estatura para a idade, Baixa estatura para a idade ou Estatura adequada para a idade
	Grau de prematuridade	Categórica Ordinal	< 196 dias; entre 196 e 224 dias; entre 224 e 259 dias; 259 ou mais
	Perímetro cefálico	Numérica Contínua	Qualquer número real
	Apgar 1''	Categórica Nominal	Adequado ou Asfíxia
	Apgar 5''	Categórica Nominal	Adequado ou Asfíxia
	Comprimento ao nascer	Numérica Contínua	Qualquer número real
	Baixo Peso ao Nascer	Categórica Nominal	Sim ou Não
Preditores do 1º ano de vida	Idade da criança	Numérica Contínua	0 a 24
	Mamou exclusivamente no peito até 6 meses	Categórica Nominal	Sim ou Não
	Entubação nos primeiros 3 meses de vida	Categórica Nominal	Sim ou Não
	Numero de filhos	Numérica discreta	Qualquer número natural
	Nasceu pretermo	Categórica Nominal	Sim ou Não
Desfecho	Não <i>catch-up</i> de comprimento	Categórica Nominal	Sim ou Não

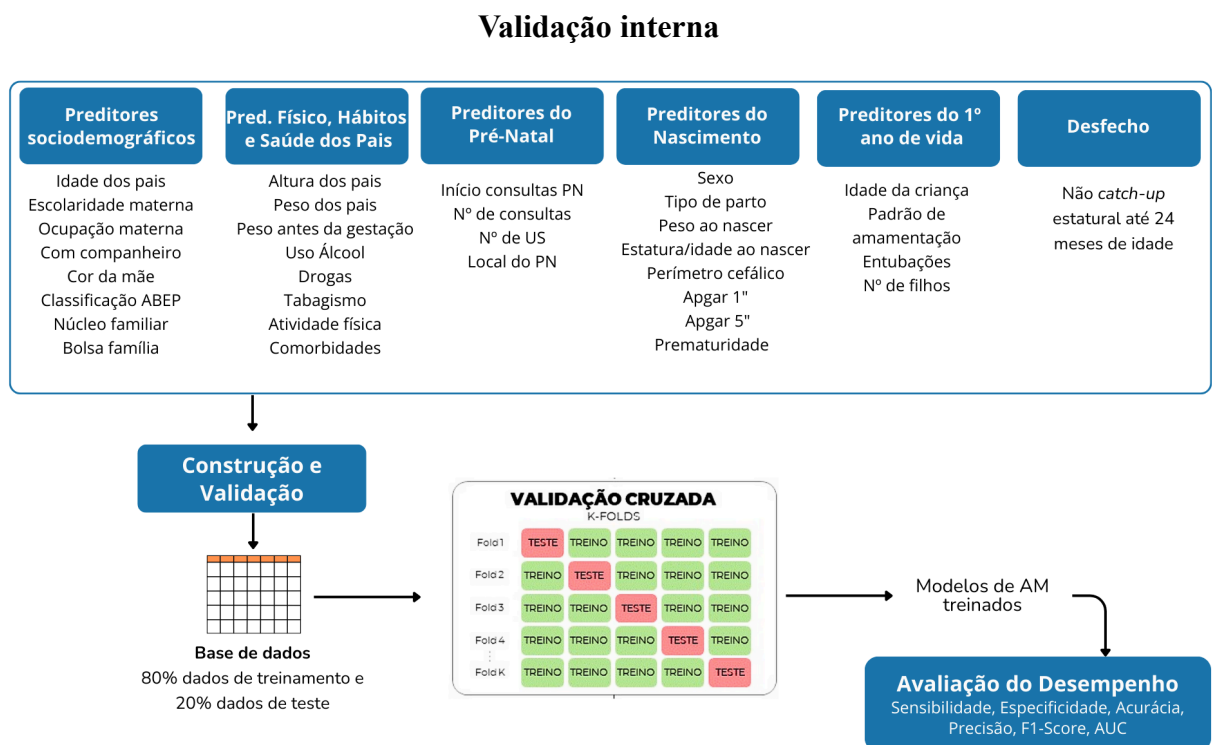
Fonte: Elaborado pelos autores (2025).

4.7 Processamento e Análise dos Dados

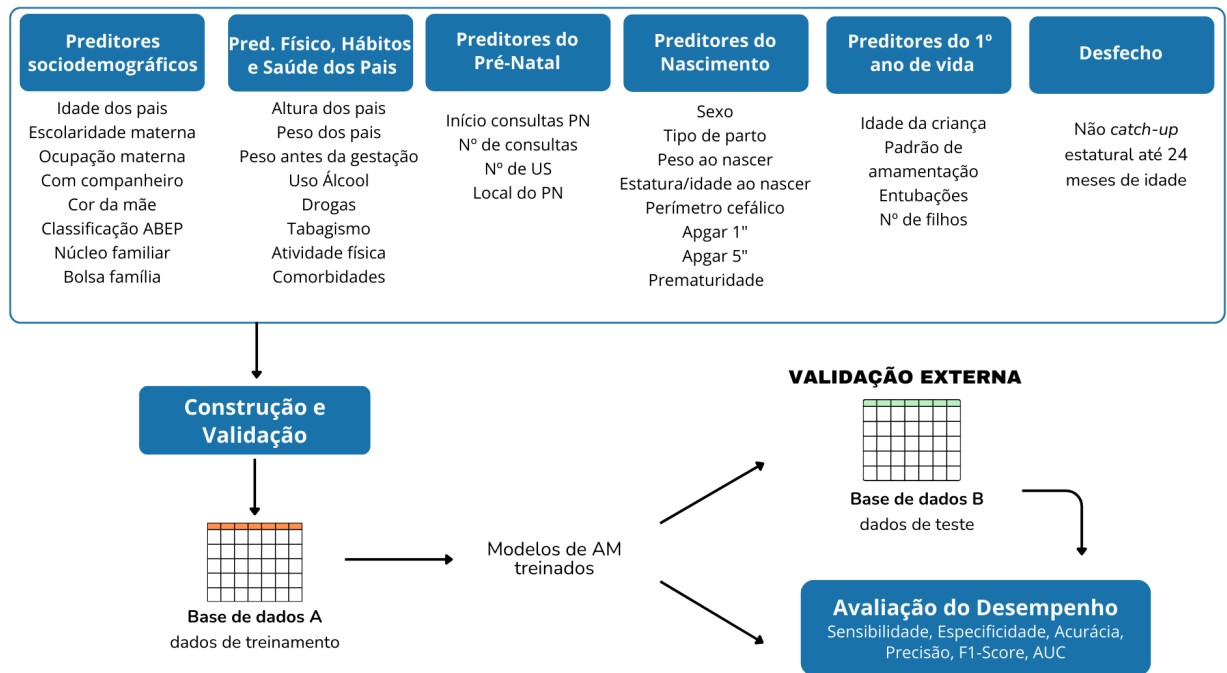
A linguagem de programação R (versão 4.4.2) e o *RStudio* (versão 2024.09.0), um ambiente de desenvolvimento integrado (IDE), foram adotados. Entre as vantagens, estão a disponibilidade gratuita para uso em código aberto, no qual é possível facilmente alterar as análises por linha de código, e assim otimizar o tempo comparado aos softwares via clique; possibilidade de trabalhos colaborativos, que contribuem para a melhoria de códigos; garantia de reprodutibilidade. Além disso, conta com uma comunidade ampla de usuários e diferentes pacotes, úteis para AM (Gibbons; Gibbons, 2019).

Para a condução dos experimentos desta pesquisa, adotou-se o fluxo de trabalho para estudos envolvendo AM apresentado por Lo Vercio, 2020. Ele consiste em cinco etapas principais. Na primeira, o problema é definido. Em seguida, os dados utilizados são processados para posterior modelagem. Então, algoritmos de AM apropriados são selecionados para a inferência. Na sequência, os métodos são de fato aplicados aos dados e os modelos de AM são gerados. Por fim, os modelos executados e os resultados preditivos são avaliados. É possível observar abaixo o modelo teórico construído para este estudo (Figura 5). Na sequência, as etapas computacionais desse fluxo são descritas.

Figura 5. Modelos teóricos do estudo com validações interna e externa, respectivamente



Validação externa



Fonte: Elaborado pelos autores (2025).

4.7.1 Preparação dos Dados

Nesta etapa, também conhecida como pré-processamento de dados, alguns aspectos relevantes que podem interferir no desempenho dos modelos foram considerados no conjunto de treinamento (e aplicados posteriormente no conjunto de teste): tratamento de valores *missing*, transformação dos dados, observação ao vazamento de dados e a eliminação de variáveis muito correlacionadas e/ou irrelevantes.

Foram excluídas variáveis com mais de 95% de resposta a um atributo, uma vez que não seria possível ao algoritmo extrair alguma informação relevante para aprender diante de uma "variância próxima de zero" (Kuhn, 2008). Da mesma forma, foram descartadas aquelas mal preenchidas, considerando mais de 50% de *missing values* (Pfob; Lu; Sidey-Gibbons, 2022).

A transformação de dados, por sua vez, é uma forma de modificar os dados originais, criando variáveis secundárias, pela aplicação de funções matemáticas, produtos, proporções ou combinações mais avançadas (Badillo et al., 2020). Em R, uma variável do tipo *factor* (fator) é usada para representar dados categóricos. Então, dados qualitativos (nominais ou

ordinais) foram transformados para permitir a realização das manipulações e análises estatísticas.

No entanto, como alguns algoritmos lidam melhor com dados numéricos, também empregou-se a técnica *One-hot encoding*, que converte variáveis categóricas em numérica. Nela, cada categoria é representada por um vetor binário, no qual apenas um dos valores é 1 (“hot”, indicando a presença da categoria) e os outros são “0” (“cold”, ausência da categoria). Os novos atributos às vezes são chamados de atributos fictícios. A vantagem desse artifício é sua facilidade de implementação, além de evitar que o modelo interprete uma relação de ordem que não existe entre as categorias. Entretanto, isso pode aumentar expressivamente o número de colunas conforme as categorias, levando a um maior uso de memória e tempo de processamento (Géron, 2019).

Por sua vez, a normalização ou padronização de variáveis, também chamada de padronização Z-score, é outra ferramenta que foi utilizada. Transformam-se os dados para uma distribuição com média 0 e desvio padrão 1, o que torna as variáveis comparáveis e facilita a operacionalização de determinados algoritmos. O método de vizinhos mais próximos (k Nearest Neighbors - kNN) foi utilizado para a imputação de dados. Ele opera identificando observações semelhantes àquela com valores *missing* e utilizando a média dos valores disponíveis de uma determinada variável para realizar o preenchimento (Batista; Chiavegatto Filho, 2019).

Como em muitos experimentos na área da Saúde, nem todas as variáveis coletadas são relevantes ao desfecho considerado. Isso decorre, em parte, pela amplitude das pesquisas, que podem abranger múltiplos desfechos ou mesmo desfechos não previstos no momento da coleta. Assim, variáveis pouco relevantes ou redundantes podem surgir. Embora muitos algoritmos de AM possuam mecanismos implícitos de ponderação/seleção de variáveis, pode ser útil recorrer a um método externo. Aqui, optou-se pelo Boruta, dado seu bom desempenho em diversos domínios (Kursa, 2022). Ele foca na identificação de variáveis relevantes de acordo com o critério de importância do algoritmo de classificação *Random Forest* (Mello; Ponti, 2018). Ele opera comparando as variáveis com suas versões aleatórias e progressivamente eliminando aquelas com poder preditivo nulo, conservando as demais.

Os dados foram analisados em 5 cenários diferentes:

1. Junção da base de dados de pré-natal e nascimento de SL e RP;
2. Base de dados SL isoladamente;
3. Base de dados RP isoladamente;

4. Base de dados SL como treino e RP como teste;
5. Base de dados RP como treino e SL como teste.

Nos três primeiros cenários, houve a validação interna, na qual 80% dos dados foram subdivididos em treino e 20% em teste. Nos dois últimos, a validação externa possibilitou treinar os algoritmos com uma base e testar os modelos gerados com a outra base.

4.7.2 Treinamento dos Algoritmos

Com os dados pré-processados, pode-se realizar o treinamento dos modelos de predição. Tal processo refere-se à aplicação de algoritmos de AM aos dados, a fim de produzir um mapeamento entre as variáveis preditoras e os desfechos. Com isso, amostras não vistas durante esse processo podem ter seus desfechos estimados. Nesta tarefa, são empregadas aqui as abordagens LDA, RLR, SVM Polinomial e XGBoost, discutidas na Seção 3.6. Eles representam famílias de algoritmos com diferentes vieses indutivos e contribuem para uma análise mais abrangente do problema sob estudo. À exceção do LDA, os métodos empregados possuem hiperparâmetros específicos que devem ter seus valores definidos para obter os melhores resultados de predição. O Quadro 1 apresenta os hiperparâmetros considerados e uma breve descrição de cada um. Seus valores foram definidos utilizando a abordagem *grid search*, que avalia múltiplas combinações de valores em busca daquela com melhor desempenho (Faceli et al., 2021).

Quadro 2. Hiperparâmetros dos algoritmos usados e suas respectivas descrições.

Algoritmo	Hiperparâmetro	Descrição
RLR	cost	Custo de erros de classificação.
	loss	Tipo de função de perda utilizada.
	epsilon	Tolerância para convergência do algoritmo.
SVM	degree	Grau do polinômio do <i>kernel</i> Polinomial.
	scale	Fator de escala aplicada às variáveis.
	C	Custo de erros de classificação.
XGBoost	nrounds	Número de iterações no <i>boosting</i> .
	max_depth	Profundidade máxima das árvores.
	eta	Taxa de aprendizado.
	gamma	Penalidade por ramos extras nas árvores.
	colsample_bytree	Percentual de variáveis usadas em cada árvore.

	<code>min_child_weight</code>	Número mínimo de observação para dividir um nó.
	<code>subsample</code>	Percentual de amostras usadas em cada iteração.

Fonte: Documentação do pacote caret do R.

A implementação de todos os algoritmos e análises realizadas neste trabalho foi feita utilizando o pacote caret, pROC, Boruta, boot, dplyr e ggplot2 do R.

4.7.3 Teste dos Algoritmos e avaliação de desempenho

Uma vez selecionados os algoritmos de AM, deve-se testar os modelos preditivos gerados e avaliar os resultados obtidos. Para tanto, é preciso definir que estratégia de avaliação utilizar, que métricas de desempenho considerar e que ferramentas estatísticas aplicar para interpretar os resultados. Esses aspectos serão abordados a seguir.

Para fazer melhor uso dos dados disponíveis, considerou-se neste estudo um particionamento iterativo do conjunto de dados, conhecido como validação cruzada com K partições (do inglês *K-fold cross validation*). Nesta abordagem, os dados são divididos em K partições de tamanhos aproximadamente iguais, em que K-1 partições são utilizadas para treinar o modelo preditivo, enquanto a partição remanescente é utilizada para a testar o modelo e estimativa de seu desempenho. O processo se repete até que todas as partições sejam testadas uma vez, resultando assim em K estimativas de desempenho (Batista; Chiavegatto Filho, 2019). Neste trabalho, utilizou-se K=5, como recomendado por Tohka; Van Gils, 2021. Ademais, para mitigar os efeitos aleatórios da validação cruzada e tornar o processo de estimação de desempenho mais robusto, repetiu-se o processo 10 vezes.

Para mensurar o desempenho dos modelos em cada partição de testes, é necessário definir métricas de avaliação apropriadas. Como tem-se um desfecho binário (ocorrência ou não de *catch-up* estatural), indicadores próprios para o contexto de classificação foram empregados (Gibbons; Gibbons, 2019). Aqui, utilizaram-se as seguintes métricas, observáveis na Figura 6: sensibilidade (*recall*), especificidade (*specificity*), acurácia (*accuracy*), precisão (*precision*). Sensibilidade é a proporção de verdadeiros positivos corretamente identificados pelo teste; especificidade é a proporção de verdadeiros negativos corretamente identificados pelo teste; precisão é a proporção de vezes que o classificador está correto (Batista; Chiavegatto Filho, 2019).

Figura 6. Matriz de confusão para problemas de duas classes

		Predicted labels		
		1	0	
Actual labels (observations)	1	True Positive (TP)	False Negative (FN)	Recall=TPR (True Positive Rate) $TPR = \frac{TP}{TP+FN}$
	0	False Positive (FP)	True Negative (TN)	Specificity = $\frac{TN}{TN+FP}$ False Positive Rate: $FPR = \frac{FP}{FP+TN}$
		Precision $\frac{TP}{TP+FP}$	False Negative Rate $\frac{FN}{TN+FN}$	Accuracy $\frac{TP+TN}{TP+TN+FP+FN}$

Essa estrutura indica o quão bem-sucedido o algoritmo foi na previsão de rótulos em um problema e classificação binária, no qual os rótulos assumem valores 0 (“negativo”) ou 1 (“positivo”) avaliando o previsto *versus* o real. Fonte: Badillo et al. (2020)

Além dessas, há o *F1-score*, combinação entre precisão e sensibilidade de maneira equilibrada, e a AUC (*Area Under Curve*), que explica a probabilidade de uma amostra aleatória ser classificada corretamente por cada algoritmo e fornece um valor único que representa o desempenho esperado (Gibbons; Gibbons, 2019). Como a AUC é uma porção da área do quadrado unitário, seu valor estará sempre entre 0 e 1; quanto mais próximo de 1 for o valor da AUC, melhor é o desempenho do algoritmo (Fawcett, 2006).

A fim de analisar com mais propriedade os resultados preditivos dos modelos de AM, devem-se incluir testes estatísticos apropriados no estudo. Isso é importante para lidar com flutuações aleatórias dos valores obtidos, decorrentes da limitação natural de dados, da própria estratégia de avaliação empregada e da natureza estocástica de algumas abordagens de aprendizagem. Em um trabalho abrangente, Ranio et al, 2024 compilou as métricas de avaliação mais comumente utilizadas para diversas tarefas de AM e delineou uma série de recomendação de quais testes utilizar em que situações. Com base neste trabalho e em referências ali contidas, optou-se por aplicar aqui o teste não paramétrico de Friedman para avaliar diferenças globais entre os métodos. Quando esse teste identificar diferenças significativas, aplica-se então o teste de Wilcoxon para comparações par a par, com correção de Bonferroni para ajustar o nível de significância em função do número de comparações realizadas. O teste de McNemar também foi utilizado para comparar modelos com base em um único conjunto de testes. Neste trabalho, de maneira geral, foi adotado um nível de significância de 95%.

4.8 Aspectos Éticos e Legais

Em SL, o projeto fornecedor dos dados foi aprovado pelo Comitê de Ética em Pesquisa do Hospital Universitário da Universidade Federal do Maranhão (nº do processo 4771/2008-30). Em RP, no Comitê de Ética em Pesquisa do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo (FMRP/USP), com o número de parecer 4116/2008. A concordância em participar do estudo ocorreu mediante a assinatura do Termo de Consentimento Livre e Esclarecido (TCLE) dos responsáveis pelas crianças.

5 RESULTADOS

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

5.1 Artigo

***Catch-up* Estatural em Crianças nos Primeiros Mil Dias de Vida: um Estudo de Predição usando Aprendizagem de Máquina**

(a ser submetido à Revista Preventive Medicine Reports. Fator de impacto 2.4, Qualis A3)

***Catch-up* Estatural em Crianças nos Primeiros Mil Dias de Vida: um estudo de Predição usando Aprendizagem de Máquina**

Height Catch-up growth in Children in the First Thousand Days of Life: a Prediction Study using Machine Learning

Cibelly de Fátima Vieira Ferreira

Programa de Pós Graduação em Saúde Coletiva, Universidade Federal do Maranhão, São Luís, Maranhão, Brasil.

Erika Bárbara Abreu Fonseca Thomaz

Departamento de Saúde Pública), Universidade Federal do Maranhão, São Luís, Maranhão, Brasil

Bruno Feres de Souza

Departamento de Engenharia da Computação, Universidade Federal do Maranhão, São Luís, Maranhão, Brasil.

Correspondência: cibelly.estudos@gmail.com

Resumo

Objetivo: Predizer a não ocorrência de *catch-up* estatural até 24 meses, de crianças que nasceram com comprimento abaixo do esperado, por meio de Aprendizagem de Máquina.

Métodos: Utilizaram-se as bases de dados das coortes de pré-natal e nascimento BRISA, nas cidades brasileiras Ribeirão Preto (RP) e São Luís (SL). 659 crianças classificadas em baixa ou muito baixa estatura para idade ao nascer, segundo gráfico INTERGROWTH-21st, não gemelares, foram selecionadas e avaliadas quanto ao *catch-up* estatural (mudança entre o escore Z de estatura até os 2 anos de idade e o escore Z de comprimento ao nascer $>0,67$) e potenciais preditores (variáveis socioeconômicas, hábitos, características dos pais, entre outros). Os algoritmos *Linear Discriminant Analysis* (LDA), *Regularized Logistic Regression* (RLR), *Support Vector Machine* (SVM) e *Extreme Gradient Boosting* (XGBoost) foram empregados para construir modelos de predição, em 5 diferentes cenários, no programa R studio. **Resultados:** Neste estudo, a especificidade foi uma importante medida de desempenho. Utilizando a base RP e SL conjugada com validação cruzada, os modelos em LDA e RLR com Boruta tiveram os melhores desempenhos em sensibilidade (72,8% e 72,7%), porém sem ganhos expressivos para acurácia $\approx$ 68 e 69%, especificidade $\approx$ 64% e 65%, precisão $\approx$ 68%, F1 score $\approx$ 70% e AUC $\approx$ 75%, cada. Por outro lado, o cenário da base de SL, utilizando o LDA com o Boruta teve o melhor desempenho: de modo geral: acurácia de 70,9%, sensibilidade de 75%, precisão de 70,5%, F1 Score de 72,5% e AUC de 74,4%, porém

com especificidade de 66,5%. A maior especificidade foi em RP, como treino, e SL, como teste (variando de 78,3% a 87,3%), no qual o LDA sem Boruta teve o maior valor obtido. Todos os modelos testados foram estatisticamente superiores à classe majoritária (até 51,5%). As variáveis que tiveram maior contribuição aos modelos foram: peso ao nascer, grau de prematuridade, altura da mãe, estatura para idade ao nascer, peso para idade ao nascer, fumo. O grau de prematuridade apareceu em todos os modelos. **Conclusão:** Variáveis de fácil acesso relativas aos primeiros mil dias de vida foram capazes de predizer, com acurácia, o não *catch-up* estatural até 2 anos idade de nascidos com déficit de comprimento.

Palavras-chave: Saúde da Criança. Antropometria. Aprendizagem de Máquina Supervisionada. Aprendizagem por Probabilidade. Crescimento Corporal.

1 Introdução

Uma das formas mais simples e comuns de acompanhar o crescimento infantil é o registro antropométrico periódico, em gráficos internacionalmente aceitos, que permitem uma fácil identificação de desvios do crescimento (Villar, 2014; Brasil, 2021). No Brasil, essa avaliação contínua é feita preferencialmente por profissionais da Atenção Primária à Saúde (APS) via Caderneta da Criança, formulada pelo Ministério da Saúde (Palombo et al., 2014; Brasil, 2021).

A estatura é um indicador mais estável ao longo do tempo do que outros parâmetros, como peso ou IMC. Quando registrado por ocasião do nascimento, o comprimento reflete o desenvolvimento intrauterino (Brasil, 2012) e, ao decorrer dos anos de acompanhamento, pode indicar anormalidades e problemas graves, como a desnutrição crônica (Leroy; Frontillo, 2020; Soliman et al., 2021; Roro; Deressa; Lindtjorn, 2022). Déficits no crescimento em estatura estão associados a uma perpetuação de menores índices de altura, aumento da morbimortalidade, impactos neurocognitivos, motores, de autoestima, diminuição dos rendimentos na vida adulta, dentre outros efeitos negativos (Martinez et al., 2020; Soliman et al., 2021; Nguyen et al., 2021; Sartika et al., 2021).

Em indivíduos que nasceram pequenos para idade, seja devido ao baixo peso em geral ou, especificamente, devido à prematuridade, ou crescimento intrauterino restrito (CIUR), é possível observar o fenômeno do *catch-up* como uma propriedade do crescimento humano pelo qual a criança retorna à sua trajetória genética depois de um período de atraso no

crescimento (Ong et al., 2000). Observa-se, então, um crescimento acelerado, a fim de recuperar um déficit prévio (Silveira et al, 2024), permitindo que esses indivíduos consigam nivelar o seu crescimento aos padrões esperados (Rodrigues et al., 2018).

No entanto, crianças cujas curvas de crescimento não se aproximam dos percentis mínimos da normalidade e/ou apresentam padrão descendente, requerem investigação (Vila; Macedo; Melo, 2017). Então, prever a ocorrência dessa “recuperação espontânea” se torna importante, uma vez que o crescimento compensatório nos primeiros anos de vida está fortemente associado ao estado de saúde na infância e desenvolvimento cognitivo (Freitas et al., 2012). Dentro do período dos primeiros mil dias de vida, as intervenções se tornam ainda mais estratégicas por serem muito precoces, podendo causar impacto positivo ao longo da vida dessas crianças.

Para auxiliar nessa predição, a Aprendizagem de Máquina (AM) surge como uma opção viável para a análise de dados, ajudando a encontrar relações não triviais entre variáveis preditoras e o desfecho de interesse (Mello; Ponti, 2018). De fato, alguns trabalhos utilizando AM para prever desfechos antropométricos nos primeiros anos de vida foram desenvolvidos recentemente (Sun et al., 2016; Kuhle et al., 2018; Li et al., 2020; Harrison et al., 2020; Lu et al., 2020; Saw et al., 2021; Tao et al., 2021; Adeeba et al., 2022; Bai et al., 2022; Khan et al., 2022; Islam Pollob et al., 2022; Bekele, 2022; Bitew, Sparks, Nyarko, 2022; Arayeshgari et al., 2023). Entretanto, de modo geral, percebe-se que tais estudos focaram em predições envolvendo o peso da criança como medida primária, em detrimento de outras grandezas antropométricas potencialmente úteis para avaliar o crescimento infantil, como a estatura e mais especificamente, o *catch-up* estatural.

Portanto, considerando a relevância desse tema, este estudo visa realizar uma análise preditiva do crescimento de crianças, que nasceram com comprimento abaixo do esperado, em duas coortes brasileiras (*Brazilian Ribeirão Preto and São Luís Cohort Studies - BRISA*). Utilizando dados do pré-natal, nascimento e pós-nascimento, buscou-se desenvolver e validar modelos de AM para prever a não ocorrência do *catch-up* estatural, contribuindo para a identificação precoce de crianças em risco e auxiliando na implementação de estratégias de intervenção.

2 Material e métodos

2.1. Desenho do estudo

Trata-se de um estudo prospectivo, vinculado às coortes BRISA “Estudo de Coortes Brasileiras de Pré-Natal e Nascimento de Ribeirão Preto e São Luís” (Silva et al., 2014). Para este recorte, o fenômeno da não ocorrência do *catch-up* estatural foi estudado a partir da quantificação de dados preditores dos primeiros mil dias de vida. Por meio da aplicação de modelos preditivos de AM, buscou-se identificar padrões capazes de apontar o desfecho conhecido, bem como avaliar os modelos propostos. Para isso, utilizaram-se os dados das coortes de ambas as cidades brasileiras., inicialmente em 2010 e seguidas entre 2011 a 2013 (Silva et al, 2014).

2.2 Aspectos éticos

Os estudos foram aprovados pelos Comitês de Ética em Pesquisa do Hospital Universitário da Universidade Federal do Maranhão (nº 4771/2008-30), em São Luís (SL), e do Hospital das Clínicas da Faculdade de Medicina de Ribeirão Preto, da Universidade de São Paulo (nº 4116/2008), em Ribeirão Preto (RP). Todos os responsáveis pelas crianças participantes assinaram o Termo de Consentimento Livre e Esclarecido (TCLE).

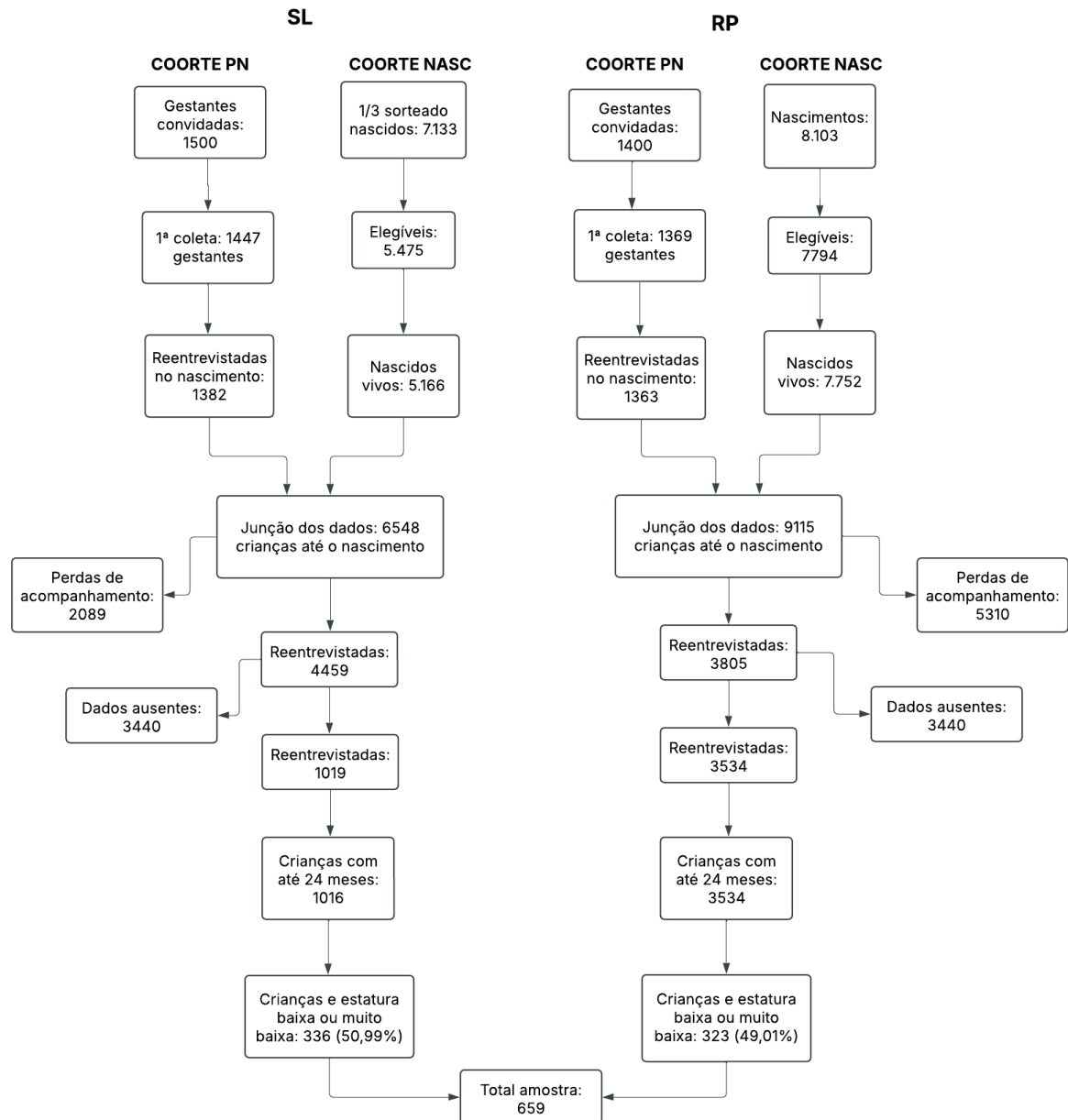
2.3 Amostra

Em ambas as cidades brasileiras, houve duas coortes: a primeira iniciada no pré-natal e a segunda, no momento do nascimento. Os locais de recrutamento foram serviços de saúde públicos e privados. Detalhes das amostras podem ser encontrados em Morais et al, 2017, Ribeiro et al., 2019, D’êça Junior et al, 2019 e Silva et al, 2019. As bases de dados não estão disponíveis publicamente.

Foram incluídas crianças de gestação única, com até 24 meses na fase de seguimento pós-nascimento, e que apresentaram estatura baixa ou muito baixa (escore Z de comprimento ao nascer <-3 ou ≥ -3 e <-2), conforme o gráfico INTERGROWTH-21st, de comprimento para idade de 0 a 2 anos (Villar et al, 2014).

Entre perdas de seguimento e dados faltantes, decorrentes da não resposta a algumas variáveis-chave (comprimento ao nascer e/ou até 24 meses), o total da amostra foi de 659 participantes, conforme pode ser observado no fluxograma a seguir.

Figura 4. Fluxograma dos estudos de coortes envolvidos, seus respectivos inícios, fases e quantitativo de participantes.



Legenda: SL=São Luís; RP=Ribeirão Preto; PN=Pré-Natal; Nasc=Nascimento. Fonte: Elaborado pelos autores (2025)

2.4 Variáveis do estudo

Foram selecionadas manualmente 40 preditores para a análise, conforme a plausibilidade teórica. Incluíram-se variáveis sobre condição socioeconômica dos pais, pré-natal e hábitos maternos, entre outros. Foram agrupadas nos 5 blocos seguintes:

- **Características sociodemográficas:**

Idade materna (<15anos, 15-35 anos, >35 anos, para gravidez de risco (Brasil, 2022)); Idade paterna (<15anos, 15-35 anos, >35 anos), padronizando conforme idade materna; Escolaridade materna, em anos de estudo (até 8anos, >8anos); Ocupação materna (Desempregada, Manual não qualificada, Manual especializada/semi-especializada, Qualificada não manual); Com companheiro (sim, não); Cor da mãe (Branca, Não Branca); Classe econômica - ABEP (A/B, C, D/E); Núcleo familiar (até 4, 5 ou mais pessoas); Beneficiário do bolsa família (sim, não).

- **Características físicas, comportamentais e de saúde dos pais:**

Altura do pai (numérico); Altura da mãe (numérico); Peso da mãe (numérico); Peso do pai (numérico); Peso da mãe antes de engravidar (numérico); Etilismo materno (sim, não); Uso materno de drogas ilícitas (sim, não); Fumo materno (sim, não); Atividade física materna - IPAQ (ativa, inativa); N° de morbidades maternas (0 a 6); Diabetes materna (sim, não); Hipertensão materna (sim, não);

- **Pré-Natal (PN):**

Início das consultas do PN no 1º trimestre (Sim, Não); realizou mais de 6 consultas PN (Sim, Não); N° de ultrassonografias (até 2, >3)

- **Nascimento:**

Sexo do recém-nascido (Masculino ou Feminino); Tipo de parto (Normal ou Cesárea);_Peso ao Nascer OU para idade (numérico); Estatura para idade ao nascer (numérico); Perímetro cefálico (numérico); Apgar 1º minuto (adequado, asfixia); Apgar 5º minuto (adequado, asfixia);

- **1º ano de vida da criança:**

Idade da criança (0 a 24); Amamentação exclusivamente no peito até 6 meses (Sim ou Não); Entubação nos primeiros 3 meses de vida (Sim ou Não).

O desfecho deste estudo (não *catch-up* estatural) foi criado a partir da diferença entre o escore Z das crianças até 24 meses e o escore Z das crianças ao nascer. Crianças com uma mudança no escore Z de comprimento para idade $> 0,67$, são definidas como as que apresentaram *catch-up* estatural (Ong et al, 2000). Posteriormente, essa variável foi

categorizada em “Sim”, para quem teve o *catch-up* e “Não”, em casos negativos, determinando a variável final de desfecho.

2.5. Processamento e Análise dos Dados

Alguns aspectos relevantes que podem interferir no desempenho dos modelos foram considerados no conjunto de treinamento (e aplicados posteriormente no conjunto de teste): tratamento de valores *missing*, transformação dos dados, observação ao vazamento de dados e a eliminação de variáveis muito correlacionadas e/ou irrelevantes.

Foram excluídas variáveis com mais de 95% de resposta a um atributo, "variância próxima de zero" (Kuhn, 2008). Da mesma forma, foram descartadas aquelas mal preenchidas (com mais de 50% de *missing values*) (Pfob; Lu; Sidey-Gibbons, 2022). Variáveis do tipo *factor* (dados categóricos: nominais ou ordinais) foram transformadas para permitir a realização das manipulações e análises estatísticas.

Empregou-se a técnica *One-hot encoding*, que converte variáveis categóricas em numérica. Nela, cada categoria é representada por um vetor binário, no qual apenas um dos valores é 1 (“*hot*”, indicando a presença da categoria) e os outros são “0” (“*cold*”, ausência da categoria) (Géron, 2019).

Transformam-se os dados para uma distribuição com média 0 e desvio padrão 1, o que torna as variáveis comparáveis e facilita a operacionalização de determinados algoritmos (normalização ou padronização de variáveis). O k-NN foi usado para imputação de dados. (Batista; Chiavegatto Filho, 2019).

O Boruta, um algoritmo de seleção de variáveis, foi usado para identificar quais variáveis são mais importantes para um modelo de Aprendizagem de Máquina. Esse método é baseado em *Random Forest* e busca selecionar apenas as variáveis que contribuem significativamente para a precisão do modelo, ignorando variáveis irrelevantes ou redundantes (Kursa, 2022).

Os dados foram analisados em 5 cenários diferentes:

6. Junção da base de dados de pré-natal e nascimento de SL e RP;
7. Base de dados SL isoladamente;
8. Base de dados RP isoladamente;
9. Base de dados SL como treino e RP como teste;
10. Base de dados RP como treino e SL como teste.

Nos três primeiros cenários, houve a validação interna, na qual 80% dos dados foram subdivididos em treino e 20% em teste. Nos dois últimos, a validação externa possibilitou treinar os algoritmos com uma base e testar os modelos gerados com a outra base.

2.6. Treinamento dos algoritmos

Para prever o desfecho, foram empregados os seguintes algoritmos: *Linear Discriminant Analysis* (LDA), *Regularized Logistic Regression* (RLR), *Support Vector Machine* (SVM), *Extreme Gradient Boosting* (XGboost). Devido ao desequilíbrio na proporção de amostras foi empregada a técnica de sobreamostragem de minoria sintética para evitar sobreamostragem.

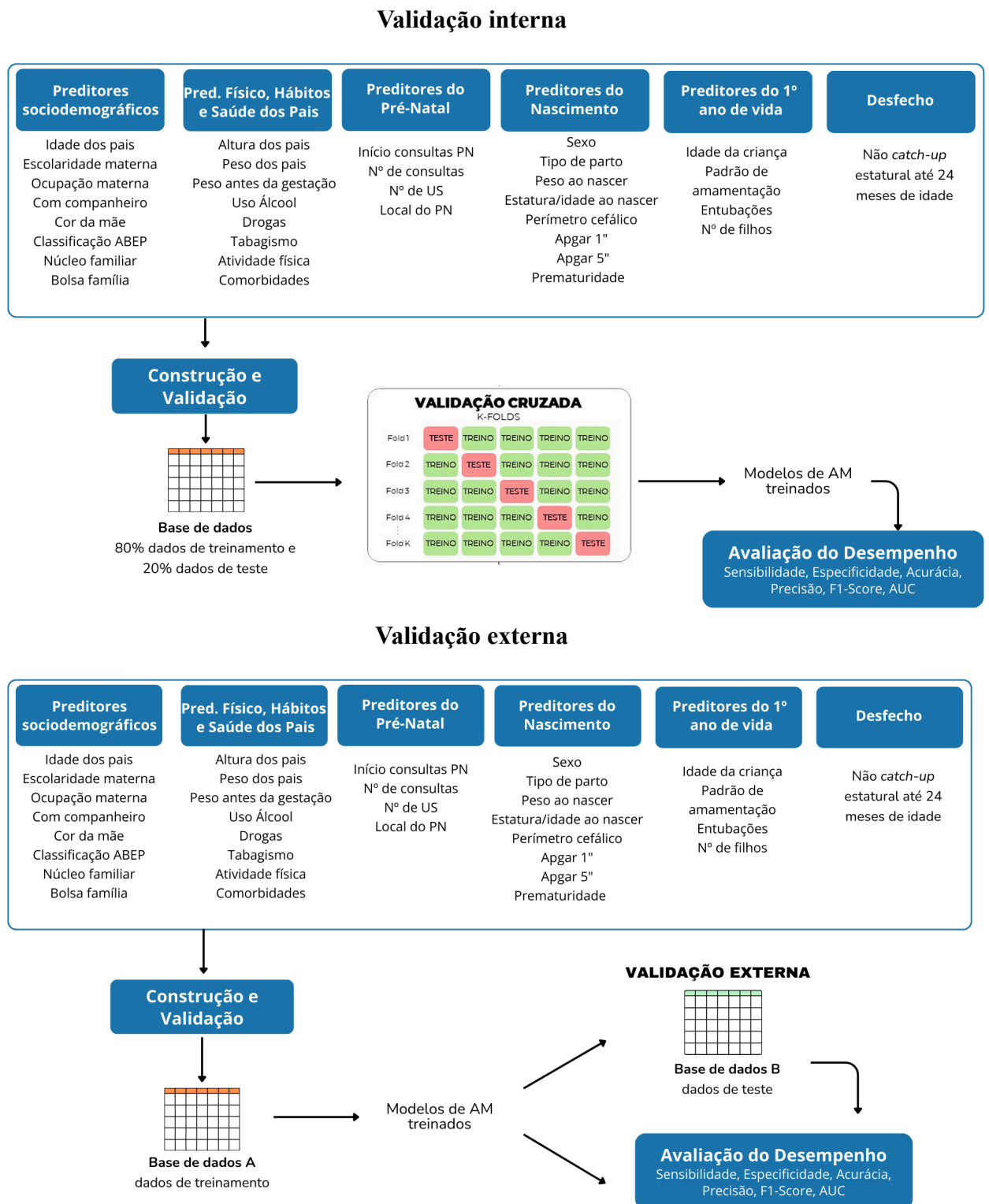
Para fazer melhor uso dos dados disponíveis, considerou-se neste estudo um particionamento iterativo da amostra, conhecido como validação cruzada com K partições (do inglês *k-fold cross validation*): dados divididos em K partes de tamanhos aproximadamente iguais, em que K-1 partes foram utilizadas para treinar o modelo preditivo, enquanto a parte remanescente foi utilizada para a validação do modelo e estimativa de seu desempenho. O processo se repetiu até que todas as partes participaram tanto do treinamento quanto da validação do modelo, resultando assim em K estimativas de desempenho (Batista; Chiavegatto Filho, 2019). Esse processo também foi utilizado para selecionar os hiperparâmetros dos algoritmos.

2.7 Teste dos Algoritmos e Avaliação de desempenho

Uma vez treinados, os modelos de AM tiveram seus desempenhos estimados em dados não vistos previamente. Na validação cruzada, tais dados corresponderam às partições de teste. O desempenho dos modelos é medido comparando-se as predições feitas com os desfechos reais já conhecidos (Gibbons; Gibbons, 2019). Como se tratou de um desfecho binário, os modelos de predição foram do tipo classificação.

Utilizaram-se as seguintes métricas para os cálculos de desempenho: sensibilidade, especificidade, acurácia, precisão, *F1-score* e AUC. O modelo teórico que sintetiza a proposta do trabalho, pode ser observado na Figura 7.

Figura 7. Modelos teóricos do estudo com validações interna e externa, respectivamente



Fonte: Elaborado pelos autores (2025).

3 Resultados

Nesta seção, são reportados os resultados da aplicação de algoritmos de classificação para a predição da ocorrência ou não de *catch-up* estatural adequado, com base em fatores sociodemográficos, pré-natais, perinatais e pós-natais. Os experimentos e análises correspondentes foram conduzidos em cinco cenários. Nos três iniciais, os experimentos preditivos foram executados em três bases de dados separadas: na primeira, dados das coortes de São Luís e Ribeirão Preto foram conjugados para formar uma base de dados mais ampla; na segunda, dados apenas da coorte de São Luís foram considerados e; na terceira, tem-se dados apenas da coorte de Ribeirão Preto. Nos dois últimos cenários, as bases de dados das duas coortes foram utilizados como treino e testes de forma alternada.

Em todos os cenários, foram analisados os comportamentos preditivos dos quatro algoritmos de classificação previamente discutidos: LDA, RLR, SVM e XGBoost. As métricas de desempenho utilizadas foram acurácia, sensibilidade, especificidade, precisão, F1 Score e AUC. Nos cenários com bases de dados separadas, a estratégia de avaliação empregada foi a validação cruzada com K=5 e 10 repetições, e os testes estatísticos de Friedman e Wilcoxon foram realizados. Nos cenários com bases de dados de treino e teste, os modelos de AM foram construídos usando a primeira base e avaliados na segunda, e o teste de McNemar foi executado. O método Boruta foi aplicado para selecionar as variáveis mais relevantes para o desfecho. A seguir, os resultados são apresentados, depois de uma descrição panorâmica dos dados.

Na Tabela 1, observa-se a distribuição dos dados das coortes de São Luís e Ribeirão Preto conjugadas, conforme as condições sociodemográficas, características físicas, comportamentais e de saúde dos pais, bem como dados do nascimento e do primeiro ano de vida das crianças da amostra, conforme discutido na Seção 2.4. Das 659 baixas ou muito baixas da amostra estudada em ambas as cidades, 332 não tiveram *catch-up* estatural (classe de referência) e 327 experimentaram a recuperação.

Tabela 1. Condições sociodemográficas, características físicas, comportamentais e de saúde dos pais, dados do nascimento e do primeiro ano de vida das crianças da amostra, em ambas as cidades, n=659.

<i>Variável</i>	<i>Freq.</i>	<i>%</i>	<i>Md (dp)</i>	<i>Min</i>	<i>Max</i>	<i>IC</i>
Características sociodemográficas						
Idade Materna						
< 15 anos	131	19.88				[0.173; 0.233]

15-35 anos	503	76.33	-	-	-	[0.748; 0.812]
> 35 anos	11	1.67				[0.007; 0.027]
Idade Paterna						
< 15 anos	-	-				-
15-35 anos	512	77.69	-	-	-	[0.766; 0.823]
>35 anos	133	20.18				[0.177; 0.236]
Anos de estudo						
até 8anos	140	21.40	-	-	-	[0.183; 0.246]
> 8anos	514	78.59				[0.754; 0.817]
Ocupação Materna						
Desempregada	341	51.74				[0.480; 0.556]
Manual não qualificada	80	12.14	-	-	-	[0.097; 0.147]
Manual esp./semi-esp.	88	13.35				[0.108; 0.160]
Qualificada não manual	149	22.61				[0.194; 0.258]
Com companheiro						
Sim	520	78.91	-	-	-	[0.758; 0.820]
Não	139	21.10				[0.180; 0.242]
Cor materna						
Branca	247	37.48	-	-	-	[0.338; 0.412]
Não branca	411	62.37				[0.588; 0.662]
Classe econômica (ABEP)						
A/B	93	14.11	-	-	-	[0.118; 0.173]
C	353	53.57				[0.514; 0.591]
D/E	193	29.29				[0.266; 0.338]
Núcleo familiar						
Ate 4	501	76.02	-	-	-	[0.728; 0.793]
5 ou mais	158	23.98				[0.207; 0.272]
Beneficiário bolsa família						
Não	534	81.03	-	-	-	[0.784; 0.844]
Sim	122	18.51				[0.156; 0.216]
Características físicas, comportamentais e de saúde dos pais						
Etilismo						
Não	520	78.90	-	-	-	[0.758; 0.820]
Sim	139	21.09				[0.180; 0.242]
Fumo atual ou pregresso						
Não	541	82.09	-	-	-	[0.793; 0.851]
Sim	117	17.75				[0.149; 0.207]

adequado: 7-10	584	88.62	-	-	-	[0.884; 0.929]
asfixia: 0-6	60	9.10				[0.071; 0.116]
Apgar 5o minuto						
adequado: 7-10	636	96.5	-	-	-	[0.979; 0.996]
asfixia: 0-6	8	1.21				[0.004; 0.021]
Sexo da criança						
Feminino	314	47.64	-	-	-	[0.438; 0.515]
Masculino	345	52.35				[0.485; 0.562]
Tipo de parto						
Normal	283	42.94	-	-	-	[0.392; 0.467]
Cesárea	376	57.05				[0.533; 0.608]
Peso para estatura	-	-	-196(1.56)	-5.45	1.37	[-2.05; -87]
Nasceu PT						
Extremo PT	1	0.6	-	-	-	[-0.006; 0.018]
Muito PT	5	3	-	-	-	[0.004; 0.056]
Moderado	72	43.11	-	-	-	[0.356; 0.506]
Atermo	89	53.29	-	-	-	[0.457; 0.609]
Perímetro cefálico			32.5 (1.85)	27,37	-	[32.18;32.75[]
Variáveis do 1º ano de vida						
Amamentação exclusiva até os 6 meses						
Não	557	84.527	-	-	-	[0.82; 0.875]
Sim	100	15.17				[0.125; 0.180]
Nº filhos						
-	-	-	1.75 (3.9)	1	8	[1.5;1.7]
Entubação 3 meses de vida						
Não	583	88.52	-	-	-	[0.884; 0.929]
Sim	60	15.17				[0.071; 0.116]
Catch-up estatural						
Não	332	50.38	-	-	-	[0.458; 0.534]
Sim	327	49.62				[0.466; 0.542]

Legenda: Freq=Freqüência absoluta; dp=desvio-padrão; Min=Mínimo; Máx=Máximo; IC=Intervalo de Confiança. Manual esp./semi-esp=Manual especializada ou semi-especializada. Hipert=Hipertensão; PN=Pré-Natal; US=Ultrassonografia; PT=Pretermo. Fonte: elaborada pelos autores (2025).

A maioria das mães tinha entre 15 e 35 anos (76,33%), enquanto apenas 1,67% tinham mais de 35 anos. Quanto à idade paterna, 77,69% dos pais estavam na faixa etária de 15 a 35 anos. A escolaridade materna predominante foi superior a oito anos de estudo (78,59%). Em relação à ocupação, 51,74% das mães estavam desempregadas, e 22,61% exerciam funções qualificadas não manuais. A maioria (78,91%) vivia com um companheiro. A cor materna foi predominantemente não branca (62,37%). A classe econômica mais frequente foi a classe C (53,57%), seguida das classes D/E (29,29%). A maioria das famílias tinha até quatro membros (76,02%). Quanto ao recebimento de benefícios sociais, 18,51% eram beneficiários do Bolsa Família.

A prevalência de tabagismo materno atual ou pregresso foi de 17,75%, enquanto o etilismo foi reportado por 21,09%. A prática de atividade física foi alta, com 91,65% das mães consideradas ativas. Cerca de 40,06% das mães apresentaram uma morbidade, enquanto 28,37% não relataram nenhuma. O diagnóstico de diabetes prévio ou durante a gestação foi relatado por 5,77%, e a hipertensão gestacional ocorreu em 25,49% das mães. As médias da altura paterna e materna foram 172 cm e 157 cm, respectivamente.

O início do pré-natal no primeiro trimestre foi relatado por 76,93% das mães. A maioria realizou mais de seis consultas (69,04%), e 72,99% fizeram mais de três ultrassonografias. O parto ocorreu majoritariamente no SUS (65,85%). O escore de Apgar no primeiro minuto foi adequado (7-10) em 88,62% dos casos, e no quinto minuto, essa proporção aumentou para 96,5%. O sexo da criança foi equilibrado, com 52,35% do sexo masculino. O parto cesáreo foi mais frequente (57,05%).

Quanto à idade gestacional, 53,29% dos nascidos foram a termo, e 43,11% foram moderadamente pré-termo. O perímetro cefálico médio ao nascer foi de 32,5 cm. A amamentação exclusiva até os seis meses foi relatada por apenas 15,17% das mães. A média de filhos por mãe foi de 1,75. O histórico de entubação nos primeiros três meses de vida foi observado em 15,17% dos casos.

3.1 Cenário 1: amostra conjugada de São Luís e Ribeirão Preto

Neste cenário, a base de dados compreende um total de 359 observação, descritas por 40 variáveis preditoras, distribuídas entre 332 crianças que não tiveram *catch-up* estatural (classe de referência) e 327 que experimentaram a recuperação. Os resultados preditivos podem ser observados na Tabela 2. Ela exhibe, para cada par classificador/métrica a média dos

50 valores obtidos e o intervalo de confiança correspondente. São apresentados também os resultados da classificação quando a seleção de variáveis é realizada.

Antes da aplicação da técnica de seleção, os modelos apresentaram desempenhos similares, com acurácia variando entre 68,1% e 69%, sendo o SVM o que obteve maior sensibilidade (71,7%). A especificidade foi levemente inferior para todos os modelos, variando entre 65,8% e 67%. O F1-score e a AUC variaram com valores máximos de 69,7% e 75,3%, respectivamente. Vale ressaltar que o modelo Nulo pode ser incluído na comparação, servido de base para o desempenho dos demais métodos. Ele representa o classificador que sempre prediz de acordo com a classe majoritária. Sendo assim, para esta abordagem ou a sensibilidade ou a especificidade é sempre 0 e o AUC é 50%.

Utilizando os testes estatísticos de Friedman (para comparação global) e Wilcoxon com correção de Bonferroni (para comparação par a par), ambos com nível de significância de 95%, tem-se que todos os classificadores foram melhores que o modelo Nulo em termos de acurácia e AUC. Entre si, a significância estatística da diferença dos classificadores varia de acordo com a métrica de desempenho. Assim, para a acurácia, há diferença entre os pares: LDA/RLR e LDA/SVM, apesar da sobreposição dos intervalos de confiança. Para a sensibilidade, as diferenças ocorrem entre LDA/SVM e SVM/XGBoost. Para a especificidade e precisão, são diferentes os pares LDA/RLR. Para o F1-Score, observam-se diferenças entre LDA/RLR, LDA/SVM, RLR/XGBoost e SVM/XGBoost. Para a AUC, há diferenças entre LDA/RLR e RLR/SVM.

Com a aplicação do Boruta, houve um leve aumento na sensibilidade dos modelos, especialmente para o LDA e o RLR (72,8% e 72,7%, respectivamente), indicando uma melhor identificação dos casos nos quais as crianças não experimentaram o *catch-up* estatural, de fato. No entanto, a especificidade apresentou uma leve redução, particularmente para o SVM (64,3%). A acurácia manteve-se estável, e a AUC permaneceu praticamente inalterada. Novamente, todos os classificadores aqui mostraram-se superiores ao modelo Nulo, considerando-se a acurácia e o AUC. Considerando comparação par a par, tem-se as seguinte diferenças estatisticamente significativas, agrupadas por métrica: acurácia (LDA/LRL, LDA/SVM, RLR/SVM e RLR/XGBoost), sensibilidade (LDA/SVM, LDA/XGBoost, RLR/SVM, RLR/XGBoost e SVM/XGBoost), especificidade (LDA/RLR, LDA/XGBoost e SVM/XGBoost), precisão (LDA/RLR, LDA/SVM, RLR/SVM e SVM/XGBoost), F1-Score (LDA/SVM, LDA/XGBoost, RLR/SVM e RLR/XGBoost) e AUC (LDA/SVM e RLR/SVM).

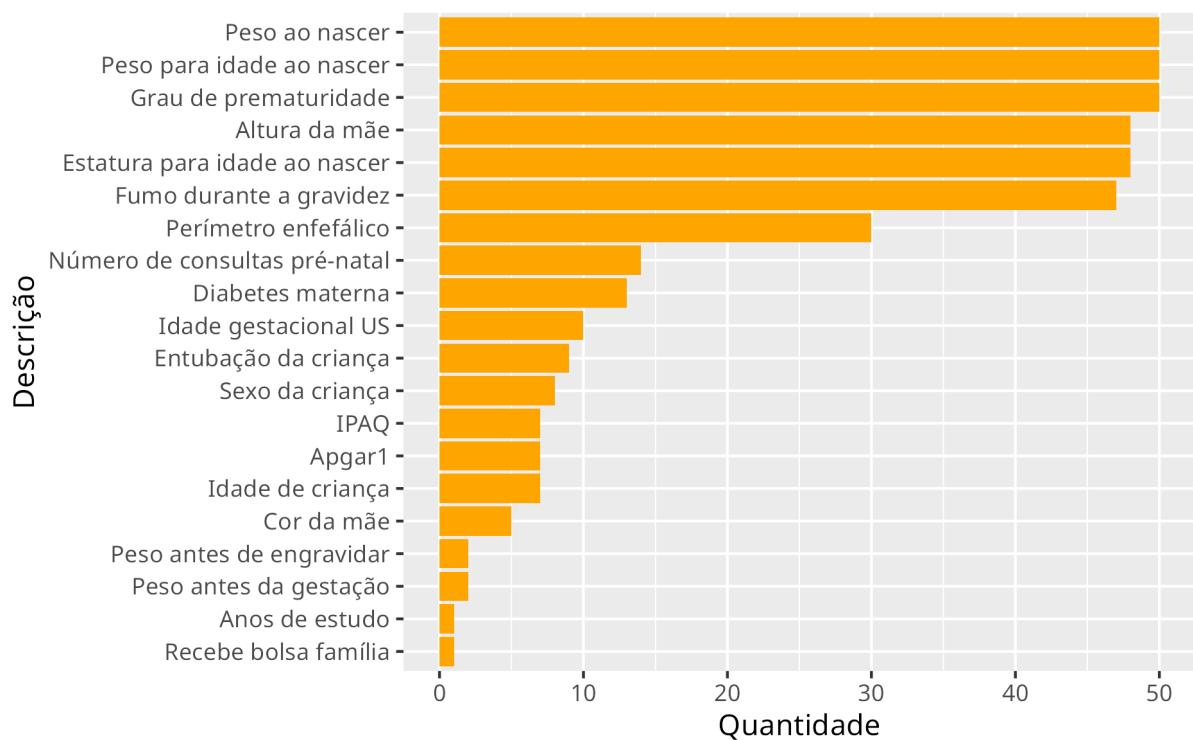
Tabela 2. Resultado comparativo entre os algoritmos, considerando os desempenhos com e sem o selecionador de atributos Boruta, na base de dados conjunta.

	Algoritmo	Acurácia	Sensibilidade	Especificidade	Precisão	F1-Score	AUC
Sem Boruta	LDA	0.682 [0.671,0.692]	0.705 [0.687,0.722]	0.658 [0.641,0.675]	0.678 [0.667,0.688]	0.69 [0.678,0.701]	0.746 [0.734,0.757]
	RLR	0.69 [0.678,0.702]	0.71 [0.693,0.727]	0.67 [0.652,0.689]	0.688 [0.675,0.70]	0.697 [0.685,0.709]	0.753 [0.741,0.765]
	SVM	0.69 [0.677,0.702]	0.717 [0.699,0.734]	0.662 [0.645,0.679]	0.684 [0.672,0.696]	0.669 [0.686,0.711]	0.746 [0.733,0.759]
	XGBoost	0.681 [0.671,0.69]	0.701 [0.685,0.717]	0.66 [0.642,0.678]	0.678 [0.668,0.689]	0.688 [0.678,0.698]	0.747 [0.736,0.758]
	Nulo	0.504[0.504, 0.504]	1 [1,1]	0 [0,0]	0.504 [0.504,0.504]	0.67 [0.67, 0.67]	0.5 [0.5,0.5]
Com Boruta	LDA	0.686 [0.674,0.698]	0.728 [0.712,0.744]	0.643 [0.625,0.66]	0.675 [0.664,0.687]	0.70 [0.688,0.711]	0.747 [0.733,0.761]
	RLR	0.69 [0.677,0.703]	0.727 [0.71,0.745]	0.652 [0.633,0.67]	0.681 [0.668,0.693]	0.702 [0.689,0.715]	0.749 [0.735,0.762]
	SVM	0.676 [0.665,0.688]	0.709 [0.691,0.727]	0.643 [0.628,0.657]	0.669 [0.658,0.679]	0.687 [0.675,0.70]	0.741 [0.729,0.754]
	XGBoost	0.68 [0.669,0.692]	0.697 [0.681,0.714]	0.663 [0.646,0.68]	0.679 [0.667,0.69]	0.687 [0.675,0.689]	0.741 [0.728,0.754]
	Nulo	0.504[0.504, 0.504]	1 [1,1]	0 [0,0]	0.504 [0.504,0.504]	0.67 [0.67, 0.67]	0.5 [0.5,0.5]

Fonte: elaborada pelos autores (2025).

De forma geral, os resultados sugerem que a aplicação do Boruta trouxe algum ganho para a análise, principalmente na sensibilidade dos modelos, sem impacto significativo na AUC. A seleção de variáveis, portanto, pode ter melhorado a capacidade dos modelos de identificar corretamente as crianças sem *catch-up* estatural, mas sem grandes alterações no desempenho geral. Abaixo, é possível observar a Figura 8, que mostra os atributos mais frequentemente selecionados pelo Boruta, ao longo de 50 execuções.

Figura 8. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados completa.



Fonte: elaborada pelos autores (2025).

Dentre as características que tiveram maior contribuição para os modelos em ambas as cidades, as que aparecem em pelo menos 45 vezes foram: peso ao nascer, peso para idade ao nascer, prematuridade, altura da mãe, estatuta para idade ao nascer, fumo. Alguns destes preditores apareceram em todas as rodadas, sugerindo se tratar de atributos de maior relevância. Estão elas relacionadas ao peso, ao nascimento e à prematuridade.

3.2 Cenário 2: amostra de São Luís

Neste cenário, a base de dados compreende um total de 324 observação, descritas por 40 variáveis preditoras, distribuídas entre 167 crianças que não tiveram *catch-up* estatural (classe de referência) e 157 que experimentaram a recuperação.

A Tabela 3 exibe os resultados de forma semelhante à tabela anterior. É possível notar que os modelos apresentaram desempenhos moderados, com acurácia variando entre 67,2% e 69,9% e AUC entre 72,4% e 75,8%, indicando capacidade preditiva superior ao acaso (modelo Nulo: AUC = 50%). A inclusão da seleção de variáveis com Boruta resultou em pequenas melhorias, especialmente na sensibilidade e precisão dos modelos LDA, RLR e SVM. Os testes estatísticos confirmam a superioridade de todos os classificadores frente ao modelo Nulo. A partir deste cenário, as demais comparações par a par serão publicadas em material suplementar futuro.

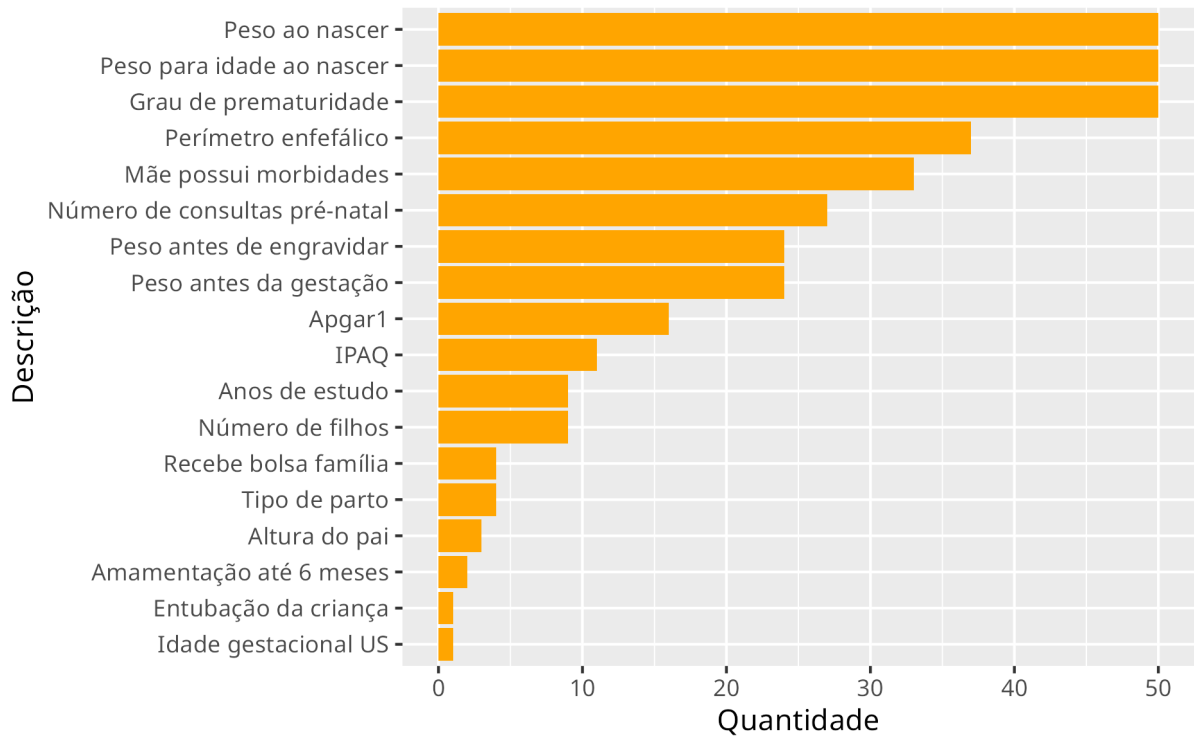
O RLR se destacou como o modelo mais equilibrado, apresentando maior acurácia de (71,3%) e maior AUC (74,5%). O SVM e o LDA tiveram desempenhos semelhantes. O XGBoost, embora amplamente utilizado para modelagem preditiva, apresentou os menores valores de acurácia, F1-score e AUC (65,5%, 67,2% e 70,9%, respectivamente), sugerindo que pode não ser a melhor abordagem para este conjunto de dados. Portanto, a seleção de variáveis pode melhorar em algum grau o desempenho, mas a escolha do algoritmo ainda tem maior impacto nos resultados. O RLR pode ser uma opção interessante, ao conseguir equilibrar sensibilidade e especificidade.

Tabela 3. Resultado comparativo entre os algoritmos, considerando SL.

	Algoritmo	Acurácia	Sensibilidade	Especificidade	Precisão	F1-Score	AUC
Sem Boruta	LDA	0.688 [0.672,0.703]	0.732 [0.707,0.757]	0.64 [0.616,0.663]	0.685 [0.67,0.701]	0.705 [0.688,0.723]	0.743 [0.726,0.76]
	RLR	0.699 [0.683,0.715]	0.725 [0.699,0.75]	0.671 [0.648,0.695]	0.703 [0.687,0.718]	0.711 [0.693,0.729]	0.758 [0.74,0.775]
	SVM	0.688 [0.67,0.706]	0.721 [0.695,0.747]	0.652 [0.624,0.68]	0.691 [0.672,0.711]	0.703 [0.684,0.721]	0.745 [0.727,0.762]
	XGBoost	0.672 [0.656,0.687]	0.693 [0.666,0.721]	0.648 [0.626,0.669]	0.677 [0.662,0.692]	0.683 [0.663,0.702]	0.724 [0.704,0.744]
	Nulo	0.515[0.515, 0.515]	1 [1,1]	0 [0,0]	0.515 [0.515,0.515]	0.68 [0.68, 0.68]	0.5 [0.5,0.5]
Com Boruta	LDA	0.709 [0.692,0.726]	0.75 [0.727,0.773]	0.665 [0.643,0.686]	0.705 [0.689,0.721]	0.725 [0.708,0.742]	0.744 [0.725,0.763]
	RLR	0.713 [0.696,0.73]	0.739 [0.715,0.763]	0.685 [0.663,0.707]	0.715 [0.699,0.732]	0.725 [0.708,0.743]	0.745 [0.726,0.764]
	SVM	0.713 [0.697,0.729]	0.739 [0.715,0.763]	0.684 [0.661,0.707]	0.716 [0.7,0.731]	0.725 [0.709,0.742]	0.742 [0.724,0.759]
	XGBoost	0.655 [0.64,0.67]	0.688 [0.666,0.711]	0.619 [0.595,0.644]	0.66 [0.645,0.676]	0.672 [0.656,0.687]	0.709 [0.692,0.727]
	Nulo	0.515[0.515, 0.515]	1 [1,1]	0 [0,0]	0.515[0.515, 0.515]	0.68 [0.68, 0.68]	0.5 [0.5,0.5]

Fonte: elaborada pelos autores (2025).

Figura 9. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados SL.



Fonte: elaborada pelos autores (2025).

Dentre as características que tiveram maior contribuição para os modelos com validação cruzada em SL, as que aparecem em todas as rodadas foram: peso ao nascer, peso para idade ao nascer, prematuridade, sugerindo se tratar de atributos de maior relevância.

3.3 Cenário 3: amostra de Ribeirão Preto

Neste cenário, a base de dados compreende um total de 335 observação, descritas por 40 variáveis preditoras, distribuídas entre 165 crianças que não tiveram catch-up estatural (classe de referência) e 170 que experimentaram a recuperação.

A Tabela 4 exibe os resultados preditivos deste cenário. Os modelos apresentaram um desempenho relativamente próximo entre si, com acurácia variando entre 64,8% e 67,6% e AUC entre 71,1% e 73,4%. O XGBoost obteve a maior acurácia (67,6%). O RLR apresentou a segunda melhor acurácia (66,6%) com desempenho próximo ao do XGBoost quanto a sensibilidade (66,6%) e AUC (73,2%). O SVM, por outro lado, apresentou o menor desempenho geral, com acurácia (64,8%) e F1-score (63,9%) abaixo dos demais modelos. Os

testes estatísticos confirmam a superioridade de todos os classificadores frente ao modelo Nulo.

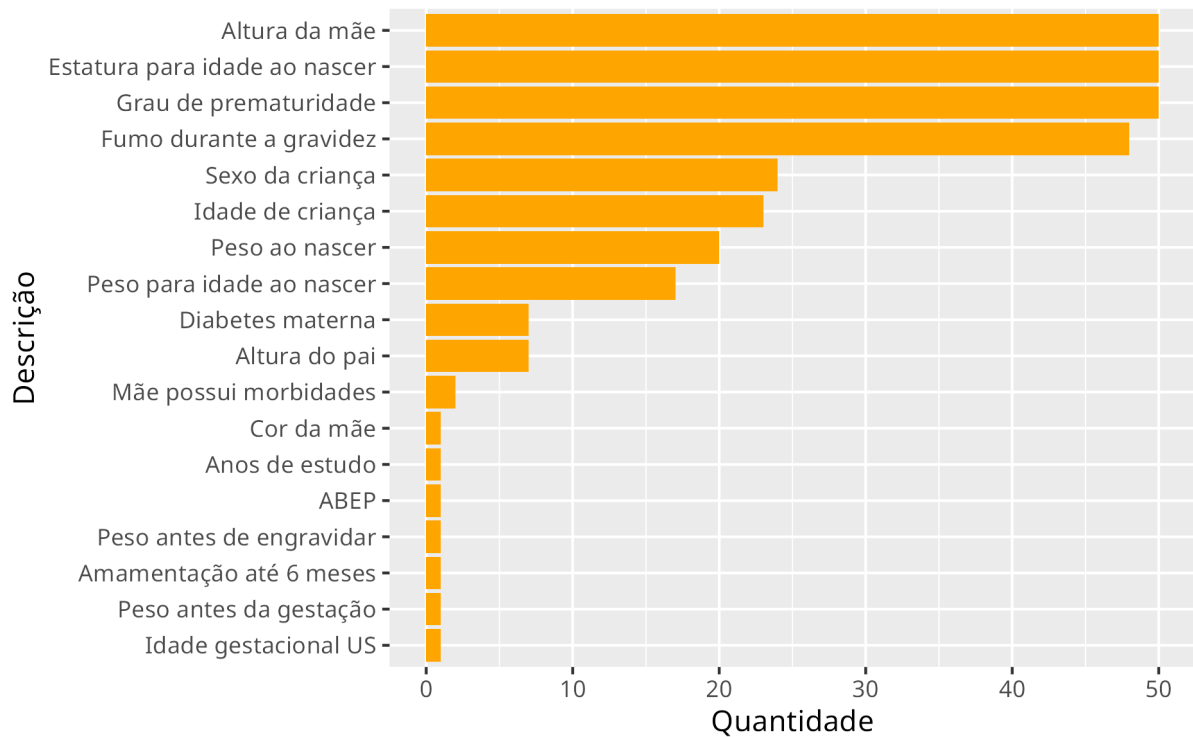
Todos os algoritmos apresentaram valores maiores para acurácia, com o uso do Boruta: LDA=69,4%, RLR=69%, SVM=70,5% e Xgboost=69%. Da mesma forma, tiveram aumento para sensibilidade, especificidade, precisão, *F1-score* e AUC para todos os algoritmos analisados. O SVM apresentou o maior ganho, atingindo em acurácia e AUC, tornando-se o melhor modelo após a aplicação do selecionador de variáveis, sugerindo que esse algoritmo pode ter dificuldades para lidar com o conjunto de dados sem uma seleção prévia de variáveis relevantes. O RLR e o LDA também apresentaram melhoria em suas métricas, com aumentos sutis na sensibilidade e especificidade. Já o XGBoost, apesar da melhora na AUC (74,2%), não obteve um aumento expressivo na acurácia em comparação aos outros modelos.

Tabela 4. Resultado comparativo entre os algoritmos, considerando RP.

	Algoritmo	Acurácia	Sensibilidade	Especificidade	Precisão	F1-Score	AUC
Sem Boruta	LDA	0.658 [0.644,0.671]	0.653 [0.632,0.673]	0.662 [0.636,0.688]	0.657 [0.64,0.674]	0.652 [0.638,0.665]	0.726 [0.712,0.74]
	RLR	0.666 [0.653,0.679]	0.666 [0.642,0.69]	0.666 [0.64,0.692]	0.664 [0.649,0.679]	0.661 [0.647,0.675]	0.732 [0.718,0.746]
	SVM	0.648 [0.634,0.661]	0.637 [0.613,0.661]	0.658 [0.633,0.683]	0.684 [0.632,0.663]	0.639 [0.624,0.654]	0.711 [0.696,0.727]
	XGBoost	0.676 [0.663,0.688]	0.666 [0.641,0.691]	0.685 [0.662,0.708]	0.676 [0.661,0.69]	0.667 [0.652,0.682]	0.734 [0.716,0.752]
	Nulo	0.507[0.507, 0.507]	0 [0,0]	1 [1,1]	NA [NA,NA]	NA [NA,NA]	0.5 [0.5,0.5]
Com Boruta	LDA	0.694 [0.681,0.706]	0.687 [0.665,0.709]	0.701 [0.678,0.723]	0.694 [0.679,0.708]	0.687 [0.673,0.701]	0.763 [0.75,0.775]
	RLR	0.69 [0.677,0.704]	0.692 [0.67,0.715]	0.689 [0.665,0.713]	0.687 [0.672,0.703]	0.687 [0.672,0.705]	0.76 [0.748,0.773]
	SVM	0.705 [0.694,0.717]	0.69 [0.667,0.713]	0.721 [0.699,0.743]	0.709 [0.695,0.723]	0.696 [0.682,0.71]	0.766 [0.753,0.778]
	XGBoost	0.69 [0.679,0.701]	0.686 [0.662,0.71]	0.694 [0.669,0.719]	0.69 [0.676,0.704]	0.684 [0.671,0.697]	0.742 [0.729,0.755]
	Nulo	0.507[0.507, 0.507]	0 [0,0]	1 [1,1]	NA [NA,NA]	NA [NA,NA]	0.5 [0.5,0.5]

Fonte: elaborada pelos autores (2025).

Figura 10. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados RP.



Fonte: elaborada pelos autores (2025).

Dentre as características que tiveram maior contribuição para os modelos com validação cruzada em RP, as que aparecerem em mais de 45 das 50 rodadas foram: altura da mãe, estatura para idade ao nascer, prematuridade e fumo, sugerindo se tratar de atributos de maior relevância. Outras, em menor frequência, também foram apontadas: sexo do bebê, idade da criança, peso no nascimento.

3.4 Cenário 4: treino com São Luís e teste com Ribeirão Preto

Neste cenário, os algoritmos de AM são treinados com a amostra de São Luís e o modelo resultante é testado com a amostra de Ribeirão Preto. O racional deste cenário é propiciar ao classificador uma validação completamente externa. Os resultados preditivos podem ser observados na Tabela 5. Ela exhibe, para cada par classificador/métrica o valor pontual do desempenho e o intervalo de confiança correspondente. Como nos cenários anteriores, são apresentados também os resultados da classificação quando a seleção de variáveis é realizada.

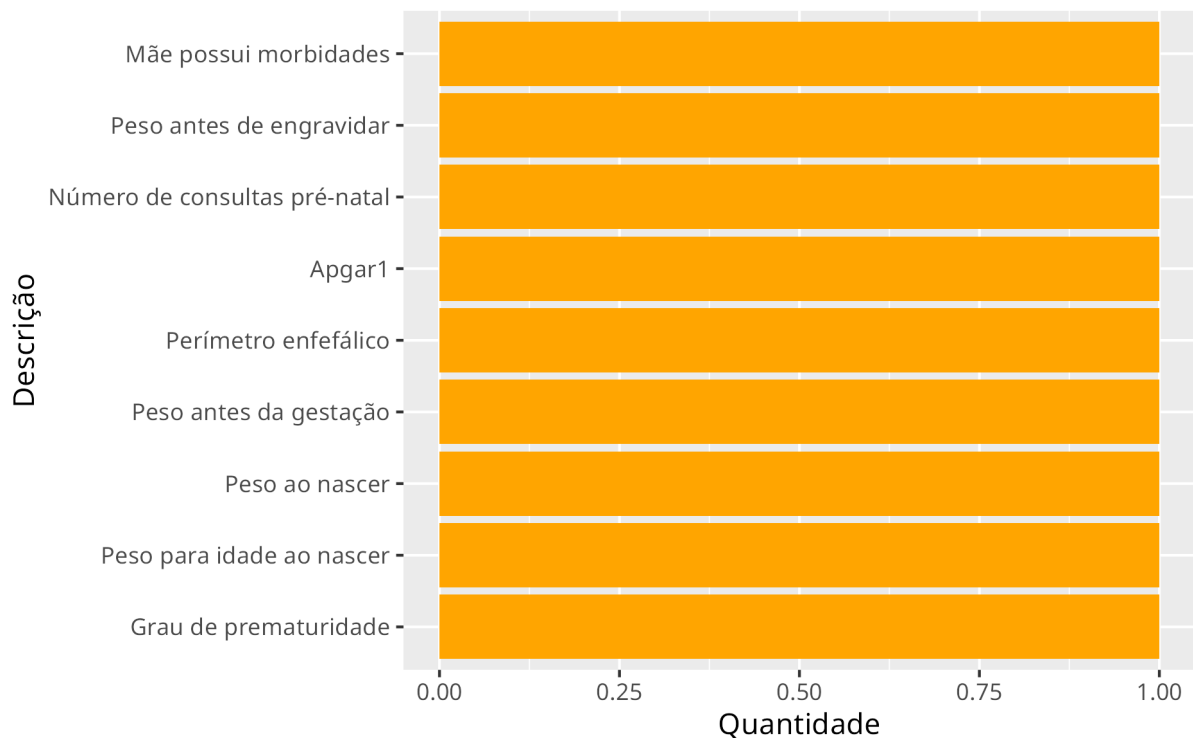
Os resultados indicam que todos os modelos apresentam desempenho moderado, com acurácia variando entre 57,5% e 62,7%, e AUC entre 61,6% e 68,3%. A seleção de variáveis não trouxe melhorias expressivas, e em alguns casos, houve até uma leve redução nos valores de acurácia e AUC (LDA, RLR e *XGboost*). Entre os modelos testados, o RLR e o LDA tiveram as maiores acurácias 62,7% e 61,5%, respectivamente sem o Boruta, enquanto o SVM apresentou leve melhora após a aplicação do método de seleção de variáveis (passando de 60,6% para 62,4%). Aplicando o teste de McNemar, verificou-se que todos os classificadores superaram o modelo Nulo quando todas as variáveis preditoras são consideradas. Com a aplicação do Boruta, o LDA, o RLR e o SVM mostram-se superiores, enquanto o XGBoost, não.

Tabela 5. Resultado comparativo entre os algoritmos, considerando treino em SL e teste em RP.

	Algoritmo	Acurácia	Sensibilidade	Especificidade	Precisão	F1-Score	AUC
Sem Boruta	LDA	0.615 [0.558,0.669]	0.655 [0.577,0.728]	0.576 [0.5,0.655]	0.6 [0.523,0.673]	0.626 [0.56,0.686]	0.669 [0.612,0.726]
	RLR	0.627 [0.57,0.678]	0.697 [0.623,0.766]	0.559 [0.479,0.636]	0.605 [0.525,0.677]	0.648 [0.583,0.705]	0.683 [0.626,0.739]
	SVM	0.606 [0.552,0.657]	0.612 [0.538,0.689]	0.6 [0.52,0.673]	0.598 [0.517,0.673]	0.605 [0.537,0.667]	0.659 [0.601,0.717]
	XGBoost	0.609 [0.558,0.666]	0.685 [0.608,0.757]	0.535 [0.452,0.616]	0.589 [0.512,0.661]	0.633 [0.57,0.692]	0.642 [0.583,0.7]
	Nulo	0.493 [0.436,0.546]	1 [1,1]	0 [0,0]	0.493 [0.436,0.546]	0.66 [0.607,0.707]	0.5 [0.5,0.5]
Com Boruta	LDA	0.609 [0.552,0.663]	0.709 [0.635,0.776]	0.512 [0.433,0.589]	0.585 [0.508,0.656]	0.641 [0.575,0.699]	0.637 [0.578,0.696]
	RLR	0.609 [0.558,0.663]	0.673 [0.595,0.741]	0.547 [0.475,0.629]	0.59 [0.514,0.665]	0.629 [0.563,0.687]	0.633 [0.574,0.692]
	SVM	0.624 [0.573,0.681]	0.63 [0.554,0.702]	0.618 [0.542,0.694]	0.615 [0.538,0.693]	0.623 [0.558,0.683]	0.625 [0.565,0.685]
	XGBoost	0.576 [0.522,0.633]	0.679 [0.607,0.753]	0.476 [0.399,0.555]	0.557 [0.484,0.628]	0.612 [0.551,0.674]	0.616 [0.556,0.676]
	Nulo	0.493 [0.436,0.546]	1 [1,1]	0 [0,0]	0.493 [0.436,0.546]	0.66 [0.607,0.707]	0.5 [0.5,0.5]

Fonte: elaborada pelos autores (2025).

Figura 11. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados SL como teste e RP como treino.



Fonte: elaborada pelos autores (2025).

Dentre as características que tiveram maior contribuição para os modelos treinados em SL e testados em RP, as que aparecem em todas as rodadas foram: morbidades (quando a mãe apresenta algum tipo de doença: diabetes, hipertensão...), peso antes de engravidar, número de consultas pré-natal, apgar no 1º minuto, perímetro cefálico, peso da mãe antes da gestação, peso no nascimento, peso para idade ao nascer, prematuridade, sugerindo se tratar de atributos de maior relevância.

3.5 Cenário 5: treino com Ribeirão Preto e teste com São Luís

Neste cenário, os algoritmos de AM são treinados com a amostra de Ribeirão Preto e o modelo resultante é testado com a amostra de São Luís. Novamente, o racional do cenário é propiciar ao classificador uma validação completamente externa. Os resultados preditivos podem ser observados na Tabela 6, de forma semelhante ao visto no cenário anterior. De maneira geral, os achados foram muito aproximados para os algoritmos experimentados. A acurácia, que mede a proporção de previsões corretas em relação ao total de previsões, indica

a capacidade geral do modelo de fazer previsões corretas. Nos algoritmos testados, o melhor foi SVM, com 66%. A sensibilidade encontrada, considerando o *SVM*, revela que dentre todas as crianças que não tiveram *catch-up*, 47,3% delas de fato não experimentaram o fenômeno. Tal resultado foi superior ao RLR (46,7%), LDA (43,1%) e *XGboost* (38,9%).

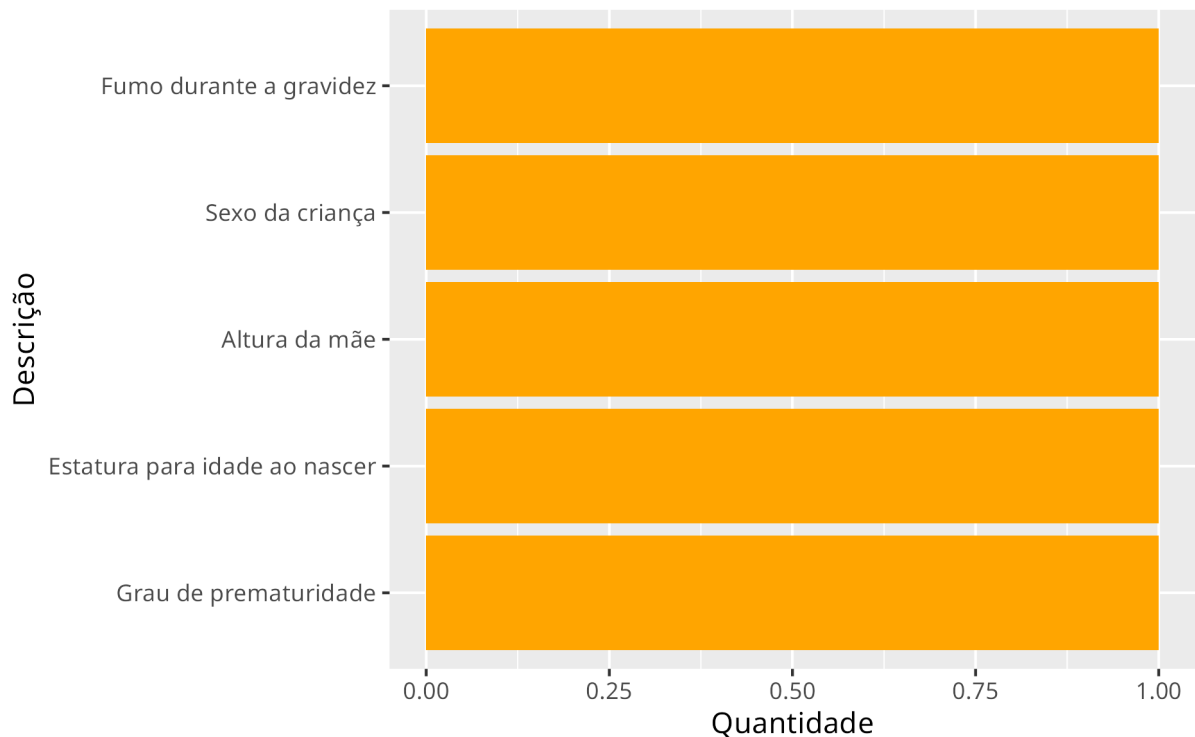
Os resultados indicam um desempenho moderado dos modelos testados, com acurácia variando entre 60,8% e 66% sem o Boruta e entre 60,5% e 64,2% com o Boruta. A especificidade dos modelos é consideravelmente maior do que a sensibilidade, sugerindo uma certa dificuldade em identificar corretamente as crianças sem *catch-up* dentre aquelas que não experimentaram. Nesse cenário, os modelos apresentam boa especificidade e baixa sensibilidade, possivelmente indicando algum viés. O teste de McNemar confirma a superioridade de todos os métodos frente ao modelo Nulo.

Tabela 6. Resultado comparativo entre os algoritmos, considerando treino em RP e teste em SL.

	Algoritmo	Acurácia	Sensibilidade	Especificidade	Precisão	F1-Score	AUC
Sem Boruta	LDA	0.645 [0.596,0.698]	0.431 [0.352,0.508]	0.873 [0.818,0.923]	0.783 [0.695,0.871]	0.556 [0.477,0.625]	0.713 [0.657,0.769]
	RLR	0.657 [0.605,0.707]	0.467 [0.394,0.546]	0.86 [0.803,0.911]	0.78 [0.697,0.862]	0.584 [0.515,0.655]	0.726 [0.671,0.781]
	SVM	0.66 [0.608,0.707]	0.473 [0.396,0.549]	0.86 [0.805,0.911]	0.782 [0.698,0.865]	0.59 [0.516,0.655]	0.73 [0.675,0.785]
	XGBoost	0.608 [0.556,0.664]	0.389 [0.317,0.464]	0.841 [0.781,0.895]	0.722 [0.625,0.815]	0.506 [0.429,0.577]	0.668 [0.61,0.727]
	Nulo	0.485 [0.426,0.537]	0 [0,0]	1 [1,1]	NA [NA,NA]	NA [NA,NA]	0.5 [0.5,0.5]
Com Boruta	LDA	0.627 [0.574,0.676]	0.425 [0.352,0.503]	0.841 [0.783,0.892]	0.74 [0.651,0.824]	0.54 [0.467,0.611]	0.743 [0.689,0.797]
	RLR	0.63 [0.577,0.679]	0.431 [0.358,0.509]	0.841 [0.783,0.892]	0.742 [0.654,0.827]	0.545 [0.471,0.614]	0.74 [0.686,0.795]
	SVM	0.642 [0.593,0.694]	0.479 [0.406,0.553]	0.815 [0.753,0.873]	0.734 [0.65,0.815]	0.58 [0.511,0.643]	0.7 [0.643,0.757]
	XGBoost	0.605 [0.549,0.657]	0.437 [0.361,0.515]	0.783 [0.713,0.844]	0.682 [0.588,0.77]	0.533 [0.46,0.602]	0.688 [0.63,0.746]
	Nulo	0.485 [0.426,0.537]	0 [0,0]	1 [1,1]	NA [NA,NA]	NA [NA,NA]	0.5 [0.5,0.5]

Fonte: elaborada pelos autores (2025).

Figura 12. Seleção de atributos pelo algoritmo Boruta, considerando a base de dados RP como teste e SL como treino.



Fonte: elaborada pelos autores (2025).

Dentre as características que tiveram maior contribuição para os modelos treinados em RP e testados em SL, as que aparecerem em todas as rodadas foram: fumo, sexo do bebê, altura da mãe, estatura para idade ao nascer e prematuridade, sugerindo se tratar de atributos de maior relevância. Estão elas relacionadas ao peso, ao nascimento e à prematuridade.

4 Discussão

De modo geral, a ligeira maioria das crianças da amostra estudada, que nasceram baixas ou muito baixas para idade, não atingiram o *catch-up* estatural (50.38% contra 49.62%). No contexto do crescimento, a expectativa é que a maioria dos recém-nascidos com algum déficit inicial, como tamanho reduzido, decorrentes de prematuridade e/ou CIUR, apresentem o *catch-up*, atingindo seu canal de crescimento entre os percentis de normalidade nas curvas de referência até os 2-3 anos de idade (Vila; Macedo; Melo, 2017).

Diante disso, crianças cujas curvas de crescimento não se aproximam dos percentis mínimos da normalidade e/ou apresentam padrão descendente, requerem investigação e

precisam ser identificadas. No presente estudo, a classe positiva foram as crianças terem experimentado o fenômeno, enquanto a negativa foi não terem alcançado. Assim, a especificidade se torna uma medida de desempenho muito importante e variou de 47,6% a 70,1%, considerando todos os modelos e os 5 cenários.

Entre todos, o cenário de São Luís, utilizando o LDA com Boruta teve o melhor desempenho, de modo geral: acurácia de 70,9%, sensibilidade de 75%, precisão de 70,5%, F1 Score de 72,5% e AUC de 74,4%. No entanto, apresentou uma especificidade de 66,5%. Os cenários com validação externa apresentaram os menores e maiores valores para especificidade, enquanto os cenários com validação interna foram mais estáveis para essa variável.

A menor especificidade esteve presente em SL, como treino, e RP, como teste (variando de 47,6% a 61,8%), com o modelo *Xgboost* com Boruta apresentam o menor desempenho para essa métrica. Já a maior especificidade foi em RP, como treino, e SL, como teste (variando de 78,3% a 87,3%), no qual o LDA sem Boruta teve o maior valor obtido. Houve queda nas especificidades comparando sem e com adição do Boruta nos cenários de SL e RP juntas, SL treino e RP teste, e vice-versa, sugerindo que o selecionador de atributos possa interferir no desempenho dessa medida de avaliação quando ambas as bases estavam envolvidas.

No cenário com junção de ambas as cidades para treino e teste (validação cruzada), os modelos em LDA e RLR com Boruta tiveram os melhores desempenhos em sensibilidade (72,8% e 72,7%), porém sem ganhos expressivos para acurácia $\cong$ 68 e 69%, especificidade $\cong$ 64% e 65%, precisão $\cong$ 68%, F1 score $\cong$ 70% e AUC $\cong$ 75%, cada. Todos os modelos apresentaram resultados estatisticamente superiores ao método padrão, aquele que aponta a classe majoritária como desfecho (até 51,5%).

Neste estudo, fumo materno, sexo da criança, altura da mãe, estatura para idade ao nascer, grau de prematuridade, morbidades, peso da mãe antes da gestação, quantidade de consultas de PN, apgar no 1º minuto de vida, perímetro cefálico, peso do recém nascido, peso para idade ao nascer foram as características com algum nível de contribuição para os modelos, indicadas pelo Boruta, nos cenários propostos.

Algumas dessas variáveis estão descritas na literatura associadas à temática, de forma coincidente, uma vez não é o intuito da predição com AM analisar medidas de associação ou causalidade. Parte delas foram transversais nos modelos construídos, como a prematuridade, apontada em todos, sugerindo ser um atributo importante. Estudos a associam ao atraso de

crescimento infantil e também a relacionam a efeitos adversos de curto e longo prazo, sendo a principal causa de morbidade e mortalidade perinatal em países desenvolvidos e em desenvolvimento, produzindo aumentos dos custos públicos de saúde (Sartika et al., 2021; Jiang et al., 2018).

Em RP, o fumo e a altura materna foram identificados em diferentes cenários, parecendo se tratar de variáveis relevantes para o desfecho nessa cidade. O fumo materno também foi encontrado por Rotem et al. (2020) e Roro; Deressa; Lindtjorn (2022) como um dos fatores de risco associados a resultados desfavoráveis. Por ser um hábito modificável, o tabagismo representa uma oportunidade para promover a saúde materno-infantil, incentivando ações de cessação que beneficiem mães e crianças. Ademais, a altura materna já foi documentada por Roro, Deressa e Lindtjorn (2022), muito provavelmente pelo seu componente genético/biológico.

Como a prematuridade está relacionada ao CIUR, Batista et al (2012) revelaram que boa parte das crianças nascidas com essa condição conseguem recuperar o crescimento nos primeiros dois anos de vida. Porém, não é incomum também continuarem com essa conformação, ou seja, não experimentarem o *catch-up* estatural, quando comparados às nascidas sem CIUR, conforme também encontrado nesta pesquisa.

Crianças nascidas prematuras têm maior risco de apresentar crescimento inicial retardado devido à imaturidade dos sistemas corporais e à exposição a desafios neonatais, como complicações respiratórias e dificuldades alimentares. No entanto, Silva; Pitta; Almeida (2022) aponta que o *catch-up* de estatura pode ser mais lento ou incompleto conforme condições gerais de saúde e suporte recebido.

O mesmo fenômeno ocorre com relação à idade gestacional: crianças nascidas prematuras têm maior probabilidade de apresentar *catch-up* de altura se receberem acompanhamento adequado. Quanto ao peso ao nascer, bebês com baixo peso ao nascer podem experimentar crescimento compensatório, dependendo de condições adequadas, como nutrição e cuidados pós-natais (Kale et al., 2018; Lu et al., 2020, Oghli et al.; Roro, Deressa e Lindtjorn, 2022).

A estatura inicial é um marcador de crescimento pré-natal. Crianças que nascem com comprimento limitado, mas que não apresentam complicações graves ou condições restritivas contínuas, têm potencial de alcançar um *catch-up* significativo (Brasil, 2012; Ryckman; Berberich; Dagle, 2015). O estudo de Harrison et al. (2020) demonstrou como a importância

dessa métrica, destacando a necessidade de intervenções nos primeiros meses de vida, pode prevenir o atraso dos ganhos em altura subsequente e suas consequências.

Esses autores buscaram investigar os padrões de crescimento em bebês até os 48 meses de idade para avaliar se o crescimento de crianças com baixa estatura eventualmente melhorou, bem como os potenciais preditores de crescimento, por meio de modelos de AM, mais especificamente pelo *Random Forest*. Descobriu-se que das 107 crianças, o estado de atraso no crescimento da maioria delas aos 48 meses era o mesmo que aos 18 meses. O comprimento ao nascer foi a característica mais importante na previsão aos 18 meses. Dos biomarcadores, AGP (glicoproteína ácida alfa-1), PCR (proteína C reativa) e IL1 (interleucina-1) foram identificados como fortes preditores de crescimento subsequente em ambos os modelos de classificação e regressão (Harrison et al., 2020).

Semelhante a esse estudo, Harrison et al. (2020) referenciaram o comprimento (em escore z de altura para idade) como medida primária para sua investigação, especialmente no tocante à atrofia/déficit deste. No entanto, diferentemente do presente estudo, o desfecho foi distinto e as variáveis analisadas nem sempre são facilmente disponibilizadas para análise, como os biomarcadores. Outro fato importante foi o quantitativo, no qual foram selecionadas 6x mais do que a amostra de Harrison et al. (2020), que também explorou apenas por um algoritmo, com 78% de precisão, enquanto que nesta pesquisa foram considerados quatro algoritmos de diferentes famílias, comparando seus desempenhos que variaram de 55,7% a 78,3% quanto a precisão, entre os modelos.

Bitew, Sparks, Nyarko (2022), por outro lado, desenvolveram algoritmos preditivos para estado nutricional, considerando como um dos 3 indicadores, a altura para idade, em menores de 5 anos. Entre as técnicas utilizadas, estiveram o *XGBTree*, *k-NN*, *Random Forest*, rede neural e modelos lineares generalizados. O *XGBTree* demonstrou uma capacidade razoavelmente melhor para predição em comparação com os outros analisados e apresentou preditores importantes associados ao acesso ao abastecimento de água, histórico de anemia e baixo peso materno, entre outros.

No presente estudo, investigou-se uma faixa etária menor que o estudo de Bitew, Sparks, Nyarko (2022), considerando menores de 24 meses, dentro do intervalo dos primeiros mil dias de vida. Em alguns cenários, o *XGBoost* também teve destaque positivo, e obteve melhores modelos sem a seleção de variáveis. Por outro lado, algoritmos mais simples como o LDA e o PLR se beneficiaram da aplicação do Boruta. Isso pode indicar que o *XGBoost* já lida bem com variáveis irrelevantes internamente devido à sua estrutura baseada em árvores

de decisão. Dessa forma, os resultados sugerem que a escolha do algoritmo deve considerar também a capacidade do modelo de lidar com variáveis irrelevantes. O uso do Boruta demonstrou ser benéfico, especialmente para algoritmos mais sensíveis à escolha de variáveis.

Apesar da dificuldade de manter sob acompanhamento todos os indivíduos originalmente integrantes da coleta ao longo do tempo, os estudos de coortes de nascimento são fundamentais para estudar temas propostas pelo DOHaD, pois o seguimento permite investigar possíveis influências em diferentes períodos da vida (Oken et al, 2023), como pôde ser observado nesta pesquisa.

É importante ressaltar que a temática em estudo se trata de um fenômeno complexo. Para que a criança atinja todo o seu potencial, além dos fatores considerados, muitos outros estão envolvidos, como a nutrição, ausência de doenças crônicas, sono adequado, prática moderada de exercícios, saúde emocional, uso de telas, etc. A coleta do banco de dados utilizado não abrangeu todos esses aspectos. Portanto, não havia essas informações registradas, que podem também ser úteis para melhorar o desempenho dos modelos.

Este trabalho se alinha à Política Nacional de Atenção Integral à Saúde da Criança (PNAISC), importante estratégia instituída pelo Ministério da Saúde (MS), para o cuidado infantil (Brasil, 2015). Entre os eixos temáticos, está inserido no eixo III, que trata da promoção e acompanhamento do crescimento e do desenvolvimento integral, e IV acerca da atenção integral a crianças com agravos prevalentes na infância e com doenças crônicas. Por essa razão, tem o potencial de ser introduzido ao sistema via Prontuário Eletrônico do Cidadão (PEC), o software gratuito do MS, utilizado para registrar e armazenar informações sobre a saúde dos cidadãos.

O intuito é que o conhecimento gerado possa ser utilizado por profissionais que acompanham regularmente o crescimento infantil, desde a gravidez até a primeira infância. Assim, enfermeiros e médicos de família e comunidade na Atenção Primária, nas consultas de puericultura, bem como profissionais da Alta Complexidade em caso de gestação de risco e pediatras na Atenção Especializada, poderão usá-los no cuidado e atenção infantil. Da mesma forma, os algoritmos e modelos encontrados poderão ser úteis para pesquisadores de saúde infantil e outras partes interessadas, especialmente em estudos envolvendo predições com AM aplicadas na área da saúde.

Considerações finais

Neste trabalho, ao analisar 5 diferentes cenários de predição (validação interna ambas as cidades, apenas SL, apenas RP, treino em SL com teste em RP e vice-versa), considerando os dados das coortes de São Luís e Ribeirão Preto, observou-se que houve diferença em termos absolutos, apesar de os modelos não apresentarem diferença significativa estatisticamente. Em todos os casos, os modelos construídos por algoritmos de AM (LDA, RLR, SVM e *XGboost*) apresentam resultados melhores que o modelo nulo para referência.

Em alguns cenários, algoritmos mais simples como o LDA e o PLR se favoreceram da aplicação do Boruta, enquanto o *XGBoost*, obteve melhores modelos sem a seleção de variáveis. Dessa forma, os resultados sugerem que a escolha do algoritmo deve considerar também a capacidade do modelo de lidar com variáveis irrelevantes. O uso do Boruta demonstrou ser benéfico, especialmente para algoritmos mais sensíveis à escolha de variáveis.

As características mais relevantes para predizer o não *catch-up* estatutal nos primeiros 24 meses de idade foram: fumo materno, sexo da criança, altura da mãe, estatura para idade ao nascer, prematuridade, morbidades, peso da mãe antes da gestação, quantidade de consultas de PN, apgar no 1º minuto de vida, perímetro cefálico, peso do recém-nascido, peso para idade ao nascer, como documentados em estudos na literatura para fatores associados.

Abstract

Height Catch-up growth in Children in the First Thousand Days of Life: a Prediction Study using Machine Learning

Objective: To predict the non-occurrence of height catch-up up to 24 months of age in children born with shorter than expected length, using Machine Learning. **Methods:** The BRISA prenatal and birth cohort databases in the Brazilian cities of Ribeirão Preto (RP) and São Luís (SL) were used. Six hundred and fifty-nine children classified as short or very short for age at birth, according to the INTERGROWTH-21st chart, not twins, were selected and evaluated for height catch-up (change between height Z-score up to 2 years of age and length Z-score at birth >0.67) and potential predictors (socioeconomic variables, habits, parental characteristics, among others). The Linear Discriminant Analysis (LDA), Regularized

Logistic Regression (RLR), Support Vector Machine (SVM) and Extreme Gradient Boosting (XGBoost) algorithms were used to build prediction models, in 5 different scenarios, in the R studio program. **Results:** In this study, specificity was an important performance measure. Using the RP and SL base combined with cross-validation, the models in LDA and RLR with Boruta had the best performances in sensitivity (72.8% and 72.7%), but without significant gains for accuracy $\cong$ 68 and 69%, specificity $\cong$ 64% and 65%, precision $\cong$ 68%, F1 score $\cong$ 70% and AUC $\cong$ 75%, each. On the other hand, the SL base scenario, using LDA with Boruta, had the best performance: overall: accuracy of 70.9%, sensitivity of 75%, precision of 70.5%, F1 Score of 72.5% and AUC of 74.4%, but with specificity of 66.5%. The highest specificity was in RP, as training, and SL, as testing (ranging from 78.3% to 87.3%), in which LDA without Boruta had the highest value obtained. All models tested were statistically superior to the majority class (up to 51.5%). The variables that had the greatest contribution to the models were: birth weight, degree of prematurity, maternal height, height for age at birth, weight for age at birth, smoking. The degree of prematurity appeared in all models. **Conclusion:** Easily accessible variables related to the first thousand days of life were able to accurately predict the lack of height catch-up up to 2 years of age in children born with a length deficit.

Keywords: Child Health. Anthropometry. Supervised Machine Learning. Probability Learning.

6 CONSIDERAÇÕES FINAIS

Neste trabalho, aplicaram-se técnicas de AM para a predição de **não** alcance estatutal de crianças de até 24 meses de idade, utilizando dados sociodemográficos, de pré-natal, nascimento e primeiro ano da criança. Assim, foi possível comparar diferentes tipos de algoritmos de AM (LDA, RLR, SVM e *XGboost*) a fim de identificar os melhores desempenhos em 5 diferentes cenários de predição. Em todos os casos, os modelos construídos por AM apresentam resultados melhores que o modelo nulo para referência.

Em alguns cenários, algoritmos mais simples como o LDA e o PLR se favoreceram da aplicação do Boruta, enquanto o *XGBoost*, obteve melhores modelos sem a seleção de variáveis. Dessa forma, os resultados sugerem que a escolha do algoritmo deve considerar também a capacidade do modelo de lidar com variáveis irrelevantes. O uso do Boruta mostrou ser benéfico, especialmente para algoritmos mais sensíveis à escolha de variáveis para o desempenho geral.

No presente estudo, como a classe positiva foram as crianças terem experimentado o fenômeno, enquanto a negativa foi não terem alcançado, a especificidade se tornou uma medida de desempenho muito importante. Assim, a menor especificidade esteve presente em SL, como treino, e RP, como teste (variando de 47,6% a 61,8%), com o modelo *Xgboost* com Boruta apresentam o menor desempenho para essa métrica. Já a maior especificidade foi em RP, como treino, e SL, como teste (variando de 78,3% a 87,3%), no qual o LDA sem Boruta teve o maior valor obtido. Houve queda nas especificidades comparando sem e com adição do Boruta nos cenários de SL e RP juntas, SL treino e RP teste, e vice-versa, sugerindo que o selecionador de atributos possa interferir no desempenho dessa medida de avaliação quando ambas as bases estavam envolvidas.

As características mais relevantes para predizer a não ocorrência do *catch-up* estatutal nos primeiros 24 meses de idade foram: fumo materno, sexo da criança, altura da mãe, estatura para idade ao nascer, prematuridade, morbidades, peso da mãe antes da gestação, quantidade de consultas de PN, apgar no 1º minuto de vida, perímetro cefálico, peso do recém-nascido, peso para idade ao nascer.

Dessa forma, estudos considerando as variáveis supracitadas são interessantes para futuras projeções sobre *catch-up* estatutal. Sugere-se também analisar os 3 parâmetros de crescimento (peso, altura e perímetro cefálico), de forma conjunta. Melhorias nos modelos

para aumentar os desempenhos podem ser realizadas, visando também investigações e testes com o desfecho na forma numérica (problema de regressão).

Essas descobertas podem estimular pesquisadores da área de saúde infantil e outras partes interessadas, a aperfeiçoar os modelos e explorar outros algoritmos para predição de *catch-up* estatural. O intuito é que o conhecimento possa ser futuramente utilizado por profissionais que acompanham regularmente o crescimento infantil, desde a gravidez até a primeira infância, como enfermeiros e médicos de família e comunidade na APS, nas consultas de puericultura, bem como profissionais da Alta Complexidade em caso de gestação de risco e pediatras na Atenção Especializada.

REFERÊNCIAS

- ADEEBA, S et al. A Comparative Study of Machine Learning Algorithms for Predicting Weight Range of Neonate. International Conference on Decision Aid Sciences and Applications (DASA), Chiangrai, Thailand, p. 869-873, 2022.
- AKOMBI, B. J. et al. Stunting and severe stunting among children under-5 years in Nigeria: A multilevel analysis. *BMC pediatrics*, 17(1), 15, 2017.
- ALEXANDER, GR et al. A united states national reference for fetal growth. *Obstetrics & Gynecology*, vol 87, n 2, p 163- 168, 1996.
- ALBERNAZ, ALG; COUTO, MCV. A puericultura no SUS: o cuidado da criança na perspectiva da atenção integral à saúde. *Saúde Debate*, Rio De Janeiro, V. 46, N. Espel 5, . 236-248, 2022
- ANDERSON, C et al. Comparing predictive abilities of longitudinal child growth models. *Statistics in medicine*, 38(19), 3555–3570, 2019.
- ARAYESHGARI, M *et al.* Machine Learning-based Classifiers for the Prediction of Low Birth Weight. *Healthcare Informatics Research*, v. 29, n. 1, p. 54-63, 2023.
- BADILLO, S., et al. An Introduction to Machine Learning. *Clin. Pharmacol. Ther.*, vol. 107, p. 871-885, 2020.
- BARKER DJ. In utero programming of chronic disease. *Clin Sci (Lond)*. 1998 Aug; 95(2):115-28
- BAI, X, et al. Development and Evaluation of a Machine Learning Prediction Model for Small-for-Gestational-Age Births in Women Exposed to Radiation before Pregnancy. *J Pers Med.*, vol 12, n. 4, p 550, 2022.
- BARBIERI, MA. Fatores etiológicos do nascimento pré-termo e consequências dos fatores perinatais na saúde da criança: coortes de nascimentos em duas cidades brasileiras; Projeto de estudo de coorte. *Fundação de Amparo à Pesquisa do Estado de São Paulo*, Ribeirão Preto, 2008.
- BATISTA, R. F., et al. Factors associated with height catch-up and catch-down growth among schoolchildren. *PloS one*, vol 7, n. 3, e32903, 2012.
- BATISTA, A. F. M. ; CHIAVEGATTO FILHO, A.D.P. . Machine Learning aplicado à Saúde. In: Artur Ziviani; Natalia Castro Fernandes; Débora Christina Muchaluat Saade. (Org.). Livro de Minicursos. 1ed.Niterói, RJ: Sociedade Brasileira de Computação, 2019, p. 1-42.
- BEKELE, WT. Machine learning algorithms for predicting low birth weight in Ethiopia. *BMC Medical Informatics and Decision Making*, v. 22, n. 1, p. 232, 2022.
- BITEW, F. H., SPARKS, C. S., & NYARKO, S. H. Machine learning algorithms for predicting undernutrition among under-five children in Ethiopia. *Public health nutrition*, v. 25, n. 2, p. 269–280, 2022.

BLACK, M. M. et al. Early childhood development coming of age: science through the life course. *Lancet (London, England)*, 389(10064), 77–90, 2017

BLACK, MM; TRUDE, ACB; KOWALSKI, AJ. Growth in the first 1000 days lays the foundation for human capital formation. *Am J Clin Nutr.*, vol. 113, n. 4, p. 767-768, 2021.

BRASIL. Ministério da Saúde. Vigilância do desenvolvimento da criança. Caderneta de saúde da criança - menina. Brasília, 2021. Disponível em:
<https://saude.rs.gov.br/upload/arquivos/202201/26153336-caderneta-da-crianca- menina.pdf>

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção básica. Atenção ao pré-natal de baixo risco, 1. ed. rev. – Brasília : Editora do Ministério da Saúde, 318 p.: il. – (Cadernos de Atenção Básica, nº 32), 2013.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. Saúde da criança: crescimento e desenvolvimento, Brasília : Ministério da Saúde. 272 p.: il. – (Cadernos de Atenção Básica, nº 33), 2012.

Brasil. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. Saúde da criança : crescimento e desenvolvimento / Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de AtençãoBásica. – Brasília : Ministério da Saúde, 2012.

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). Disponível em:
https://bvsms.saude.gov.br/bvs/saudelegis/gm/2015/prt1130_05_08_2015.html

BRASIL. Ministério da Saúde. Leite materno passa por transformações de acordo com cada etapa de desenvolvimento do bebê. 2022. Disponível em:
<https://www.gov.br/saude/pt-br/assuntos/noticias/2022/agosto/leite-materno-passa-por-transfor macoes-de-acordo-com-cada-etapa-de-desenvolvimento-do-bebe>. Acesso em: 10 jan. 2025.

BRASIL. Ministerio da Saúde. PORTARIA Nº 1.058, de 4 DE JULHO DE 2005 Institui a disponibilização gratuita da “Caderneta de Saúde da Criança”, e dá outras providências. Disponível em:
https://bvsms.saude.gov.br/bvs/saudelegis/gm/2005/prt1058_04_07_2005.html

BRASIL. Ministério da Saúde. Secretaria de Atenção Primária à Saúde. Departamento de Ações Programáticas. Manual de gestação de alto risco [recurso eletrônico] / Ministério da Saúde, Secretaria de Atenção Primária à Saúde. Departamento de Ações Programáticas. – Brasília : Ministério da Saúde, 2022. 692 p. : il. Disponívem em:
https://bvsms.saude.gov.br/bvs/publicacoes/manual_gestacao_alto_risco.pdf

CAMERON, N; SCHELL, LM. Human Growth and Development. 3rd edition, United Kingdom: Elsevier, 2021.

CHAGAS, DC et al. Efeitos do ganho de peso gestacional e do aleitamento materno na retenção de peso pós-parto em mulheres da coorte BRISA, *Cad. Saúde Pública* 2017, vol. 33, n. 5, e00007916, 2017.

- CHIAVEGATTO FILHO, A; BATISTA, AFDM; SANTOS, HG. Data Leakage in Health Outcomes Prediction With Machine Learning. Comment on “Prediction of Incident Hypertension Within the Next Year: Prospective Study Using Statewide Electronic Health Records and Machine Learning”, *J Med Internet Res* 2021;23(2):e10969, 2021
- CHRISTIAN, P. et al. Risk of childhood undernutrition related to small-for-gestational age and preterm birth in low- and middle-income countries. *International Journal of Epidemiology*, v. 42, n. 5, p. 1340–1355, 2013.
- COLLINS GS 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.
- CUNHA, JPZ. Um estudo comparativo das técnicas de validação cruzada aplicadas a modelos mistos. Dissertação de Mestrado em Ciências. Instituto de Matemática e Estatística da Universidade de São Paulo, 2019.
- D'EÇA JUNIOR, A et al. Aleitamento materno complementado e fatores associados: coorte de nascimento brisa. *Rev. baiana saúde pública* ; 43(1)2019.
- FACELI, K. et al. Inteligência Artificial: Uma Abordagem de Aprendizado de Máquina. Editora: LTC, 2021.
- FAWCETT, T. An introduction to ROC analysis, *Pattern Recognition Letters*, vol 27, n 8, p. 861-874, 2006.
- FENTON TR, KIM JH. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. *BMC Pediatr*. 2013;13:59.
- FLEMING, TP. et al. Origins of lifetime health around the time of conception: causes and consequences. *Lancet*, vol 391, n 10132, p. 1842-1852, 2018.
- FREITAS et al. (2012). Crescimento do prematuro: revisão sobre as curvas de referência. *Revista de Pediatria SOPERJ*, 2012. Disponível em: http://revistadepediatriasoperj.org.br/detalhe_artigo.asp?id=548. Acesso em: 12 jan. 2025;
- GERON, A. (2019). Hands-on machine learning with Scikit-Learn, Keras and TensorFlow: concepts, tools, and techniques to build intelligent systems (2ª ed.). O'Reilly.
- GIBBONS, J. S.; GIBBONS, C.S. Machine learning in medicine: a practical introduction. *BMC Med Res Methodol*, vol. 19, n. 64, 2019.
- GONZÁLEZ-MUNIESA, P et al. Obesity. *Nature*, vol 3, n 17034, 2017.
- HARRISON, E. et al. Machine learning model demonstrates stunting at birth and systemic inflammatory biomarkers as predictors of subsequent infant growth – a four-year prospective study. *BMC Pediatr* 20, 498 (2020).
- HASTIE, T., TIBSHIRANI, R., & FRIEDMAN, J. The elements of statistical learning: data

mining, inference, and prediction. 2nd ed. New York, Springer. 2009.

HOCQUETTE, A; DUROX, M; WOOD, R et al. International versus national growth charts for identifying small and large-for-gestational age newborns: A population-based study in 15 European countries. *Lancet Reg Health Eur.*, vol 8, p. 100167, 2021.

HOKKEN-KOELEGA, A. C. S. et al. International Consensus Guideline on Small for Gestational Age (SGA): Etiology and Management from Infancy to Early Adulthood. *Endocr Rev.*, vol. 44, n. 3, p. 539-565, 2023.

ISLAM POLLOB, S. M. A. et al. Predicting risks of low birth weight in Bangladesh with machine learning. *Plos One*, v. 17, n. 5, p. e0267190, 2022.

JIANG, M et al. A case control study of risk factors and neonatal outcomes of preterm birth, *Taiwanese Journal of Obstetrics and Gynecology*, vol. 57, n. 6, p. 814-818, 2018

JUAN, J; YANG, H. Early life 1000 days: opportunities for preventing adult diseases. *Chin Med J (Engl)*, vol 135, n. 5, p. 516-518, 2021.

KALE, P. L. et al. Adequação do peso ao nascer para idade gestacional de acordo com a curva INTERGROWTH-21 st e fatores associados ao pequeno para idade gestacional. *Cad. saúde colet*, vol. 26, n. 4, 2018.

KARAKOCHUK, C. D.; WHITFIELD, K. C.; GREEN, T. J.; KRAEMER, K. The biology of the first 1,000 days. Taylor & Francis, p. 1-519, 2018.

KHAN, W; ZAKI, N; MASUD, MM; et al. Infant birth weight estimation and low birth weight classification in United Arab Emirates using machine learning algorithms. *Sci Rep.*, vol 12, n. 1, p 12110, 2022.

KRAMER, M. S., PLATT, R. W., WEN, S. W., JOSEPH, K. S., ALLEN, A., ABRAHAMOWICZ, M., BLONDEL, B., BRÉART, G., & Fetal/Infant Health Study Group of the Canadian Perinatal Surveillance System. A new and improved population- based Canadian reference for birth weight for gestational age. *Pediatrics*, 108(2), E35, 2001.

KUHLE, S; MAGUIRE, B; ZHANG, H. et al. Comparison of logistic regression with machine learning methods for the prediction of fetal growth abnormalities: a retrospective cohort study. *BMC Pregnancy Childbirth*, vol, 333, 2018.

Kuhn, Max (2008). "Building Predictive Models in R Using the caret Package." *Journal of Statistical Software*, 28(5), 1–26. doi:10.18637/jss.v028.i05.

Kursa, M. B. Boruta for those in a hurry. (2022). Disponível em: <https://cran.r-project.org/web/packages/Boruta/vignettes/inahurry.pdf>. Acesso em: 18 dez. 2024.

LEROY, JL; FRONTILLO, EA., Perspective: What Does Stunting Really Mean? A Critical Review of the Evidence, *Advances in Nutrition*, Volume 10, edição 2, 2019, páginas 196–204.

LI, J et al. Comparison of Different Machine Learning Approaches to Predict Small for Gestational Age Infants. *IEEE Transactions on Big Data*, vol. 6, no. 2, pp. 334-346, 1 June

2020.

LIMA, ACO. Acompanhamento do estado nutricional de crianças beneficiárias do programa bolsa família da 10ª região de saúde do Estado do Ceará. Monografia. Curso de Pós-Graduação Lato Sensu em Saúde da Família/Gestão em Saúde. Universidade da Integração Internacional da Lusofonia Afro-Brasileira. 2018. Limoeiro do Norte - CE. Disponível em: <https://repositorio.unilab.edu.br/jspui/bitstream/123456789/1536/1/ANA%20CARMEM%20DE%20OLIVEIRA%20LIMA%20TCC%20Artigo.pdf>

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.

LO VERCIO, L.; AMADOR, K.; BANNISTER, J.J.; et al. Supervised machine learning tools: a tutorial for clinicians. *Journal of Neural Engineering*, v.17, n.6, nov. 2020.

LU, Y. et al. Prediction of fetal weight at varying gestational age in the absence of ultrasound examination using ensemble learning. *Artificial Intelligence in Medicine*, vol 102, 101748, 2020.

MARTINEZ, AR. et al. Height and body-mass index trajectories of school-aged children and adolescents from 1985 to 2019 in 200 countries and territories: a pooled analysis of 2181 population-based studies with 65 million participants. *The lancet*, vol 396, n 10261, 2020, p 1511-1524.

MARQUES, G. Impacto do aleitamento materno exclusivo no sobrepeso e obesidade em filhos de mães com excesso de peso. 2020. 116 f. Dissertação de Mestrado - Faculdade de Medicina de Ribeirão Preto, Universidade de São Paulo, Ribeirão Preto, 2020.

MAULUD, DH; ABDULAZEEZ, AM. A Review on Linear Regression Comprehensive in Machine Learning. *Journal of Applied Science and Technology Trends*, Vol. 01, No.04, pp 140–147, 2020.

MELLO, R. F. de; PONTI, M. A. Machine Learning: A Practical Approach on the Statistical Learning Theory. Springer, 2018. ISBN 978-3-319-94988-8.

MILLWARD D. J. Nutrition, infection and stunting: the roles of deficiencies of individual nutrients and foods, and of inflammation, as determinants of reduced linear growth of children. *Nutrition research reviews*, 30(1), 50–72, 2017.

MILLER, SL; HUPPI, PS; MALLARD, C. The consequences of fetal growth restriction on brain structure and neurodevelopmental outcome, *The Journal of Physiology*, vol. 594, n. 4, p. 807–823, 2016.

MONARI, F; MENICHINI, D; SPANO-BASCIO, L; et al. A first trimester prediction model for large for gestational age infants: a preliminary study. *BMC Pregnancy Childbirth*, vol. 21, n. 1, p. 654, 2021.

MORAIS, AODS et al. Sintomas depressivos e de ansiedade maternos e prejuízos na relação mãe/filho em uma coorte pré-natal: uma abordagem com modelagem de equações estruturais.

Saúde Pública 2017; 33(6):e00032016

NARDOZZA LM, CAETANO AC, ZAMARIAN AC, MAZZOLA JB, SILVA CP, MARÇAL VM, et al. Fetal growth restriction: current knowledge. *Arch Gynecol Obstet.*, vol. 295, n. 5, p.1061-77, 2017.

Nass EMA, et al. Breastfeeding and diseases prevalent in the first two years of a child's life: a cross-sectional study. *Rev Bras Enferm.* 2022;75(6):e20210534. <https://doi.org/10.1590/0034-7167-2021-0534pt> 2022.

NEVES, FJ; FERREIRA, AA; WELCH, JR. Estado nutricional e fatores associados ao déficit estatural em crianças menores de cinco anos de comunidades remanescentes de quilombos do Nordeste brasileiro. *Cad. Saúde Pública*, vol 37, n, 7, p. e00060220, 2021.

NGUYEN, PH et al. Child Linear Growth During and After the First 1000 Days Is Positively Associated with Intellectual Functioning and Mental Health in School-Age Children in Vietnam. *J Nutr*; vol 151, n. 9, p. 2816-2824, 2021

NORO, G; GON, MCC. Epigenética, Cuidados Maternais e Vulnerabilidade ao Estresse: Conceitos Básicos e Aplicabilidade 2015. Psicologia Do Desenvolvimento. *Psicol. Reflex. Crit.* 28 (4) • Oct-Dec 2015

OBERMEYER Z, EMANUEL EJ. Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. *N Engl J Med* 2016;375:1216-1219. DOI: 10.1056/nejmp1606181.

OGHLI MG; SHABANZADEH, A; MORADI, S, et al. Automatic fetal biometry prediction using a novel deep convolutional network architecture. *Phys Med.* 2021;88:127-137. doi:10.1016/j.ejmp.2021.06.020

OKEN, E.; et al. A nearly continuous measure of birth weight for gestational age using a United States national reference. *BMC Pediatr* 3, 6 (2003).

OKEN, E. et al. When a birth cohort grows up: challenges and opportunities in longitudinal developmental origins of health and disease (DOHaD) research. *Journal of developmental origins of health and disease*, 14(2), 175–181, 2023.

ONG, K. K, et al. Association between postnatal catch-up growth and obesity in childhood: prospective cohort study. *BMJ*, v. 8, n. 320, p. 967-71, 2000.

ONIS M & BRANCA F. Childhood stunting: a global perspective. *Mat Child Nutr* 12, Suppl. 1, 12–26, 2016.

PALOMBO, CNT et al. Uso e preenchimento da caderneta de saúde da criança com foco no crescimento e desenvolvimento. *Rev Esc Enferm USP* 2014; 48(Esp):60-7.

PEREZ-MUÑOZ, C; CARRETERO-BRAVO, J; ORTEGA-MARTIN, E. et al. Interventions in the first 1000 days to prevent childhood obesity: a systematic review and quantitative content analysis. *BMC Public Health*, vol 22, n. 2367, 2022.

Pfob, A., Lu, SC. & Sidey-Gibbons, C. Aprendizado de máquina em medicina: uma introdução prática a técnicas de pré-processamento de dados, ajuste de hiperparâmetros e

comparação de modelos. *BMC Med Res Methodol* 22 , 282 (2022).
<https://doi.org/10.1186/s12874-022-01758-8>

PONTES, M. J. C. Algoritmo das Projeções Sucessivas para Seleção de Variáveis Espectrais em Problemas de Classificação. [S.l.]: Universidade Federal da Paraíba, 2009.

PRADER A. Catch-up growth. *Postgrad Med J*. 1978;54 Suppl 1:133-46. PMID: 724586.

RAHMAN, MS et al. Association of Low-Birth Weight with Malnutrition in Children under Five Years in Bangladesh: Do Mother's Education, Socio-Economic Status, and Birth Interval Matter? *PLoS One* . 2016;11(6):e0157814–e0157814. pmid:27355682

RAINIO, Oona; TEUHO, Jarmo; KLÉN, Riku. Evaluation metrics and statistical tests for machine learning. *Scientific Reports*, v. 14, n. 1, p. 6086, 2024.

RIBEIRO, SVO et al. Violência e sintomas de depressão na gestação e materna na coorte BRISA: uma abordagem com modelagem de equações estruturais. *Rev. Bras. Saúde Mater. Infant.*, Recife, vol. 19, n.1, p. 185-196 jan. / mar., 2019

RICHMOND, E.; ROGOL, A. Diagnostic approach to children and adolescents with short stature. *UpToDate*. 2020.

RORO, M; DERESSA, W; LINDTJORN. Influence of intrauterine factors on birth weight and on child linear growth in rural Ethiopia: A prospective cohort study. *PLoS ONE*, vol 17, n 8, p. e0272620, 2022.

ROSSI, CE; VASCONCELOS, FAG. Peso ao nascer e obesidade em crianças e adolescentes: uma revisão sistemática. *Rev. bras. epidemiol.* 13 (2) • Jun 2010 •
<https://doi.org/10.1590/S1415-790X2010000200007>

ROTEM, R et al. Trends of change in the individual contribution of risk factors for small for gestational age over more than 2 decades. *Arch Gynecol Obstet*, vol 302, n. 5, p. 1159-1166, 2020

Rugolo LM. Crescimento e desenvolvimento a longo prazo do prematuro extremo. *J Pediatr* (Rio J). 2005;81(1 Supl):S101-S110.

RYCKMAN, KK; BERBERICH, SL; DAGLE JM. Predicting gestational age using neonatal metabolic markers. *Am J Obstet Gynecol*, vol 214, n. 4, p. 515.e1-515.e13, 2016.

SANTOS EM et al. Avaliação do aleitamento materno em crianças até dois anos assistidas na atenção básica do Recife, Pernambuco, Brasil. *Ciênc. saúde colet.* 24 (3), Mar 2019

SARKER, I.H. Machine Learning: Algorithms, Real-World Applications and Research Directions. *SN Computer Science*, v.2, n.160, 2021.

SARTIKA, AN. et al, Prenatal and postnatal determinants of stunting at age 0–11 months: A cross-sectional study in Indonesia. *PLOS ONE* 16(7): e0254662, 2021.

SAW, SN; BISWAS, A; MATTAR, CNZ; LEE, HK; YAP, CH. Machine learning improves early prediction of small-for-gestational-age births and reveals nuchal fold thickness as unexpected predictor. *Prenat Diagn.*, vol 41, n 4, p 505-516, 2021.

SILVA, AAM et al. A protocol to identify non-classical risk factors for preterm births: the Brazilian Ribeirão Preto and São Luís prenatal cohort (BRISA). *Reproductive health* vol. 11,n .1, p. 79, 2014.

SILVA, DO. Otimização de hiper-parâmetros de algoritmos de Machine Learning aplicado no contexto de análise de risco de crédito. Especialização em Ciências de dados, Universidade Tecnológica Federal do Paraná 2022

SILVA, JP; PITTA, MGR; ALMEIDA, AR. Análise dos custos do cuidado aos recém-nascidos pré-termo internados em uma UTI Neonatal pública na cidade de Recife-PE, 2022. Disponível em <https://repositorio.ufpe.br/bitstream/123456789/48672/1/DISSERTA%C3%87%C3%83O%20Josineide%20Pereira%20da%20Silva.pdf>

SILVA FS et al, Foco e cobertura do programa Bolsa Família em crianças das coortes de nascimento BRISA, Ribeirão Preto (São Paulo) e São Luís (Maranhão), Brasil *Cad. Saúde Pública* 2019; 35(6):e00159718

SILVEIRA RC et al. Manual de seguimento do Recém-nascido de alto risco/ Departamento Científico de Neonatologia da Sociedade Brasileira de Pediatria. - 2.ed.- São Paulo: SBP, 2024. 142 f 2024. Disponível em: https://www.sbp.com.br/fileadmin/user_upload/24651c-ManSeguimento_RN_AltoRisco_MIOLO.pdf

SISVAN. Estado Nutricional dos Usuários da Atenção Básica - Norma Técnica da Vigilância Alimentar e Nutricional – SISVAN, 2004. Disponível em: http://tabnet.datasus.gov.br/cgi/SISVAN/CNV/notas_sisvan.html

SHARMA, D., SHASTRI, S., & SHARMA, P. Intrauterine Growth Restriction: Antenatal and Postnatal Aspects. *Clin Med Insights Pediatr*, v. 10, p. 67–83, 2016.

SOLIMAN, A; SANCTIS, V; ALAARANJ, N; AHAMED S; ALYAFEI F; HAMED N. Early and Long-term Consequences of Nutritional Stunting: From Childhood to Adulthood. *Acta Biomed*, vol 92, n. 1, p e2021168, 2021.

SPYRIDES, MHC et al. Amamentação e crescimento infantil: um estudo longitudinal em crianças do Rio de Janeiro, Brasil, 1999/2001 *Cad. Saúde Pública*, Rio de Janeiro, 21(3):756-766, mai-jun, 2005

SUN J et al. A data-driven approach to predict Small-for-Gestational-Age infants. *EEE 13th International Conference on Networking, Sensing, and Control (ICNSC)*, Mexico City,

Mexico, pp. 1-6, 2016.

TAO, J.; YUAN, Z.; SUN, L.; YU, K; ZHANG, Z. Fetal birthweight prediction with measured data by a temporal machine learning method. *BMC Medical Informatics and Decision Making*, vol. 21, n. 26, p. 2-10, 2021.

THOMSON, LC. et al Addressing child undernutrition: evidence review, 2014.
Disponível em: <https://www.dfat.gov.au/sites/default/files/evidence-review-child-nutrition.pdf>

TOHKA, Jussi; VAN GILS, Mark. Evaluation of machine learning algorithms for health and wellness applications: A tutorial. *Computers in Biology and Medicine*, v. 132, p. 104324, 2021.

UNPD. Programa das Nações Unidas para o Desenvolvimento. Atlas do desenvolvimento humano dos municípios; 2010. Disponível:
<http://www.undp.org.br/content/brazil/pt/home/idh0/atlas-do-desenvolvimento-umano/atlas-dos-municipios.html>.

VICTORA, Cesar G. et al. Association between breastfeeding and intelligence, educational attainment, and income at 30 years of age: a prospective birth cohort study from Brazil. *The Lancet Global Health*, 2015. Disponível em: [https://www.thelancet.com/journals/langlo/article/PIIS2214-109X\(15\)70002-1/fulltext](https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(15)70002-1/fulltext).

VILA, KLC; MACEDO, DJN; MELO, CYSV. Catch up growth até dois anos de idade de crianças nascidas pré-termo acompanhadas em follow up de método canguru de hospital de referência. TCC, 2017. Disponível em: <https://tcc.fps.edu.br/jspui/handle/fpsrepo/376>

VILLARES, JMM. Los mil primeros días de vida y la prevención de la enfermedad en el adulto. *Nutr Hosp*;33(Supl. 4):8-11, 2016.

VILLAR, J; GIULIANI, F; FENTON, TR; OHUMA, EO; ISMAIL, LC; KENNEDY, SH. INTERGROWTH-21st very preterm size at birth reference charts. *The Lancet*, vol. 387, n. 10021, p. 844-845, 2016.

VILLAR, J; et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. *Lancet*; vol 384, p. 857–68, 2014.

WADHWA, P. D et al. Stress, infection and preterm birth: a biobehavioural perspective. *Paediatr Perinat Epidemiol*. 2001;15(Suppl 2):17–29.

WADHWA, P. D. et al. Developmental origins of health and disease: brief history of the approach and current focus on epigenetic mechanisms. *Seminars in reproductive medicine*, 27(5), 358–368, 2009.

Wit JM, Boersma B. Catch-up growth: definition, mechanisms, and models. *J Pediatr Endocrinol Metab*. 2002 Dec;15 Suppl 5:1229-41. PMID: 12510974.

WONG, SC. et al. Growth and the growth hormone-insulin like growth factor 1 axis in children with chronic inflammation: current evidence, gaps in knowledge and future

directions. *Endocr Rev* 37, 62–110, 2016.

WHO. Multicentre Growth Reference Study Group. WHO Child Growth Standards based on length/height, weight and age. *Acta paediatrica* (Oslo, Norway : 1992), Supplement vol. 450, 76-85, 2006.

YANG, L; SHAMI, A. On hyperparameter optimization of machine learning algorithms: Theory and practice, *Neurocomputing*, vol 415, Pages 295-316, 2020

YU Z, HAN S, ZHU J, SUN X, JI C, GUO C. Pre-pregnancy body mass index in relation to infant birth weight and offspring overweight/obesity: a systematic review and meta-analysis. *PLOS ONE*, vol 8(4):61627, 2013.

ZOHDI, V. et al. Low Birth Weight due to Intrauterine Growth Restriction and/or Preterm Birth: Effects on Nephron Number and Long-Term Renal Health. *Int J Nephrol.*, 136942, 2012.

ANEXOS

O anexo A contém os questionários utilizados para extração das variáveis que possibilitaram o desenvolvimento do trabalho. O anexo B contém as Normas da revista a qual será submetido o artigo.

QUESTIONÁRIO DO 1º ANO MÃE

BLOCO J – SAÚDE DA MÃE E CONTRACEPÇÃO SATISFAÇÃO/SEXUALIDADE

1J. Em geral, como a Sra. considera sua saúde:

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

()

2J. A Sra. teve algum problema de saúde causado pelo parto, depois que <CRIANÇA> nasceu, até agora? (**Só aplicar para a mãe biológica**)

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

Passe para a questão 4J

()

3J. Se sim, qual (is) _____ fazer lista

()()

4J. Depois que <CRIANÇA> nasceu, a Sra. foi internada alguma vez?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

Passe para a questão 9J

()

Qual motivo da internação?

5J. Internação1 _____

88. ☐ Não se aplica

99. ☐ Não sabe

()()

6J. Quantos dias1 _____ dias

88. ☐ Não se aplica

99. ☐ Não sabe

()()

7J. Internação2 _____

88. ☐ Não se aplica

99. ☐ Não sabe

()()

8J. Quantos dias2 _____ dias

88. ☐ Não se aplica

99. ☐ Não sabe

()()

9J. A Sra. engravidou depois que <CRIANÇA> nasceu?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

Passe para a questão 11J

()

10J. O que aconteceu com essa gravidez?

1. ☐ Aborto
2. ☐ Nascimento
3. ☐ Está grávida atualmente
8. ☐ Não se aplica

()

11J. A Sra. está fazendo alguma coisa para não engravidar?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

Passe para a questão 13J

Passe para a questão 13J

()

12J. O que a Sra. está fazendo para não engravidar?

- 01. ☐ Pílula
- 02. ☐ Injetável trimestral
- 03. ☐ Coito interrompido
- 04. ☐ Preservativo
- 05. ☐ DIU
- 06. ☐ Laqueadura tubária
- 07. ☐ Tabela
- 08. ☐ Ele fez vasectomia
- 09. ☐ Não tem parceiro
- 10. ☐ Não menstruou
- 11. ☐ Não voltou a ter relações sexuais
- 12. ☐ Fez ligadura
- 13. ☐ Está amamentando
- 14. ☐ Outro _____
- 88. ☐ Não se aplica (não está fazendo nada)

() ()

13J. Foi oferecida alguma orientação na enfermaria do hospital onde você deu a luz para iniciar algum método de prevenção de gravidez?

- 1. ☐ Sim
- 2. ☐ Não **Passe para questão 15J**
- 9. ☐ Não sabe **Passe para questão 15J**

()

14J. Qual orientação para evitar gravidez a Sra. recebeu no hospital?

- 1. ☐ Procurar posto para ser prescrito medicação para evitar gravidez
- 2. ☐ Já iniciaram o método na enfermaria
- 3. ☐ Saiu com a receita para iniciar método após 30 dias do parto
- 4. ☐ Não recebeu orientação
- 5. ☐ Outro
- 9. ☐ Não sabe

()

Em relação ao parto responda:

15J. Pensando no que aconteceu na hora do parto e depois, a Sra. está satisfeita com o parto?

- 1. ☐ Sim
- 2. ☐ Não
- 9. ☐ Não sabe

()

16J. Com respeito ao seu parto, a Sra. sentiu que o atendimento, tratamento e orientações da equipe médica foram:

- 1. ☐ Muito bons
- 2. ☐ Bons
- 3. ☐ Nem bons, nem ruins
- 4. ☐ Ruins
- 5. ☐ Muito ruins
- 9. ☐ Não Sabe

()

17J. Durante o seu trabalho de parto e ou parto, a Sra. contou com a presença de acompanhante?

- 1. ☐ Sim
- 2. ☐ Não **Passe para questão 19J**
- 9. ☐ Não sabe **Passe para questão 19J**

()

18J. Se SIM, quem acompanhou seu trabalho de parto ? _____

()

19J. Com respeito ao acompanhante durante o trabalho de parto a Sra. acha que ter acompanhante é:

- 1. ☐ Muito bom
- 2. ☐ Bom
- 3. ☐ Nem bom, nem ruim
- 4. ☐ Ruim
- 5. ☐ Muito ruim
- 9. ☐ Não sabe

()

20J. Baseada na sua experiência, que tipo de parto a Sra. recomendaria para uma amiga?

- 1. ☐ Parto vaginal normal

()

2. ☐ Cesariana

9. ☐ Não sabe

21J. A Sra. reiniciou a sua atividade sexual apos o nascimento da criança?

1. ☐ Sim

2. ☐ Não

Passe para questão 24J

9. ☐ Não sabe

Passe para questão 24J

☐

22J. Com quanto tempo depois do nascimento retomou a atividade sexual?

1. ☐ Entre 0 a 14 dias após o parto

2. ☐ Entre 15 a 30 dias após o parto

3. ☐ Entre o primeiro e o terceiro mês

4. ☐ Entre o terceiro e o sexto mês

5. ☐ Entre o sexto e o nono mês

6. ☐ Após o nono mês

7. ☐ Ainda não retomou a atividade sexual

9. ☐ Não sabe

☐

23J. Com respeito a sua vida sexual após o parto, a Sra. considera que seu nível de prazer sexual nas relações é:

1. ☐ Muito bom

2. ☐ Bom

3. ☐ Nem bom, nem ruim

4. ☐ Ruim

5. ☐ Muito ruim

6. ☐ Não se aplica

9. ☐ Não sabe

☐

24J. Houve alguma mudança com os seus hábitos urinários após o nascimento da criança?

1. ☐ Sim

2. ☐ Não

Passe para questão 26J

9. ☐ Não sabe

Passe para questão 26J

☐

25J. Se SIM, qual foi a mudança ? _____ lista

☐☐

26J. Desde que seu bebê nasceu, a Sra. sente dor no pé da barriga?

1. ☐ Sim

2. ☐ Não

Passe para questão 30J

9. ☐ Não sabe

Passe para questão 30J

☐

27J. Há quanto tempo a Sra. apresenta a dor? _____ anos _____ meses

☐☐

28J. Qual o número de vezes que sentiu a dor na última semana?

1. ☐ Menos que 1x na semana

2. ☐ 1 a 2 xx na semana

3. ☐ 3 a 4 xx na semana

4. ☐ 5 a 7 xx na semana

9. ☐ não se aplica

☐

BLOCO L – HÁBITOS DE VIDA

1L. A Sra. fumou desde que <CRIANÇA> nasceu?

1. ☐ Sim

2. ☐ Não

Passe para a questão 4L

9. ☐ Não sabe

☐

2L. Quantos cigarros a Sra. fumava por dia? _____ cigarros / dia

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

3L. Do total, quantos cigarros a Sra. fumava por dia dentro de casa perto da criança? _____ cigarros / dia

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

4L. A Sra. tem o hábito de fumar?

1. ☐ Sim

2. ☐ Não **Passe para a questão 8L**

9. ☐ Não sabe

☐

5L. Se SIM, há quanto tempo? ____ anos e ____ meses

☐☐

6L. Quantos cigarros a Sra. fuma por dia? ____ cigarros / dia

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

7L. Do total, quantos cigarros a Sra. fuma por dia dentro de casa perto da criança? ____ cigarros / dia

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

8L. Das pessoas que moram na sua casa, alguém fuma?

1. ☐ Sim

2. ☐ Não

Passe para a questão 10L

8. ☐ Não se aplica **(Se a mãe mora sozinha)**

9. ☐ Não sabe

☐

9L. Quantos cigarros as pessoas fumam por dia? ____ cigarros / dia

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

10L. A Sra. tem o hábito de tomar cerveja?

1. ☐ Sim

2. ☐ Não

Passe para a questão 14L

9. ☐ Não sabe

☐

11L. Quantos dias por semana? ____

8. ☐ Não se aplica

9. ☐ Não sabe

☐

12L. Quanto toma por dia (número de vasilhas) ____

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

13L. 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

☐

14L. A Sra. tem o hábito de tomar vinho?

1. ☐ Sim

2. ☐ Não

Passe para a questão 18L

9. ☐ Não sabe

☐

15L. Quantos dias por semana? ____

8. ☐ Não se aplica

9. ☐ Não sabe

☐

16L. Quanto toma por dia (número de vasilhas) ____

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

17L. 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

☐

18L. A Sra. tem o hábito de tomar algum outro tipo de bebida como uísque, vodka, gim, rum, pinga, cachaça?

1. () Sim

2. () Não **Passe para a questão 22L**

9. () Não sabe

19L. Quantos dias por semana? _____

8. () Não se aplica

9. () Não sabe

20L. Quanto toma por dia (número de vasilhas) _____

88. () Não se aplica

99. () Não sabe

21L. 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

22L. Com que frequencia a sra. consome 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

BLOCO M – EXAME CLÍNICO DA MÃE E DADOS DO PAI DA CRIANÇA

1M. Altura do pai da <CRIANÇA> , m

2M. Peso do Pai da <CRIANÇA> _____, kg

3M. Circunferência da cintura (mãe) _____, cm

Colocar os dados da Tanita

4M. Peso _____, _____ g

5M. Altura _____, _____ cm

6M. Gordura corporal _____, %

7M. Água, %

8M. Gordura visceral _____ %

9M. Massa muscular

10M. Massa óssea

11M. Idade metabólica anos

QUESTIONÁRIO DO 1º ANO 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

3ª casela: F. Avaliação da mãe no 1º ano
A. Avaliação no 1º ano RN 1
B. Avaliação no 1º ano RN 2
C. Avaliação no 1º ano RN 3
D. Avaliação no 1º ano RN 4

4ª e 5ª caselas: QG. Questionário geral da criança
QM. Questionário de saúde da mulher
QP. Questionário da psiquiatria
SG. Sangue
SR. Soro
HM. Hemograma da criança
TB. Teste de Bayley
AO. Avaliação Odontológica
ID. Identificação
NT. Questionário nutricional

() () () () () () () ()

6ª à 9ª. caselas: número seqüencial para cada cidade (colocar o mesmo número seqüencial do nascimento)

2A. Cidade:

1. () Ribeirão Preto
2. () São Luís

()

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

() () () () () () () ()

Entrevistador (a) : _____

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

Nome completo da criança (não abreviar):

BLOCO B – DADOS DE CONTATO

1B. Qual o seu endereço completo? _____

Telefone residencial: ____-____-____ Outro telefone: ____-____-____ Celular: ____-____-____

2B. Para facilitar futuros contatos, a sra. poderia nos fornecer o nome, relação de parentesco ou amizade, endereço e telefone fixo ou celular de parentes ou pessoas próximas com quem a Sra. tem contato frequente?

Nome da pessoa: _____

Parentesco/Amizade: _____

Endereço: _____

Telefone residencial: ____-____-____ Outro telefone: ____-____-____ Celular: ____-____-____

Nome da pessoa: _____

Parentesco/Amizade: _____

Endereço: _____

Telefone residencial: ____-____-____ Outro telefone: ____-____-____ Celular: ____-____-____

3B. A Sra. poderia nos fornecer o endereço e o telefone do seu trabalho?

Endereço: _____

Telefone comercial: ____-____-____

Telefone comercial: ____-____-____

4B. Se a Sra. pretende mudar de cidade, poderia nos informar o nome, endereço e o telefone de contato de algum parente ou alguém que more próximo à sua nova residência?

Nome da pessoa: _____

Endereço: _____

Telefone residencial: ____-____-____ Outro telefone: ____-____-____ Celular: ____-____-____

BLOCO C – IDENTIFICAÇÃO E ALIMENTAÇÃO DA CRIANÇA

1C. Qual a cor de <CRIANÇA>?

1. ☐ Branca
2. ☐ Preta/negra
3. ☐ Parda/mulata/cabocla/morena
4. ☐ Amarelo/oriental
5. ☐ Indígena
9. ☐ Não sabe

☐

2C. <CRIANÇA> vai à escolinha ou creche?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

☐

3C. Se sim, desde que idade <CRIANÇA> frequenta a escolinha ou creche? _____ meses

88. ☐ Não se aplica
98. ☐ Não sabe

☐☐

4C. A criança ONTEM recebeu leite do peito?

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

☐

5C. Se **NÃO**, até que idade a criança mamou leite do peito?

_____ meses _____ dias

8888. ☐ Não se aplica (nunca mamou)
9999. ☐ Não sabe

☐☐☐☐

6C. Até que idade seu filho ficou em aleitamento materno exclusivo? (ler para a mãe: aleitamento materno exclusivo é só leite do peito, sem chá, água, outros leites, outras bebidas ou alimentos)

_____ meses _____ dias

8888. ☐ Não se aplica
9999. ☐ Não sabe

☐☐☐☐

7C. Considerando apenas os últimos três meses, a senhora tem o hábito de oferecer o peito para <CRIANÇA> depois que ele(a) já adormeceu à noite

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

☐

Quando a senhora inseriu estes alimentos ou bebidas na rotina alimentar de <CRIANÇA>?

8C. Leite (líquido ou pó) _____ meses _____ dias 8888. ☐ Nunca 9999. ☐ Não sabe

☐☐☐☐

9C. Leite tipo fórmula? _____ meses _____ dias 8888. ☐ Nunca 9999. ☐ Não sabe

☐☐☐☐

10C. Outros líquidos (chás, sucos) _____ meses _____ dias 8888. ☐ Nunca 9999. ☐ Não sabe

☐☐☐☐

11C. Semi-sólido ou sólido? _____ meses _____ dias 8888. ☐ Nunca 9999. ☐ Não sabe

☐☐☐☐

12C. A Sra. usa açúcar (ou mel, nescau, toddy ou algo doce) para adoçar alguns desses alimentos?

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

☐

13C.Considerando apenas os últimos três meses, <CRIANÇA> costuma usar mamadeira durante a noite, após já ter adormecido?

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

()

14C.Atualmente <CRIANÇA> usa chupeta?

1. ☐ Sim, usa atualmente
2. ☐ Não, parou de usar
3. ☐ Nunca usou
9. ☐ Não sabe

Passe para a questão 18C

()

15C.Qual era a idade da <CRIANÇA>, quando começou a usar chupeta?

____meses ____dias

8888. ☐ Não se aplica
9999. ☐ Não sabe

()

16C.Qual o tipo de chupeta <CRIANÇA> usava?

1. ☐ Ortodôntica
2. ☐ Convencional
8. ☐ Não se aplica
9. ☐ Não sabe

()

17C.Quando parou de usar chupeta?

____meses ____dias

8888. ☐ Não se aplica
9999. ☐ Não sabe

()()()()

18C.<CRIANÇA> chupa dedo?

1. ☐ Sim, chupa atualmente
2. ☐ Não, parou de chupar
3. ☐ Nunca chupou
9. ☐ Não sabe

Passe para a questão 1D

()

19C.Qual a era a idade de <CRIANÇA>, quando começou a chupar dedos?

____meses ____dias

8888. ☐ Não se aplica
9999. ☐ Não sabe

()

20C.Quando parou de chupar dedo?

____meses ____dias

8888. ☐ Não se aplica
9999. ☐ Não sabe

()

BLOCO D –SAÚDE DA CRIANÇA

Agora vou fazer algumas perguntas sobre como está a saúde de <CRIANÇA>

1D. Em geral, a Sra. considera a saúde de <CRIANÇA>:

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

()

2D. CRIANÇA> teve tosse desde <dia da semana> da semana passada?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe ☐

3D. <CRIANÇA> teve respiração difícil desde <dia da semana> da semana passada?

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

4D. Estava com cansaço ou falta de ar?

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

5D. Estava com o nariz entupido?

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

6D. Estava com ronqueira ou catarro?

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

7D. Teve febre?

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

8D. <CRIANÇA> teve diarreia desde <dia da semana> de duas semanas atrás?

1. ☐ Sim
2. ☐ Não **Passe para a questão 17D**
9. ☐ Não sabe ☐

9D. A Sra. deu para <CRIANÇA> algo para tratar a diarreia?

1. ☐ Sim
2. ☐ Não **Passe para a questão 17D**
8. ☐ Não se aplica **Passe para a questão 17D**
9. ☐ Não sabe ☐

Se SIM, o que a Sra. deu?

- | | | | | | |
|--------------------------|---------------------------------|---------------------------------|---|--------------------------------------|--------------------------|
| 10D. Soro pacotinho | 1. ☐ sim | 2. ☐ Não | 8. ☐ Não se aplica | 9. ☐ Não sabe | ☐ |
| 11D. Soro caseiro | 1. ☐ sim | 2. ☐ Não | 8. ☐ Não se aplica | 9. ☐ Não sabe | ☐ |
| 12D. Soro farmácia | 1. ☐ sim | 2. ☐ Não | 8. ☐ Não se aplica | 9. ☐ Não sabe | ☐ |
| 13D. Outra solução _____ | 1. ☐ sim | 2. ☐ Não | 8. ☐ Não se aplica | 9. ☐ Não sabe | ☐ |
| 14D. Água | 1. ☐ sim | 2. ☐ Não | 8. ☐ Não se aplica | 9. ☐ Não sabe | ☐ |
| 15D. Chá | 1. ☐ sim | 2. ☐ Não | 8. ☐ Não se aplica | 9. ☐ Não sabe | ☐ |
| 16D. Remédio | 1. ☐ sim | 2. ☐ Não | 8. ☐ Não se aplica | 9. ☐ Não sabe | ☐ |

17D. Desde que <CRIANÇA> nasceu alguma vez teve diarreia com sangue?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe ☐

18D.<CRIANÇA> teve algum outro problema de saúde desde <dia da semana> de duas semanas atrás?

1. ☐ Sim

2. ☐ Não **Passe para a questão 22D**

9. ☐ Não sabe

☐

Caso tenha apresentado problema(s), nos diga qual(is) foi (foram)

19D.Problema 1 _____

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

20D.Problema 2 _____

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

21D.Problema 3 _____

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

22D.<Criança > recebeu sulfato ferroso ou vitamina com ferro desde que nasceu?

1. ☐ Sim

2. ☐ Não **Passe para a questão 26D**

9. ☐ Não sabe

☐

23D.Qual a idade do início__ __ (meses)

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

24D.Qual a idade do término __ __ (meses)

15. ☐ Ainda usa

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

25D.Quantas gotas ou ml foram dadas por vez? __ __ __ (colocar em gotas 1ml=20gts)

☐☐☐

26D.<Criança > recebeu algum remédio desde <dia da semana> de duas semanas atrás, inclusive vitamina ou remédio para febre?

1. ☐ Sim

2. ☐ Não **Passe para a questão 30D**

9. ☐ Não sabe

☐

Caso tenha tomado, qual o nome do remédio?

27D.Remédio1_____

77. ☐ Ainda usa

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

28D.Remédio2_____

77. ☐ Ainda usa

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

29D.Remédio3_____

77. ☐ Ainda usa

88. ☐ Não se aplica

99. ☐ Não sabe

☐☐

30D.Seu bebê teve chiado no peito, bronquiolite, bronquite ou sibilância alguma vez na vida?

1. ☐ Sim

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

31D.Quantos episódios de chiado no peito (bronquite ou sibilância) ele já teve?

1. ☐ Menos de 3 episódios
2. ☐ De 3 a 6 episódios
3. ☐ Mais de 6 episódios
8. ☐ Não se aplica ☐
9. ☐ Não sabe ☐

32D.<Criança > recebeu remédio por nebulização ou inalação tipo “bombinha” ?

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

33D.Se sim, qual _____

88. ☐ Não se aplica ☐
99. ☐ Não sabe ☐

34D.<Criança > recebeu remédio na boca?

1. ☐ Sim
2. ☐ Não **Passe para a questão 36D**
8. ☐ Não se aplica
9. ☐ Não sabe ☐

35D.Se sim, qual _____

88. ☐ Não se aplica ☐
99. ☐ Não sabe ☐

36D.Desde que <Criança > nasceu o chiado no peito (bronquite ou sibilância) do seu bebê foi tão intenso a ponto de ser necessário levá-lo a um serviço de emergência (Hospital, Clínica ou Posto de Saúde)?

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

37D.Algum médico já lhe disse alguma vez que seu bebê tem asma?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe ☐

38D.Seu bebê tem pai, mãe ou irmão/irmã com asma?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe ☐

39D. Seu bebê tem pai, mãe ou irmão/irmã com alergia no nariz ou rinite alérgica?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe ☐

40D. Seu bebê tem pai, mãe ou irmão/irmã com dermatite atópica **ou** eczema (eczema; alergia de pele caracterizada por erupção na pele com coceira intensa, que vai e volta, em qualquer área do corpo, exceto ao redor dos olhos e nariz, e região da fralda)?

1. ☐ Sim ☐

2. ☐ Não
9. ☐ Não sabe

41D.Desde que <Criança >nasceu, algum médico já diagnosticou dermatite atópica (eczema; alergia de pele caracterizada por erupção na pele com coceira intensa, que vai e volta, em qualquer área do corpo, exceto ao redor dos olhos e nariz, e região da fralda)

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

☐

42D. O seu bebe teve episódios breves de resfriado, com nariz escorrendo, espirros, obstrução nasal, tosse leve, com ou sem febre baixa, nos primeiros 3 meses de vida?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

☐

43D. Algum médico já lhe disse alguma vez que seu bebê tem rinite alérgica?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

☐

44D. A Sra. tem ar condicionado em casa?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

☐

45D.Desde que seu filho nasceu a Sra. teve ou tem algum animal de estimação (cachorro, gato, passarinho, coelho) em sua casa?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

☐

46D.A Sra. tem carpete em sua casa?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

☐

47D.Existe mofo (bolor) ou manchas de umidade em sua casa?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

☐

Seu bebê tem ou teve alergia/reação/problema quando comeu ou bebeu algum desses alimentos?

- | | | | |
|-------------------|---------------------------------|---------------------------------|--------------------------------------|
| 48D.Leite de vaca | 1. ☐ sim | 2. ☐ Não | 9. ☐ Não sabe |
| 49D.Trigo | 1. ☐ sim | 2. ☐ Não | 9. ☐ Não sabe |
| 50D.Soja | 1. ☐ sim | 2. ☐ Não | 9. ☐ Não sabe |
| 51D.Amendoim | 1. ☐ sim | 2. ☐ Não | 9. ☐ Não sabe |
| 52D.Peixe | 1. ☐ sim | 2. ☐ Não | 9. ☐ Não sabe |
| 53D.Ovo | 1. ☐ sim | 2. ☐ Não | 9. ☐ Não sabe |
| 54D.Outro_____ | 1. ☐ sim | 2. ☐ Não | 9. ☐ Não sabe |

☐

☐

☐

☐

☐

☐

☐

55D.Algum médico já lhe disse que o seu bebê tem alergia a algum alimento?

1. ☐ Sim
2. ☐ Não **Passe para a questão 57D**
9. ☐ Não sabe

☐

56D.Se sim, qual alimento?_____

88. ☐ Não se aplica
99. ☐ Não sabe

☐☐

57D.<CRIANÇA> já teve pneumonia alguma vez? ()

1. () Sim
2. () Não
9. () Não sabe

58D.Desde que nasceu, <CRIANÇA> teve infecção urinária? ()

1. () Sim
2. () Não
9. () Não sabe

59D.<CRIANÇA> foi internada alguma vez desde o nascimento até agora? ()

1. () Sim
2. () Não **Passe para a questão 64D**
9. () Não sabe **Passe para a questão 64D**

Quantas vezes <CRIANÇA> foi internada desde o nascimento até agora? (colocar as 2 últimas internações)

60D. Idade1 da internação (meses)____ 88. () Não se aplica 99. () Não sabe ()()

61D. Causa da internação1 _____ ()()

Idade2 da internação (meses)____ 88. () Não se aplica 99. () ()()

62D. Causa da internação2 _____ ()()

63D.Nos primeiros três meses de vida, <CRIANÇA> foi entubada por mais de 24 horas?

1. () Sim
2. () Não
9. () Não sabe ()

64D.<CRIANÇA> tem cartão de vacinas?

1. () Sim, visto
2. () Sim, não visto
3. () Tinha, mas perdeu
4. () Nunca teve
5. () Outro _____
9. () Não sabe ()

Quantas doses de vacina já recebeu? (verificar o cartão da criança e anotar com X as doses aplicadas)

		1ª dose	2ª dose	3ª dose	R□forço	
65D.BCG	9. Não sabe					()
66D.VHB (hepatite B)	999. Não sabe					()()()
67D.VORH (rotavírus)	99. Não sabe					()()
68D.VOP (oral contra pólio)	9999. Não sabe					()()()()
69D.Tetavalente (DTP +Hib)	999. Não sabe					()()()
70D.SRC (tríplice viral)	9. Não sabe					()
71D.Febre amarela	9. Não sabe					()
72D.DTP (tríplice bacteriana)	9999. Não sabe					()()()()
73D.Hemófilo sabe	9999. Não					()()()()

74D. Antipneumocócica	9999. Não sabe					() () () ()
75D. Antimeningocócica	999. Não sabe					() () ()
76D. Influenza (gripe)	999. Não sabe					() () ()
77D. Hepatite A	9. Não sabe					()
78D. Varicela (catapora)	9. Não sabe					()

79D. Onde <CRIANÇA> foi vacinada na maioria das vezes?

1. () Posto de saúde
 2. () Consultório ou clínica particular
 8. () Não se aplica
 9. () Não sabe ()

80D.<CRIANÇA> já foi ao médico ou posto de saúde ou hospital para consultar por doença?

1. () Sim
 2. () Não
 9. () Não sabe ()

81D.<CRIANÇA> já foi ao médico ou posto de saúde ou hospital só para vacinar, pesar ou acompanhar o crescimento e desenvolvimento?

1. () Sim
 2. () Não
 9. () Não sabe ()

82D.Alguma vez a Sra. já levou <CRIANÇA> para consultar no Pronto Socorro municipal ou em algum Pronto Atendimento?

1. () Sim
 2. () Não **Passe para a questão 85D**
 9. () Não sabe **Passe para a questão 85D** ()

83D. Por quê? (última consulta) _____

88. () Não se aplica
 99. ☐ Não sabe () ()

84D.Alguma vez a Sra. tentou levar <CRIANÇA> para consultar, hospitalizar ou vacinar e não conseguiu?

1. () Sim
 2. () Não **Passe para a questão 87D**
 9. () Não sabe **Passe para a questão 87D** ()

85D. Por quê? (última consulta) _____

88. () Não se aplica
 99. () Não sabe () ()

86D. <CRIANÇA> teve algum tipo de acidente nos últimos 15 dias?

1. () Sim
 2. () Não **Passe para a questão 89D**
 9. () Não sabe **Passe para a questão 89D** ()

87D.Qual (is)? _____

88. () Não se aplica
 99. () Não sabe () () ()

88D. <CRIANÇA> tem alguma doença, problema físico ou retardo?

1. () Sim
 2. () Não **Passe para a questão 91D**
 9. () Não sabe **Passe para a questão 91D** ()

89D.Qual? _____

88. () Não se aplica

99. () Não sabe

()()

Em relação ao desenvolvimento de <CRIANÇA> , o que ela já está sendo capaz de fazer?

- | | | | | |
|--|-----------|-----------|----------------|-----|
| 90D.Apresenta sorriso social | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| 91D.Firma pescoço | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| 92D.Iniciou lalação | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| 93D.Consegue sentar com apoio | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| 94D.Senta sem apoio | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| 95D.Já fala algumas palavras | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| 96D.Consegue dar passos com auxílio | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| 97D.Andar sem auxílio | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| 98D.É capaz de correr e pular com facilidade | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |
| Fala sem dificuldade ou gagueira | 1. ()sim | 2. ()Não | 9. ()Não sabe | () |

99D.<CRIANÇA> Apresentou crise convulsiva ou quadro sugestivo de perda de consciência súbita (desmaio) sem causa aparente?

1. () Sim, e associada a quadro febril
2. () Sim, e não associada a quadro febril
3. () Não
9. () Não sabe

()

100D. Existe história de epilepsia na família (irmãos, pais, tios, avós)?

1. () Sim, e não associada a quadro febril
2. () Não
9. () Não sabe

()

101D. Existe parentesco (consangüinidade) entre o casal (pais da criança)?

1. () Sim
2. () Não
9. () Não sabe

Passe para a questão 105D

()

102D. Se sim, em que grau?

1. () Primeiro (pai, mãe e filho)
2. () Segundo (irmão e irmã, neto/neta, avô/avó)
3. () Terceiro (primo e prima, tio e tia)
8. () Outros
9. () Não sabe

()

103D. Existe história de câncer na família (irmãos, pais, tios, avós)?

1. () Sim
2. () Não
9. () Não sabe

Passe para a questão 108D

Passe para a questão 108D

()

104D. Se sim, quem?

1. () Irmãos
2. () Pais
3. () Tios/tias / primos/primas
4. () Avós/avôs
5. () Outros _____
8. () Não se aplica
9. () Não sabe

()

105D. Se sim, qual o tipo de câncer existente na família?

1. () Glândulas adrenais (suprarenais)

2. ☐ Mama
3. ☐ Ossos
4. ☐ Músculo
5. ☐ Sistema nervoso central (cérebro)
6. ☐ Outros: _____
8. ☐ Não se aplica
9. ☐ Não sabe

()

106D. Existe história na família de desenvolvimento de mamas antes dos 8 anos de idade?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

()

107D. Existe história na família de menstruação antes dos 9 anos?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

()

BLOCO E – GASTOS COM SAÚDE

1E. <CRIANÇA> tem plano de saúde?

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

()

2E. Se sim, qual o Plano? _____

88. ☐ Não se aplica
99. ☐ Não sabe

()()

3E. O plano de saúde cobre consultas médicas?

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

()

4E. O plano de saúde cobre exames (laboratório, RX)?

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

()

5E. O plano de saúde cobre internações em hospital?

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

()

6E. Seu plano cobre tratamento odontológico?

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

()

7E. Além da mensalidade, este plano cobra algum valor pelas consultas ou exames?

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

()

8E. A Sra. gastou algum dinheiro com remédios para <CRIANÇA>?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

()

9E. A Sra. gastou algum dinheiro com consultas médicas para <CRIANÇA>?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

()

10E. A Sra. gastou algum dinheiro com exames complementares ou raio X para <CRIANÇA>?

1. ☐ Sim
2. ☐ Não
9. ☐ Não sabe

()

11E. A Sra. gastou algum dinheiro com outras coisas relacionadas à saúde de <CRIANÇA>? (*enfermeira, óculos, fisioterapia, dentista*)

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

()

12E. Se sim, qual (is) _____

88. ☐ Não se aplica
99. ☐ Não sabe

()

BLOCO F - DADOS SOCIOECONÔMICOS E DEMOGRÁFICOS

1F. A família recebe bolsa família ou algum outro benefício de transferência de renda do governo?

1. ☐ Sim
2. ☐ Não **Passe para a questão 3F**
9. ☐ Não sabe

()

2F. Há quanto tempo recebe o benefício do governo? ____ meses

88. ☐ Não se aplica
99. ☐ Não sabe

()()

3F. Sua família é cadastrada no Programa de Saúde da Família (PSF)?

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

()

A Sra. recebeu visita do PSF no último mês pelo:

- | | | | | | |
|--------------------------------------|---------------------------------|---------------------------------|--------------------------------------|---|-----|
| 4F. Agente comunitario | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica | () |
| 5F. Medico | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica | () |
| 6F. Enfermeiro | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica | () |
| 7F. Auxiliar | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica | () |
| 8F. Dentista ou Auxiliar de dentista | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica | () |

9F. 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

()

10F. 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

() ()

11F. A Sra. mora atualmente com o marido ou companheiro?

1. () Sim

2. () Não

9. () Não sabe

()

12F. Sra. exerce alguma atividade remunerada dentro ou fora de casa?

1. () Sim

2. () Não **Passe para a questão 15F**

9. () Não sabe

()

13F. Qual a sua ocupação (o que faz atualmente no trabalho)?

88. () Não se aplica (não trabalha fora de casa)

99. () Não sabe

() ()

14F. 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

()

15F. Quem é a pessoa da família com maior renda atualmente? (considerar chefe da família aquele de maior renda). Considerar o parentesco em relação à criança (mãe DA CRIANÇA, pai DA CRIANÇA, etc.)

01. () Mãe

02. () Pai

03. () Avó

04. () Avô

05. () Madrasta

06. () Padrasto

07. () Tia

08. () Tio

09. () Irmã

10. () Irmão

11. () Outro _____

99. () Não sabe

() ()

16F. Qual a idade da pessoa da família com maior renda (anos completos)? ___

99. () () Não sabe

() ()

17F. Essa pessoa sabe ler e escrever?

1. () Sim

2. () Não

9. () Não sabe

()

18F. Essa pessoa frequenta ou frequentou escola?

1. () Sim

2. () Não **Passe para a questão 21F**

9. () Não sabe

()

19F. 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
- 5. ☐ Superior graduação completo
- 8. ☐ Não se aplica
- 9. ☐ Não sabe

Passe para a questão 21F

Passe para a questão 21F

☐

20F. Qual a série que essa pessoa frequenta ou até que série frequentou?

- 01. ☐ Primeira
- 02. ☐ Segunda
- 03. ☐ Terceira
- 04. ☐ Quarta
- 05. ☐ Quinta
- 06. ☐ Sexta
- 07. ☐ Sétima
- 08. ☐ Oitava
- 88. ☐ Não se aplica
- 99. ☐ Não sabe

☐☐

Passe par a questão 23F - quando a própria mãe da criança for a pessoa de maior renda

21F. 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

☐☐

22F. 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

☐

23F. 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\$ _____. _____, _____

88888. ☐ Não quis informar

99999. ☐ Não sabe

☐☐☐☐☐☐☐☐

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

Quantidade de itens				
0	1	2	3	4 ou mais
0	1	2	3	4
0	1	2	3	4
0	1	2	3	4
0	4	5	6	7
0	4	7	9	9
0	3	4	4	4
0	2	2	2	2
0	2	2	2	2
0	4	4	4	4
0	2	2	2	2

24F. Televisão em cores

()

25F. Rádio

()

26F. Banheiro

()

27F. Automóvel

()

28F. Empregada mensalista

()

29F. Máquina de lavar

()

30F. Videocassete ou DVD

()

31F. Geladeira

()

32F. Freezer (aparelho independente ou parte da geladeira duplex)

()

33F. Grau de Instrução da pessoa com maior renda (circule a resposta)

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

()

BLOCO G – DADOS DO COMPANHEIRO

Caso o companheiro seja a pessoa com maior renda ou não more junto com a mãe do RN, passe para o bloco H

1G. Qual a idade do companheiro atual? __ __

88. () Não se aplica – não tem companheiro atual

99. () Não sabe

()()

2G. O seu companheiro sabe ler e escrever?

1. () Sim

2. () Não

8. () Não se aplica

9. () Não sabe

()

3G. O seu companheiro frequenta ou frequentou escola?

1. () Sim

2. () Não

8. () Não se aplica

9. () Não sabe

()

Passe para a questão 6G

4G. 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

5. () Superior graduação completo

8. () Não se aplica

9. () Não sabe

()

Passe para a questão 6G

Passe para a questão 6G

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

01. ☐ Primeira
02. ☐ Segunda
03. ☐ Terceira
04. ☐ Quarta
05. ☐ Quinta
06. ☐ Sexta
07. ☐ Sétima
08. ☐ Oitava
88. ☐ Não se aplica
99. ☐ Não sabe

() ()

6G. O seu companheiro está trabalhando no momento?

1. ☐ Sim
2. ☐ Não **Passe para bloco H**
8. ☐ Não se aplica
9. ☐ Não sabe

()

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

88. ☐ Não se aplica
99. ☐ Não sabe

() ()

8G. 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

()

BLOCO H – ODONTOLOGIA

1H. <CRIANÇA> recebe algum tipo de limpeza (higiene) da genviva, lingua ou dentes ?

1. ☐ Sim
2. ☐ Não **Passe para a questão 15H**
9. ☐ Não sabe

()

2H. Quando a Sra. começou a limpar a boca ou dentes de <CRIANÇA>?

1. ☐ Desde antes de nascerem os dentes com fralda ou gaze
2. ☐ Apenas quando nasceu nasceu o primeiro dente
3. ☐ Algum tempo depois de nascerem os dentes quando <CRIANÇA> deixou escovar
3. ☐ Outras: _____
8. ☐ Não se aplica
9. ☐ Não sabe

()

Quem cuida da higiene da boca e dentes de <CRIANÇA>?

()

- | | | | | |
|-----------------------------|---------------------------------|---------------------------------|--------------------------------------|---|
| 3H. A mãe | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica |
| 4H. O pai | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica |
| 5H. Um irmão menor de idade | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica |
| 6H. A empregada | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica |
| 7H. A professora da creche | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica |
| 8H. Outros _____ | 1. ☐ Sim | 2. ☐ Não | 9. ☐ Não sabe | 8. ☐ Não se aplica |

()

9H. O que a Sra. utiliza atualmente para limpar os dentes ou genviva de <CRIANÇA>?

- 1. ☐ Gase ou fralda apenas
- 2. ☐ Escova de dentes
- 3. ☐ Gase ou fralda e escova de dentes
- 8. ☐ Não se aplica
- 9. ☐ Não sabe

()

10H. CRIANÇA> tinha que idade quando começou a usar pasta de dente? ____meses ____ dias

7777. ☐ Ainda não usa

Passe para a questão 13H

8888. ☐ Não se aplica

Passe para a questão 13H

9999. ☐ Não sabe

Passe para a questão 13H

()()()()

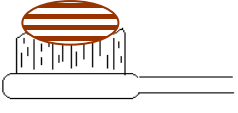
11H. Qual a pasta de dente que <CRIANÇA> mais usa? _____

88. ☐ Não se aplica

99. ☐ Não sabe

()()

12H. Escolha o desenho que mais se aproxima da quantidade de pasta de dente que é utilizada na escovação dos dentes de <CRIANÇA>:

- 1. ☐ 
- 2. ☐ 
- 3. ☐ 

8. ☐ Não se aplica

9. ☐ Não sabe

()

13H. Considerando apenas o último mês, com que frequência os dentes de <CRIANÇA> são limpos ou escovados?

- 1. ☐ Mais de uma vez por dia
- 2. ☐ Pelo menos uma vez por dia
- 3. ☐ Duas a tres vezes por semana
- 8. ☐ Não se aplica
- 9. ☐ Não sabe

()

14H. Os dentes de <CRIANÇA> são escovados ou limpos à noite depois do último alimento (ou amamentação no peito ou mamadeira) antes de colocá-lo(a) para dormir?

- 1. ☐ Sim, sempre
- 2. ☐ As vezes
- 3. ☐ Não
- 8. ☐ Não se aplica
- 9. ☐ Não sabe

()

15H. <CRIANÇA> já foi levada para consultar com dentista?

- 1. ☐ Sim
- 2. ☐ Não **Passe para a questão 17H**
- 9. ☐ Não sabe

()

16H. Qual o motivo principal de <CRIANÇA> ter sido levado ao dentista?

1. ☐ para aplicação de Flúor e orientação sobre prevenção
2. ☐ devido a presença de cárie dentária
3. ☐ para verificar porque os dentes não estavam nascendo
4. ☐ porque bateu a boca e machucou/quebrou dente
5. ☐ outro motivo _____
8. ☐ Não se aplica
9. ☐ Não sabe

()

17H. Como você classificaria hoje a saúde dos dentes de <CRIANÇA>?

1. ☐ Excelente
2. ☐ Boa
3. ☐ Razoável
4. ☐ Ruim
5. ☐ Pessima
8. ☐ Não se aplica (não tem dentes)
9. ☐ Não sabe

()

BLOCO I – EXAME CLÍNICO DA CRIANÇA

1I. Peso _____ g

()()()()()

2I. Comprimento _____ cm

()()()

3I. Perímetro cefálico _____ cm

()()()

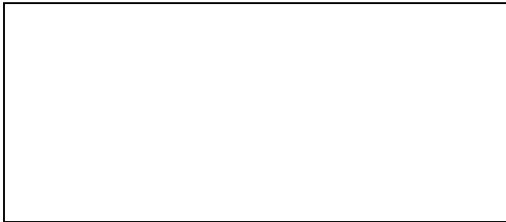
4I. Circunferência da abdominal _____ cm

()()()

5I. Coleta de sangue da criança?

1. ☐ Sim
2. ☐ Não

()



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

CIDADEP

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 sabe

IDADEG

7A. Data da ultrassonografia do recrutamento

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

DATUSR

99999999. ☐ Não sabe

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

IDGESTUSR

99. ☐ Não sabe

9A. Idade gestacional atual (pela USG):

__ semanas

IDGESTAT

99. ☐ Não sabe

BLOCO B – DADOS DE CONTATO

1B. Qual o seu endereço completo? _____

Telefone residencial: ____-____-____ Outro telefone: ____-____-____ celular: ____-____-____

2B. Para facilitar futuros contatos, a sra. poderia nos fornecer o nome, relação de parentesco ou amizade, endereço e telefone fixo ou celular de parentes ou pessoas próximas com quem a sra. tem contato frequente?

Nome da pessoa: _____

Parentesco/Amizade: _____

Endereço: _____

Telefone residencial: ____-____-____ Telefone comercial: ____-____-____ celular: ____-____-____

Nome da pessoa: _____

Parentesco/Amizade: _____

Endereço: _____

Telefone residencial: ____-____-____ Telefone comercial: ____-____-____ celular: ____-____-____

3B. A sra. poderia nos fornecer o endereço e o telefone do seu trabalho?

Endereço: _____

Telefone comercial: ____-____-____ Telefone comercial: ____-____-____

4B. Se a sra. pretende mudar de cidade, poderia nos informar o nome, endereço e o telefone de contato de algum parente ou alguém que more próximo à sua nova residência?

Nome da pessoa: _____

Endereço: _____

Telefone residencial: ____-____-____ Telefone comercial: ____-____-____ celular: ____-____-____

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

☐☐

27C. 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

RELCHFEF

☐

28C. 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

RENDAFP

☐☐☐☐☐

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

	Quantidade de itens				
	0	1	2	3	4 ou mais

29C. Televisão em cores	0	1	2	3	4
-------------------------	---	---	---	---	---

TELEVISAOP

☐

30C. Rádio	0	1	2	3	4
------------	---	---	---	---	---

RADIOP

☐

31C. Banheiro	0	4	5	6	7
---------------	---	---	---	---	---

BANHEIROP

☐

32C. Automóvel	0	4	7	9	9
----------------	---	---	---	---	---

AUTOMOVLP

☐

33C. Empregada mensalista	0	3	4	4	4
---------------------------	---	---	---	---	---

EMPREGADAP

☐

34C. Máquina de lavar	0	2	2	2	2
-----------------------	---	---	---	---	---

MAQLAVARP

☐

35C. Videocassete ou DVD	0	2	2	2	2
--------------------------	---	---	---	---	---

DVDP

☐

36C. Geladeira	0	4	4	4	4
----------------	---	---	---	---	---

GELADEIRAP

☐

37C. Freezer (aparelho independente ou parte da geladeira duplex)	0	2	2	2	2
---	---	---	---	---	---

FREEZERP

☐

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

DIATESTG

☐

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 cafezinho
 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. ☐ Tabelinha/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

MESANOUMP

☐☐☐☐☐☐

999999. ☐ Não sabe

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

DT1CPNP

☐☐☐☐☐☐☐☐

99999999. ☐ Não sabe

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

IREM2

☐☐

15H. Mês da gestação de término _
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 frequência?

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 frequência?

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 J – QUESTIONÁRIO INTERNACIONAL DE ATIVIDADE FÍSICA (IPAQ-versão curta)

Nós estamos interessados em saber que tipos de atividade física as pessoas fazem como parte do seu dia a dia. As perguntas estão relacionadas ao tempo que você gastou fazendo atividade física na **ÚLTIMA SEMANA**. As perguntas incluem as atividades que você fez no trabalho, para ir de um lugar a outro, por lazer, por esporte, por exercício ou como parte das suas atividades em casa ou no jardim. Suas respostas são **MUITO** importantes. Por favor, responda cada questão mesmo que considere que não seja ativo. Obrigado pela sua participação!

Para responder às questões lembre que:

- Atividades físicas **VIGOROSAS OU PESADAS** são aquelas que precisam de um grande esforço físico e que fazem respirar **MUITO** mais forte que o normal
- Atividades físicas **MODERADAS** são aquelas que precisam de algum esforço físico e que fazem respirar **UM POUCO** mais forte que o normal

Para responder às perguntas pense somente nas atividades que você realiza **por pelo menos 10 minutos contínuos** de cada vez:

- 1J. Em quantos dias da última semana você **CAMINHOU** por **pelo menos 10 minutos contínuos** em casa ou no trabalho, como forma de transporte para ir de um lugar para outro, por lazer, por prazer ou como forma de exercício?

0. ☐ Nenhum _____ dias por **SEMANA**

IPAQ1

☐

- 2J. Nos dias em que você caminhou por **pelo menos 10 minutos contínuos**, quanto tempo no total você gastou caminhando **por dia**?

Horas __ Minutos __

IPAQ2

☐☐☐☐

- 3J. Em quantos dias da última semana, você realizou atividades **MODERADAS** por **pelo menos 10 minutos contínuos**, como por exemplo, pedalar leve na bicicleta, nadar, dançar, fazer ginástica aeróbica leve, jogar vôlei recreativo, carregar pesos leves, fazer serviços domésticos na casa, no quintal ou no jardim como varrer, aspirar, cuidar do jardim, ou qualquer atividade que fez aumentar **moderadamente** sua respiração ou batimentos do coração (**POR FAVOR NÃO INCLUA CAMINHADA**)?

0. ☐ Nenhum _____ dias por **SEMANA**

IPAQ3

☐

- 4J. Nos dias em que você fez essas atividades moderadas por **pelo menos 10 minutos contínuos**, quanto tempo no total você gastou fazendo essas atividades por dia?

Horas __ Minutos __

IPAQ4

☐☐☐☐

- 5J. Em quantos dias da última semana, você realizou atividades **VIGOROSAS OU PESADAS** por **pelo menos 10 minutos contínuos**, como por exemplo, correr, fazer ginástica aeróbica, jogar futebol, pedalar rápido na bicicleta, jogar basquete, fazer serviços domésticos pesados em casa, no quintal ou cuidar do jardim, carregar pesos pesados ou qualquer atividade que fez aumentar **MUITO** sua respiração ou batimentos do coração?

0. ☐ Nenhum _____ dias por **SEMANA**

IPAQ5

☐

- 6J. Nos dias em que você fez essas atividades vigorosas ou pesadas por **pelo menos 10 minutos contínuos**, quanto tempo no total você gastou fazendo essas atividades por dia?

Horas __ Minutos __

IPAQ6

☐☐☐☐

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. ☐ Sim

2. ☐ Não

SANGUE

☐

10L. Foi feita coleta de urina?

1. ☐ Sim

2. ☐ Não

URINA

☐

11L. Foi feita coleta de secreção vaginal?

1. ☐ Sim

2. ☐ Não

SECVAG

☐

12L. Foi feito exame odontológico?

1. ☐ Sim

2. ☐ Não

ODONTO

☐

13L. Foi feita ultrassonografia?

1. ☐ Sim

2. ☐ Não

USGP

☐



Preventive Medicine Reports

Open access

Guide for authors

About the journal

- [Aims and scope](#)
- [Article types](#)
- [Peer review](#)
- [Open access](#)

Ethics and policies

- [Ethics in publishing](#)
- [Submission declaration](#)
- [Authorship](#)
- [Changes to authorship](#)
- [Declaration of interests](#)
- [Funding sources](#)
- [Declaration of generative AI in scientific writing](#)
- [Preprints](#)
- [Use of inclusive language](#)
- [Reporting sex- and gender-based analyses](#)
- [Jurisdictional claims](#)
- [Studies in humans and animals](#)
- [Clinical trial results](#)

Writing and formatting

- [File format](#)
- [Title page](#)
- [Abstract](#)
- [Keywords](#)
- [Highlights](#)
- [Units, classifications codes and nomenclature](#)
- [Math formulae](#)
- [Tables](#)
- [Figures, images and artwork](#)
- [Generative AI and Figures, images and artwork](#)
- [Supplementary material](#)
- [Video](#)
- [Research data](#)
- [Data statement](#)

- [Data linking](#)
- [Research Elements](#)
- [Article structure](#)
- [References](#)
- [Journal specific information](#)

Submitting your manuscript

- [Submission checklist](#)
- [Journal specific information](#)
- [Submit online](#)

After receiving a final decision

- [Article Transfer Service](#)
- [Publishing agreement](#)
- [License options](#)
- [Open access](#)
- [Permission for copyrighted works](#)
- [Proof correction](#)
- [Responsible sharing](#)

Resources for authors

- [Elsevier Researcher Academy](#)
- [Language and editing services](#)

Getting help and support

- [Author support](#)

About the journal

Aims and scope

Preventive Medicine Reports is the open access companion journal to [Preventive Medicine](#). While *Preventive Medicine* focuses on publishing articles that appeal more broadly to the preventive medicine and health promotion and public health communities, *Preventive Medicine Reports* publishes papers that have a **subspecialty angle**, are **population** or **region-specific**, describe **study protocols**, report **preliminary** or **pilot data**, **qualitative studies**, or **focus on surrogate health outcomes**. *Preventive Medicine Reports* is also interested in methodological innovation or validation of new and existing tools. It serves as a scholarly repository for the building blocks of research that inform research, practice and policy on disease prevention and health promotion as well as public health.

All articles published in *Preventive Medicine Reports* are now available through the free journal archive PubMed Central, and discoverable via PubMed.

Article types

Preventive Medicine Reports is an international journal, publishing original, scholarly manuscripts on all topics pertaining to preventive medicine, primary care, epidemiology, and public health, as well as methodology on and historical aspects of all these topics.

Key requirements for authors to take into account when preparing a submission:

- **Original Research Paper:** no more than 3500 words in the main text, no more than 6 tables or figures, no more than 50 references
- **Short Communication:** consists of concise, focused reports of original quantitative or qualitative research. These should not exceed 2000 words, 20 references, and have no more than two tables/figures. Online supplementary materials (if any) should be limited and should not include any additional results or discussions.
- **Commentary (or Guest Editorial):** are intended to provide an opportunity for authors to offer a perspective on important or emerging issues falling within the journal's scope and aligned with its aims. These submissions must have <1500 words and no more than 15 references.
- **Review Article:** no more than 4500 words, no more than six(6) tables or figures
- **Book Review:** no more than 1000 words
- **Correspondence:** are brief (<600 words) submissions that offer reflections, corrections, support or counterarguments to previous journal publications. Correspondences may have no more than 10 references and one table/figure.
- **Structured abstract** to include the headings for Objective, Methods, Results and Conclusions are required for all article types except Correspondence, Letters to the Editor and Book Reviews. (maximum 250 words)
- **No more than eight authors.** Others contributing to the manuscript but not meeting CRediT criteria may be listed in an acknowledgment.

Peer review

This journal follows a single anonymized review process. Your submission will initially be assessed by our editors to determine suitability for publication in this journal. If your submission is deemed suitable, it will typically be sent to a minimum of two reviewers for an independent expert assessment of the scientific quality. The decision as to whether your article is accepted or rejected will be taken by our editors. Authors who wish to appeal the editorial decision for their manuscript may submit a formal appeal request in accordance with the procedure outlined in [Elsevier's Appeal Policy](#). Only one appeal per submission will be considered and the appeal decision will be final.

Read more about [peer review](#).

Our editors are not involved in making decisions about papers which:

- they have written themselves.
- have been written by family members or colleagues.
- relate to products or services in which they have an interest.

Any such submissions will be subject to the journal's usual procedures and peer review will be handled independently of the editor involved and their research group. Read more about [editor duties](#).

Special issues and article collections

The peer review process for special issues and article collections follows the same process as outlined above for regular submissions, except, a guest editor will send the submissions out to the reviewers and may recommend a decision to the journal editor. The journal editor oversees the peer review process of all special issues and article collections to ensure the high standards of publishing ethics and responsiveness are respected and is responsible for the final decision regarding acceptance or rejection of articles.

Open access

We refer you to our [open access information page](#) to learn about open access options for this journal.

Ethics and policies

Ethics in publishing

Authors must follow ethical guidelines stated in [Elsevier's Publishing Ethics Policy](#).

Submission declaration

When authors submit an article to an Elsevier journal it is implied that:

- the work described has not been published previously except in the form of a preprint, an abstract, a published lecture, academic thesis or registered report. See our policy on [multiple, redundant or concurrent publication](#).
- the article is not under consideration for publication elsewhere.
- the article's publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out.
- if accepted, the article will not be published elsewhere in the same form, in English or in any other language, including electronically, without the written consent of the copyright-holder.

To verify compliance with our journal publishing policies, we may check your manuscript with our screening tools.

Authorship

All authors should have made substantial contributions to all of the following:

1. The conception and design of the study, or acquisition of data, or analysis and interpretation of data.
2. Drafting the article or revising it critically for important intellectual content.
3. Final approval of the version to be submitted.

Authors should appoint a corresponding author to communicate with the journal during the editorial process. All authors should agree to be accountable for all aspects of the work to ensure that the questions

related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

Changes to authorship

The editors of this journal generally will not consider changes to authorship once a manuscript has been submitted. It is important that authors carefully consider the authorship list and order of authors and provide a definitive author list at original submission.

The policy of this journal around authorship changes:

- All authors must be listed in the manuscript and their details entered into the submission system.
- Any addition, deletion or rearrangement of author names in the authorship list should only be made prior to acceptance, and only if approved by the journal editor.
- Requests to change authorship should be made by the corresponding author, who must provide the reason for the request to the journal editor with written confirmation from all authors, including any authors being added or removed, that they agree with the addition, removal or rearrangement.
- All requests to change authorship must be submitted using [this form](#). Requests which do not comply with the instructions outlined in the form will not be considered.
- Only in exceptional circumstances will the journal editor consider the addition, deletion or rearrangement of authors post acceptance.
- Publication of the manuscript may be paused while a change in authorship request is being considered.
- Any authorship change requests approved by the journal editor will result in a corrigendum if the manuscript has already been published.
- Any unauthorised authorship changes may result in the rejection of the article, or retraction, if the article has already been published.

Declaration of interests

All authors must disclose any financial and personal relationships with other people or organizations that could inappropriately influence or bias their work. Examples of potential competing interests include:

- Employment
- Consultancies
- Stock ownership
- Honoraria
- Paid expert testimony
- Patent applications or registrations

- Grants or any other funding

The [Declaration of Interests tool](#) should always be completed.

Authors with no competing interests to declare should select the option, "I have nothing to declare".

The resulting Word document containing your declaration should be uploaded at the "attach/upload files" step in the submission process. It is important that the Word document is saved in the .doc/.docx file format. Author signatures are not required.

We advise you to read our [policy on conflict of interest statements, funding source declarations, author agreements/declarations and permission notes](#).

Funding sources

Authors must disclose any funding sources who provided financial support for the conduct of the research and/or preparation of the article. The role of sponsors, if any, should be declared in relation to the study design, collection, analysis and interpretation of data, writing of the report and decision to submit the article for publication. If funding sources had no such involvement this should be stated in your submission.

List funding sources in this standard way to facilitate compliance to funder's requirements:

Funding: This work was supported by the National Institutes of Health [grant numbers xxxx, yyyy]; the Bill & Melinda Gates Foundation, Seattle, WA [grant number zzzz]; and the United States Institutes of Peace [grant number aaaa].

It is not necessary to include detailed descriptions on the program or type of grants, scholarships and awards. When funding is from a block grant or other resources available to a university, college, or other research institution, submit the name of the institute or organization that provided the funding.

If no funding has been provided for the research, it is recommended to include the following sentence:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Declaration of generative AI in scientific writing

Authors must declare the use of generative AI in scientific writing upon submission of the paper. The following guidance refers only to the writing process, and not to the use of AI tools to analyse and draw insights from data as part of the research process:

- Generative AI and AI-assisted technologies should only be used in the writing process to improve the readability and language of the manuscript.
- The technology must be applied with human oversight and control and authors should carefully review and edit the result, as AI can generate authoritative-sounding output that can be incorrect, incomplete or biased. Authors are ultimately responsible and accountable for the contents of the work.
- Authors must not list or cite AI and AI-assisted technologies as an author or co-author on the manuscript since authorship implies responsibilities and tasks that can only be attributed to and

performed by humans.

The use of generative AI and AI-assisted technologies in scientific writing must be declared by adding a statement at the end of the manuscript when the paper is first submitted. The statement will appear in the published work and should be placed in a new section before the references list. An example:

- Title of new section: Declaration of generative AI and AI-assisted technologies in the writing process.
- Statement: During the preparation of this work the author(s) used [NAME TOOL / SERVICE] in order to [REASON]. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

The declaration does not apply to the use of basic tools, such as tools used to check grammar, spelling and references. If you have nothing to disclose, you do not need to add a statement.

Please read Elsevier's author policy on the use of generative AI and AI-assisted technologies, which can be found in our [GenAI Policies for journals](#).

Please note: to protect authors' rights and the confidentiality of their research, this journal does not currently allow the use of generative AI or AI-assisted technologies such as ChatGPT or similar services by reviewers or editors in the peer review and manuscript evaluation process, as is stated in our [GenAI Policies for journals](#). We are actively evaluating compliant AI tools and may revise this policy in the future.

Preprints

Preprint sharing

Authors may share preprints in line with Elsevier's [article sharing policy](#). Sharing preprints, such as on a preprint server, will not count as prior publication.

We advise you to read our policy on [multiple, redundant or concurrent publication](#).

Free preprint posting on SSRN

In support of [open science](#) this journal offers authors a free preprint posting service on [SSRN](#) to ensure early registration and dissemination of research and facilitate early citations and collaboration. Posting to SSRN is subject to SSRN's standard checks.

You will be provided with the option to release your manuscript on SSRN during the submission process. Agreeing to this option will have no effect on the editorial process or outcome, and your manuscript will remain publicly available and free to read on SSRN whether our editors accept or reject your manuscript.

You will receive an email when your preprint is posted online on SSRN and a Digital Object Identifier (DOI) is assigned.

Corresponding authors must seek approval from all co-authors before agreeing to release a manuscript publicly on SSRN.

We advise you to read about [SSRN](#), including the [SSRN Terms of Use](#) and [SSRN FAQs](#) before selecting this option.

Use of inclusive language

Inclusive language acknowledges diversity, conveys respect to all people, is sensitive to differences, and promotes equal opportunities. Authors should ensure their work uses inclusive language throughout and contains nothing which might imply one individual is superior to another on the grounds of:

- age
- gender
- race
- ethnicity
- culture
- sexual orientation
- disability or health condition

We recommend avoiding the use of descriptors about personal attributes unless they are relevant and valid. Write for gender neutrality with the use of plural nouns ("clinicians, patients/clients") as default. Wherever possible, avoid using "he, she," or "he/she."

No assumptions should be made about the beliefs of readers and writing should be free from bias, stereotypes, slang, reference to dominant culture and/or cultural assumptions.

These guidelines are meant as a point of reference to help you identify appropriate language but are by no means exhaustive or definitive.

Reporting sex- and gender-based analyses

There is no single, universally agreed-upon set of guidelines for defining sex and gender. We offer the following guidance:

- Sex and gender-based analyses (SGBA) should be integrated into research design when research involves or pertains to humans, animals or eukaryotic cells. This should be done in accordance with any requirements set by funders or sponsors and best practices within a field.
- Sex and/or gender dimensions of the research should be addressed within the article or declared as a limitation to the generalizability of the research.
- Definitions of sex and/or gender applied should be explicitly stated to enhance the precision, rigor and reproducibility of the research and to avoid ambiguity or conflation of terms and the constructs to which they refer.

We advise you to read the [Sex and Gender Equity in Research \(SAGER\) guidelines](#) and the [SAGER checklist \(PDF\)](#) on the EASE website, which offer systematic approaches to the use of sex and gender information in

study design, data analysis, outcome reporting and research interpretation.

For further information we suggest reading the rationale behind and recommended [use of the SAGER guidelines](#).

Definitions of sex and/or gender

We ask authors to define how sex and gender have been used in their research and publication. Some guidance:

- Sex generally refers to a set of biological attributes that are associated with physical and physiological features such as chromosomal genotype, hormonal levels, internal and external anatomy. A binary sex categorization (male/female) is usually designated at birth ("sex assigned at birth") and is in most cases based solely on the visible external anatomy of a newborn. In reality, sex categorizations include people who are intersex/have differences of sex development (DSD).
- Gender generally refers to socially constructed roles, behaviors and identities of women, men and gender-diverse people that occur in a historical and cultural context and may vary across societies and over time. Gender influences how people view themselves and each other, how they behave and interact and how power is distributed in society.

Jurisdictional claims

Elsevier respects the decisions taken by its authors as to how they choose to designate territories and identify their affiliations in their published content. Elsevier's policy is to take a neutral position with respect to territorial disputes or jurisdictional claims, including, but not limited to, maps and institutional affiliations. For journals that Elsevier publishes on behalf of a third party owner, the owner may set its own policy on these issues.

- Maps: Readers should be able to locate any study areas shown within maps using common mapping platforms. Maps should only show the area actually studied and authors should not include a location map which displays a larger area than the bounding box of the study area. Authors should add a note clearly stating that "*map lines delineate study areas and do not necessarily depict accepted national boundaries*". During the review process, Elsevier's editors may request authors to change maps if these guidelines are not followed.
- Institutional affiliations: Authors should use either the full, standard title of their institution or the standard abbreviation of the institutional name so that the institutional name can be independently verified for research integrity purposes.

Studies in humans and animals

Authors must follow [ethical guidelines](#) for studies carried out in humans and animals.

Studies in humans

Work which involves the use of human subjects should be carried out in accordance with the World Medical Association Declaration of Helsinki: [Ethical principles for medical research involving human subjects](#).

Manuscripts should follow the [International Committee of Medical Journal Editors \(ICMJE\) recommendations](#) for the conduct, reporting, editing and publication of scholarly work in medical journals and aim to be representative of human populations in terms of sex, age and ethnicity. [Sex and gender terms](#) should be used correctly, as outlined by WHO (World Health Organization).

Manuscripts must include a statement that all procedures were performed in compliance with relevant laws and institutional guidelines and have been approved by the appropriate institutional committee(s). The statement should contain the date and reference number of the ethical approval(s) obtained.

Manuscripts must also include a statement that the privacy rights of human subjects have been observed and that informed consent was obtained for experimentation with human subjects.

This journal will not accept manuscripts that contain data derived from unethically sourced organs or tissue, including from executed prisoners or prisoners of conscience, consistent with recommendations by [Global Rights Compliance on Mitigating Human Rights Risks in Transplantation Medicine](#). For all studies that use human organs or tissues, sufficient evidence must be provided that these were procured in line with [WHO Guiding Principles on Human Cell, Tissue and Organ Transplantation](#). For clinical studies, a statement of informed consent having been obtained from a patient or their nominated representative, paired with ethical approval for the study from a suitable institution, as required by the policies of the journal, may be considered sufficient evidence, but the journal reserves the right to request additional evidence in cases where it feels this is not sufficient. The source of the organs or tissues used in clinical research must be transparent and traceable. If your manuscript describes organ transplantation you must additionally declare within the manuscript that:

- autonomous consent free from coercion was obtained from the donor(s) or their next of kin.
- organs and/or tissues were not sourced from executed prisoners or prisoners of conscience.

Studies in animals

All animal experiments should comply with [ARRIVE \(Animal Research: Reporting of In Vivo Experiments\) guidelines](#).

Studies should be carried out in accordance with [Guidance on the operation of the Animals \(Scientific Procedures\) Act 1986](#) and associated guidelines, [EU Directive 2010/63 for the protection of animals used for scientific purposes](#) or the [NIH \(National Research Council\) Guide for the Care and Use of Laboratory Animals](#) (PDF) or those of an equivalent internationally recognized body.

The sex of animals, and where appropriate, the influence (or association) of sex on the results of the study must be indicated and a statement included in your manuscript that such guidelines as listed above have been followed.

Clinical trial results

Authors must disclose all posting in registries of results of the same or closely related work.

We follow the [International Committee of Medical Journal Editors \(ICMJE\) clinical trials guidelines](#). Editors will not consider results to be a prior publication if they have already been posted in the same clinical trials

registry in which primary registration resides, as long as the results are presented in the form of a brief structured abstract (fewer than 500 words) or table.

Disclosing results in other circumstances, such as in an investors' meeting, for example, is discouraged and may jeopardise consideration of your manuscript by this journal.

Writing and formatting

File format

We ask you to provide editable source files for your entire submission (including figures, tables and text graphics). Some guidelines:

- Save files in an editable format, using the extension .doc/.docx for Word files and .tex for LaTeX files. A PDF is not an acceptable source file.
- Lay out text in a single-column format.
- Remove any strikethrough and underlined text from your manuscript, unless it has scientific significance related to your article.
- Use spell-check and grammar-check functions to avoid errors.

We advise you to read our [Step-by-step guide to publishing with Elsevier](#).

Title page

You are required to include the following details in the title page information:

- Article title. Article titles should be concise and informative. Please avoid abbreviations and formulae, where possible, unless they are established and widely understood, e.g., DNA).
- Author names. Provide the given name(s) and family name(s) of each author. The order of authors should match the order in the submission system. Carefully check that all names are accurately spelled. If needed, you can add your name between parentheses in your own script after the English transliteration.
- Affiliations. Add affiliation addresses, referring to where the work was carried out, below the author names. Indicate affiliations using a lower-case superscript letter immediately after the author's name and in front of the corresponding address. Ensure that you provide the full postal address of each affiliation, including the country name and, if available, the email address of each author.
- Corresponding author. Clearly indicate who will handle correspondence for your article at all stages of the refereeing and publication process and also post-publication. This responsibility includes answering any future queries about your results, data, methodology and materials. It is important that the email address and contact details of your corresponding author are kept up to date during the submission and publication process.
- Present/permanent address. If an author has moved since the work described in your article was carried out, or the author was visiting during that time, a "present address" (or "permanent address") can be

indicated by a footnote to the author's name. The address where the author carried out the work must be retained as their main affiliation address. Use superscript Arabic numerals for such footnotes.

Abstract

You are required to provide a concise and factual abstract which does not exceed 250 words. The abstract should briefly state the purpose of your research, principal results and major conclusions. Some guidelines:

- Abstracts must be able to stand alone as abstracts are often presented separately from the article.
- Avoid references. If any are essential to include, ensure that you cite the author(s) and year(s).
- Avoid non-standard or uncommon abbreviations. If any are essential to include, ensure they are defined within your abstract at first mention.

Keywords

You are required to provide 1 to 7 keywords for indexing purposes. Keywords should be written in English. Please try to avoid keywords consisting of multiple words (using "and" or "of").

We recommend that you only use abbreviations in keywords if they are firmly established in the field.

Highlights

You are required to provide article highlights at submission.

Highlights are a short collection of bullet points that should capture the novel results of your research as well as any new methods used during your study. Highlights will help increase the discoverability of your article via search engines. Some guidelines:

- Submit highlights as a separate editable file in the online submission system with the word "highlights" included in the file name.
- Highlights should consist of 3 to 5 bullet points, each a maximum of 85 characters, including spaces.

We encourage you to view example [article highlights](#) and read about the benefits of their inclusion.

Units, classifications codes and nomenclature

This journal requires you to use the international system of units (SI) which follows internationally accepted rules and conventions. If other units are mentioned within your article, you should provide the equivalent unit in SI.

Math formulae

- Submit math equations as editable text, not as images.
- Present simple formulae in line with normal text, where possible.
- Use the solidus (/) instead of a horizontal line for small fractional terms such as X/Y.

- Present variables in italics.
- Denote powers of e by exp.
- Display equations separately from your text, numbering them consecutively in the order they are referred to within your text.

Tables

Tables must be submitted as editable text, not as images. Some guidelines:

- Place tables next to the relevant text or on a separate page(s) at the end of your article.
- Cite all tables in the manuscript text.
- Number tables consecutively according to their appearance in the text.
- Please provide captions along with the tables.
- Place any table notes below the table body.
- Avoid vertical rules and shading within table cells.

We recommend that you use tables sparingly, ensuring that any data presented in tables is not duplicating results described elsewhere in the article.

Figures, images and artwork

Figures, images, artwork, diagrams and other graphical media must be supplied as separate files along with the manuscript. We recommend that you read our detailed [artwork and media instructions](#). Some excerpts:

When submitting artwork:

- Cite all images in the manuscript text.
- Number images according to the sequence they appear within your article.
- Submit each image as a separate file using a logical naming convention for your files (for example, Figure_1, Figure_2 etc).
- Please provide captions for all figures, images, and artwork.
- Text graphics may be embedded in the text at the appropriate position. If you are working with LaTeX, text graphics may also be embedded in the file.

Artwork formats

When your artwork is finalized, "save as" or convert your electronic artwork to the formats listed below taking into account the given resolution requirements for line drawings, halftones, and line/halftone

combinations:

- Vector drawings: Save as EPS or PDF files embedding the font or saving the text as "graphics."
- Color or grayscale photographs (halftones): Save as TIFF, JPG or PNG files using a minimum of 300 dpi (for single column: min. 1063 pixels, full page width: 2244 pixels).
- Bitmapped line drawings: Save as TIFF, JPG or PNG files using a minimum of 1000 dpi (for single column: min. 3543 pixels, full page width: 7480 pixels).
- Combinations bitmapped line/halftones (color or grayscale): Save as TIFF, JPG or PNG files using a minimum of 500 dpi (for single column: min. 1772 pixels, full page width: 3740 pixels).

Please do not submit:

- files that are too low in resolution (for example, files optimized for screen use such as GIF, BMP, PICT or WPG files).
- disproportionally large images compared to font size, as text may become unreadable.

Figure captions

All images must have a caption. A caption should consist of a brief title (not displayed on the figure itself) and a description of the image. We advise you to keep the amount of text in any image to a minimum, though any symbols and abbreviations used should be explained.

Provide captions in a separate file.

Color artwork

If you submit usable color figures with your accepted article, we will ensure that they appear in color online.

Please ensure that color images are accessible to all, including those with impaired color vision. Learn more about [color and web accessibility](#).

For articles appearing in print, you will be sent information on costs to reproduce color in the printed version, after your accepted article has been sent to production. At this stage, please indicate if your preference is to have color only in the online version of your article or also in the printed version.

Generative AI and Figures, images and artwork

Please read our policy on the use of generative AI and AI-assisted tools in figures, images and artwork, which can be found in Elsevier's [GenAI Policies for Journals](#). This policy states:

- We do not permit the use of Generative AI or AI-assisted tools to create or alter images in submitted manuscripts.
- The only exception is if the use of AI or AI-assisted tools is part of the research design or methods (for example, in the field of biomedical imaging). If this is the case, such use must be described in a

reproducible manner in the methods section, including the name of the model or tool, version and extension numbers, and manufacturer.

- The use of generative AI or AI-assisted tools in the production of artwork such as for graphical abstracts is not permitted. The use of generative AI in the production of cover art may in some cases be allowed, if the author obtains prior permission from the journal editor and publisher, can demonstrate that all necessary rights have been cleared for the use of the relevant material, and ensures that there is correct content attribution.

Supplementary material

We encourage the use of supplementary materials such as applications, images and sound clips to enhance research. Some guidelines:

- Cite all supplementary files in the manuscript text.
- Submit supplementary materials at the same time as your article. Be aware that all supplementary materials provided will appear online in the exact same file type as received. These files will not be formatted or typeset by the production team.
- Include a concise, descriptive caption for each supplementary file describing its content.
- Provide updated files if at any stage of the publication process you wish to make changes to submitted supplementary materials.
- Do not make annotations or corrections to a previous version of a supplementary file.
- Switch off the option to track changes in Microsoft Office files. If tracked changes are left on, they will appear in your published version.

Video

This journal accepts video material and animation sequences to support and enhance your scientific research. We encourage you to include links to video or animation files within articles. Some guidelines:

- When including video or animation file links within your article, refer to the video or animation content by adding a note in your text where the file should be placed.
- Clearly label files ensuring the given file name is directly related to the file content.
- Provide files in one of our [recommended file formats](#). Files should be within our preferred maximum file size of 150 MB per file, 1 GB in total.
- Provide "stills" for each of your files. These will be used as standard icons to personalize the link to your video data. You can choose any frame from your video or animation or make a separate image.
- Provide text (for both the electronic and the print version) to be placed in the portions of your article that refer to the video content. This is essential text, as video and animation files cannot be embedded in the print version of the journal.

We publish all video and animation files supplied in the electronic version of your article.

For more detailed instructions, we recommend that you read our guidelines on [submitting video content to be included in the body of an article](#).

Research data

We are committed to supporting the storage of, access to and discovery of research data, and our [research data policy](#) sets out the principles guiding how we work with the research community to support a more efficient and transparent research process.

Research data refers to the results of observations or experimentation that validate research findings, which may also include software, code, models, algorithms, protocols, methods and other useful materials related to the project.

Please read our guidelines on [sharing research data](#) for more information on depositing, sharing and using research data and other relevant research materials.

For this journal, the following instructions from our [research data guidelines](#) apply.

Option C: Research data deposit, citation and linking

You are **required** to:

- Deposit your research data in a relevant data repository.
- Cite and link to this dataset in your article.
- If this is not possible, make a statement explaining why research data cannot be shared.

Data statement

To foster transparency, you are required to state the availability of any data at submission.

Ensuring data is available may be a requirement of your funding body or institution. If your data is unavailable to access or unsuitable to post, you can state the reason why (e.g., your research data includes sensitive or confidential information such as patient data) during the submission process. This statement will appear with your published article on ScienceDirect.

Read more about the importance and benefits of providing a [data statement](#).

Data linking

Linking to the data underlying your work increases your exposure and may lead to new collaborations. It also provides readers with a better understanding of the described research.

If your research data has been made available in a data repository there are a number of ways your article can be linked directly to the dataset:

- Provide a link to your dataset when prompted during the online submission process.

- For some data repositories, a repository banner will automatically appear next to your published article on ScienceDirect.
- You can also link relevant data or entities within the text of your article through the use of identifiers. Use the following format: Database: 12345 (e.g. TAIR: AT1G01020; CCDC: 734053; PDB: 1XFN).

Learn more about [linking research data and research articles in ScienceDirect](#).

Research Elements

This journal enables the publication of research objects (e.g. data, methods, protocols, software and hardware) related to original research in [Elsevier's Research Elements journals](#).

Research Elements are peer-reviewed, open access journals which make research objects findable, accessible and reusable. By providing detailed descriptions of objects and their application with links to the original research article, your research objects can be placed into context within your article.

You will be alerted during submission to the opportunity to submit a manuscript to one of the Research Elements journals. Your Research Elements article can be prepared by you, or by one of your collaborators.

Article structure

Article sections

- Divide your article into clearly defined and numbered sections. Number subsections 1.1 (then 1.1.1, 1.1.2, ...), then 1.2, etc.
- Use the numbering format when cross-referencing within your article. Do not just refer to "the text."
- You may give subsections a brief heading. Headings should appear on a separate line.
- Do not include the article abstract within section numbering.

Theory and calculation

The theory section should lay the foundation for further work by extending the background you provided in the introduction to your article. The calculation section should represent a practical development from a theoretical basis.

Glossary

Please provide definitions of field-specific terms used in your article, in a separate list.

Footnotes

We advise you to use footnotes sparingly. If you include footnotes in your article, ensure that they are numbered consecutively.

You may use system features that automatically build footnotes into text. Alternatively, you can indicate the position of footnotes within the text and present them in a separate section at the end of your article.

Acknowledgements

Include any individuals who provided you with help during your research, such as help with language, writing or proof reading, in the acknowledgements section. Acknowledgements should be placed in a separate section which appears directly before the reference list. Do not include acknowledgements on your title page, as a footnote to your title, or anywhere else in your article other than in the separate acknowledgements section.

Author contributions: CRediT

Corresponding authors are required to acknowledge co-author contributions using [CRediT \(Contributor Roles Taxonomy\)](#) roles:

- Conceptualization
- Data curation
- Formal analysis
- Funding acquisition
- Investigation
- Methodology
- Project administration
- Resources
- Software
- Supervision
- Validation
- Visualization
- Writing – original draft
- Writing – review and editing

Not all CRediT roles will apply to every manuscript and some authors may contribute through multiple roles.

We advise you to read [more about CRediT](#) and [view an example of a CRediT author statement](#).

Funding sources

Authors must disclose any funding sources who provided financial support for the conduct of the research and/or preparation of the article. The role of sponsors, if any, should be declared in relation to the study design, collection, analysis and interpretation of data, writing of the report and decision to submit the article for publication. If funding sources had no such involvement this should be stated in your submission.

List funding sources in this standard way to facilitate compliance to funder's requirements:

Funding: This work was supported by the National Institutes of Health [grant numbers xxxx, yyyy]; the Bill & Melinda Gates Foundation, Seattle, WA [grant number zzzz]; and the United States Institutes of Peace [grant number aaaa].

It is not necessary to include detailed descriptions on the program or type of grants, scholarships and awards. When funding is from a block grant or other resources available to a university, college, or other research institution, submit the name of the institute or organization that provided the funding.

If no funding has been provided for the research, it is recommended to include the following sentence:

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Appendices

We ask you to use the following format for appendices:

- Identify individual appendices within your article using the format: A, B, etc.
- Give separate numbering to formulae and equations within appendices using formats such as Eq. (A.1), Eq. (A.2), etc. and in subsequent appendices, Eq. (B.1), Eq. (B. 2) etc. In a similar way, give separate numbering to tables and figures using formats such as Table A.1; Fig. A.1, etc.

References

References within text

Any references cited within your article should also be present in your reference list and vice versa. Some guidelines:

- References cited in your abstract must be given in full.
- We recommend that you do not include unpublished results and personal communications in your reference list, though you may mention them in the text of your article.
- Any unpublished results and personal communications included in your reference list must follow the standard reference style of the journal. In substitution of the publication date add "unpublished results" or "personal communication."
- References cited as "in press" imply that the item has been accepted for publication.

Linking to cited sources will increase the discoverability of your research.

Before submission, check that all data provided in your reference list are correct, including any references which have been copied. Providing correct reference data allows us to link to abstracting and indexing services such as Scopus, Crossref and PubMed. Any incorrect surnames, journal or book titles, publication years or pagination within your references may prevent link creation.

We encourage the use of Digital Object Identifiers (DOIs) as reference links as they provide a permanent link to the electronic article referenced.

Reference format

This journal does not set strict requirements on reference formatting at submission. Some guidelines:

- References can be in any style or format as long as the style is consistent.
- Author names, journal or book titles, chapter or article titles, year of publication, volume numbers, article numbers or pagination must be included, where applicable.
- Use of DOIs is recommended.

Our journal reference style will be applied to your article after acceptance, at proof stage. If required, at this stage we will ask you to correct or supply any missing reference data.

Reference style

All citations in the text should refer to:

- Single author: the author's name (without initials, unless there is ambiguity) and the year of publication.
- Two authors: both authors' names and the year of publication.
- Three or more authors: first author's name followed by 'et al.' and the year of publication.

Citations can be made directly (or parenthetically). Groups of references can be listed either first alphabetically, then chronologically, or vice versa. Examples: "as demonstrated (Allan, 2020a, 2020b; Allan and Jones, 2019)" or "as demonstrated (Jones, 2019; Allan, 2020). Kramer et al. (2023) have recently shown".

The list of references should be arranged alphabetically and then chronologically if necessary. More than one reference from the same author(s) in the same year must be identified by the letters 'a', 'b', 'c', etc., placed after the year of publication.

Abbreviate journal names according to the [List of Title Word Abbreviations](#) (LTWA).

Examples:

Reference to a journal publication:

Van der Geer, J., Handgraaf, T., Lupton, R.A., 2020. The art of writing a scientific article. *J. Sci. Commun.* 163, 51–59. <https://doi.org/10.1016/j.sc.2020.00372>.

Reference to a journal publication with an article number:

Van der Geer, J., Handgraaf, T., Lupton, R.A., 2022. The art of writing a scientific article. *Heliyon*. 19, e00205. <https://doi.org/10.1016/j.heliyon.2022.e00205>.

Reference to a book:

Strunk Jr., W., White, E.B., 2000. *The Elements of Style*, fourth ed. Longman, New York.

Reference to a chapter in a book:

Mettam, G.R., Adams, L.B., 2023. How to prepare an electronic version of your article, in: Jones, B.S., Smith, R.Z. (Eds.), *Introduction to the Electronic Age*. E-Publishing Inc., New York, pp. 281–304.

Reference to a website:

Cancer Research UK, 2023. Cancer statistics reports for the UK.
<http://www.cancerresearchuk.org/aboutcancer/statistics/cancerstatsreport/> (accessed 13 March 2023).

Reference to a dataset:

Oguro, M., Imahiro, S., Saito, S., Nakashizuka, T., 2015. Mortality data for Japanese oak wilt disease and surrounding forest compositions [dataset]. Mendeley Data, v1. <https://doi.org/10.17632/xwj98nb39r1>.

Reference to software:

Coon, E., Berndt, M., Jan, A., Svyatsky, D., Atchley, A., Kikinzon, E., Harp, D., Manzini, G., Shelef, E., Lipnikov, K., Garimella, R., Xu, C., Moulton, D., Karra, S., Painter, S., Jafarov, E., & Molins, S., 2020. Advanced Terrestrial Simulator (ATS) v0.88 (Version 0.88) [software]. Zenodo. <https://doi.org/10.5281/zenodo.3727209>.

Web references

When listing web references, as a minimum you should provide the full URL and the date when the reference was last accessed. Additional information (e.g. DOI, author names, dates or reference to a source publication) should also be provided, if known.

You can list web references separately under a new heading directly after your reference list or include them in your reference list.

Data references

We encourage you to cite underlying or relevant datasets within article text and to list data references in the reference list.

When citing data references, you should include:

- author name(s)
- dataset title
- data repository
- version (where available)
- year
- global persistent identifier

Add [dataset] immediately before your reference. This will help us to properly identify the dataset. The [dataset] identifier will not appear in your published article.

Preprint references

We ask you to mark preprints clearly. You should include the word "preprint" or the name of the preprint server as part of your reference and provide the preprint DOI.

Where a preprint has subsequently become available as a peer-reviewed publication, use the formal publication as your reference.

If there are preprints that are central to your work or that cover crucial developments in the topic, but they are not yet formally published, you may reference the preprint.

Reference management software

Most Elsevier journals have their reference template available in popular reference management software products. These include products that support [Citation Style Language \(CSL\)](#) such as [Mendeley Reference Manager](#).

If you use a citation plug-in from these products, select the relevant journal template and all your citations and bibliographies will automatically be formatted in the journal style. We advise you to [remove all field codes](#) before submitting your manuscript to any reference management software product.

If a template is not available for this journal, follow the format given in examples in the reference style section of this Guide for Authors.

Journal specific information

Title Page

Add word counts for the abstract and main text, excluding citations. Provide the number of references.

Tables

Please provide captions along with the tables. These must be informative to allow the table to stand alone without reference to the main manuscript. Indicate the "who", "what", "when" and "where". Do not use abbreviations in the title.

Avoid vertical rules and shading within table cells. Do not use bold type face.

Submitting your manuscript

Submission checklist

Before completing the submission of your manuscript, we advise you to read our submission checklist:

- One author has been designated as the corresponding author and their full contact details (email address, full postal address and phone numbers) have been provided.
- All files have been uploaded, including keywords, figure captions and tables (including a title, description and footnotes) included.

- Spelling and grammar checks have been carried out.
- All references in the article text are cited in the reference list and vice versa.
- Permission has been obtained for the use of any copyrighted material from other sources, including the Web.
- For gold open access articles, all authors understand that they are responsible for payment of the article publishing charge (APC) if the manuscript is accepted. Payment of the APC may be covered by the corresponding author's institution, or the research funder.

Journal specific information

All abbreviations are spelled out with first use. Please note that **two-letter abbreviations are not permitted** (excluding OR, HR and CI). Any word abbreviation not used later in the manuscript after the initial definition should not be abbreviated.

Submit online

Our online submission system guides you through the process steps of entering your manuscript details and uploading your files. The system converts your article files to a single PDF file used in the peer-review process.

Editable files (e.g., Word, LaTeX) are required to typeset your article for final publication. All correspondence, including notification of the editor's decision and requests for revision, is sent by email.

Please follow this link to [submit your paper](#).

After receiving a final decision

Article Transfer Service

If your manuscript is more suitable for an alternative Elsevier journal, you may receive an email asking you to consider transferring your manuscript via the [Elsevier Article Transfer Service](#).

The recommendation could come from the journal editor, a dedicated [in-house scientific managing editor](#), a tool-assisted recommendation or a combination.

If you agree with the recommendation, your manuscript will be transferred and independently reviewed by the editors of the new journal. You will have the opportunity to make revisions, if necessary, before the submission is complete at the destination journal.

Publishing agreement

Authors will be asked to complete a publishing agreement after acceptance. The corresponding author will receive a link to the online agreement by email. We advise you to read [Elsevier's policies related to copyright](#) to learn more about our copyright policies and your, and your employer's/institution's, additional rights for subscription and gold open access articles.

License options

Authors will be offered [open access user license options](#) which will determine how you, and third parties, can reuse your gold open access article. We advise that you review these options and any funding body license requirements before selecting a license option.

Open access

We refer you to our [open access information page](#) to learn about open access options for this journal.

Permission for copyrighted works

If excerpts from other copyrighted works are included in your article, you must obtain written permission from the copyright owners and credit the source(s) within your article using [Elsevier's permission request and license form](#) (Word).

Proof correction

To ensure a fast publication process we will ask you to provide proof corrections within two days.

Corresponding authors will be sent an email which includes a link to our online proofing system, allowing annotation and correction of proofs online. The environment is similar to Word. You can edit text, comment on figures and tables and answer questions raised by our copy editor. Our web-based proofing service ensures a faster and less error-prone process.

You can choose to annotate and upload your edits on the PDF version of your article, if preferred. We will provide you with proofing instructions and available alternative proofing methods in our email.

The purpose of the proof is to check the typesetting, editing, completeness and correctness of your article text, tables and figures. Significant changes to your article at the proofing stage will only be considered with approval of the journal editor.

Responsible sharing

We encourage you to share and promote your article to give additional visibility to your work, enabling your paper to contribute to scientific progress and foster the exchange of scientific developments within your field. Read more about how to [responsibly share and promote your article](#).

Resources for authors

Elsevier Researcher Academy

If you would like help to improve your submission or navigate the publication process, support is available via [Elsevier Researcher Academy](#).

Elsevier Researcher Academy offers free e-learning modules, webinars, downloadable guides and research writing and peer review process resources.

Language and editing services

We recommend that you write in American or British English but not a combination of both.

If you feel the English language in your manuscript requires editing to eliminate possible grammatical or spelling errors and to conform to correct scientific English, you may wish to use the [English Language Editing service](#) provided by Elsevier's Author Services.

Getting help and support

Author support

We recommend that you visit our [Journal Article Publishing Support Center](#) if you have questions about the editorial process or require technical support for your submission. Some popular FAQs:

- [How can I track the status of my submitted article?](#)
- [When will my article be published?](#)



All content on this site: Copyright © 2025 Elsevier B.V., its licensors, and contributors. All rights are reserved, including those for text and data mining, AI training, and similar technologies. For all open access content, the relevant licensing terms apply.