

**PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ  
ESCOLA POLITÉCNICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIA EM SAÚDE**

**CRISTIANE YUMI NAKAMURA**

**COORTE HISTÓRICA E MINERAÇÃO DE DADOS: PADRÕES DE ÓBITO NOS  
PRIMEIROS 1.000 DIAS DE VIDA NO MUNICÍPIO DE CURITIBA**

**CURITIBA**

**2020**

**CRISTIANE YUMI NAKAMURA**

**COORTE HISTÓRICA E MINERAÇÃO DE DADOS: PADRÕES DE ÓBITO NOS  
PRIMEIROS 1.000 DIAS DE VIDA NO MUNICÍPIO DE CURITIBA**

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em Tecnologia em Saúde da Pontifícia Universidade Católica do Paraná, como requisito parcial à obtenção do título de Mestre em Tecnologia em Saúde.

Linha de Pesquisa: Informática em Saúde.

Orientadora: Profa. Dra. Deborah Ribeiro Carvalho

Coorientadora: Profa. Dra. Simone Tetu Moysés

**CURITIBA**

2020

Dados da Catalogação na Publicação  
 Pontifícia Universidade Católica do Paraná  
 Sistema Integrado de Bibliotecas – SIBI/PUCPR  
 Biblioteca Central

**EDILENE DE OLIVEIRA DOS SANTOS CRB-9/1636**

	Nakamura, Cristiane Yumi
N163c	Coorte histórica e mineração de dados : padrões de óbito nos primeiros 1.000 dias de vida no município de Curitiba / Cristiane Yumi Nakamura ; orientadora, Deborah Ribeiro Carvalho ; coorientadora, Simone Tetu Moyses. -- 2020
2020	178 f. : il. ; 30 cm
	Dissertação (mestrado) – Pontifícia Universidade Católica do Paraná, Curitiba, 2020
	Bibliografia: f.119-126
	1. Mortalidade infantil. 2. Serviços de saúde infantil. 3. Causas de morte. 4. Condições sociais. 5. Centros de Saúde Materno-Infantil. 6. Tecnologia biomédica. I. Carvalho, Deborah Ribeiro. II. Moyses, Simone Tetu. III. Pontifícia Universidade Católica do Paraná. Programa de Pós-Graduação em Tecnologia em Saúde. IV. Título
	CDD. 20.ed. – 618.24



Pontifícia Universidade Católica do Paraná

Escola Politécnica

Programa de Pós-Graduação em Tecnologia em Saúde

### TERMO DE APROVAÇÃO DE DISSERTAÇÃO Nº 276

A Dissertação de Mestrado intitulada **“COORTE HISTÓRICA E MINERAÇÃO DE DADOS: PADRÕES DE ÓBITO NOS PRIMEIROS 1.000 DIAS DE VIDA NO MUNICÍPIO DE CURITIBA”**, defendida em sessão pública pela candidata **Cristiane Yumi Nakamura**, no dia **18 de fevereiro de 2020**, foi julgada para a obtenção do grau de Mestre em Tecnologia em Saúde, e aprovada em sua forma final, pelo Programa de Pós-Graduação em Tecnologia em Saúde.

BANCA EXAMINADORA:

Prof. Dr. Deborah Ribeiro Carvalho - Presidente – (PUCPR)

Prof. Dr. Marcia Regina Cubas – (PUCPR)

Prof. Dr. Denise Siqueira de Carvalho – (UFPR)

A via original deste documento encontra-se arquivada na Secretaria do Programa, contendo a assinatura da Coordenação após a entrega da versão corrigida do trabalho.

Curitiba, 04 de maio de 2020.

---

Prof.<sup>o</sup>. Dr.<sup>o</sup>. Percy Nohama  
**Coordenador do Programa de  
Pós-Graduação em Tecnologia em Saúde  
PUCPR**

Dedico esta dissertação às crianças  
cuja finitude precoce da vida nos  
proporciona aprendizados para melhorarmos  
como profissionais de saúde e, principalmente,  
como seres humanos.

## AGRADECIMENTOS

“Quem caminha sozinho pode até chegar mais rápido, mas aquele que vai acompanhado, com certeza vai mais longe” (Clarice Lispector).

Gratidão eterna à minha família. A meus pais, João e Iolanda, por me ensinarem o valor do estudo e conhecimento. A meus irmãos, Luciane e Fábio, e cunhados, Fábio e Flávia, por serem tão compreensivos com minha ausência. Ao meu amor, Vinícius, pelo incentivo e companheirismo, sempre. A JP, Valentina e Catarina, pelas bagunças que me davam mais energia para seguir em frente. Amo todos vocês!

À Profa. Dra. Deborah Ribeiro Carvalho, por ter me recebido de braços abertos e nunca ter desistido de mim. Sua orientação e exemplo me proporcionaram aprendizados que levarei para o resto da vida. Obrigada por ter me ajudado a ser uma pesquisadora e pessoa melhor.

À Profa. Dra. Simone Tetu Moysés, minha coorientadora, por ter plantado em mim a sementinha da saúde coletiva na graduação e continuar regando essa árvore até hoje. Seus *insights* e contribuições foram fundamentais.

Aos professores do PPGTS, pela disposição em ajudar e genuinamente repassar seus conhecimentos.

À secretária Ana, sempre generosa e atenciosa, cuidando dos trâmites burocráticos. Aos colegas do grupo de pesquisa, Yohan, Marcelo, Ana Luiza, Cristiana, Georgia, Márcia e Adans, por compartilharem suas experiências e contribuírem com a minha. Em especial ao João pelo ombro amigo, além das ajudas científicas; Everton, pela parceria desde o primeiro dia; Elziane e Lorena, pelos estudos, ideias e risadas; Ana Paula e Verônica, que sempre se fizeram presentes; Danilo, pela paciência e grande ajuda.

Aos estudantes do Projeto Institucional de Bolsas de Iniciação Científica, Ingra, Lamon e Ana Flávia, e ao colaborador Iago, cuja participação foi essencial.

À CAPES/PROSUP, pela concessão de taxa escolar para realização do mestrado.

À Micheli e equipe e Cris Marangon, pela ajuda com suas experiências na Secretaria Municipal de Saúde. Ao Alcides, Terumi, Sandra, Nilza, Sonia, Vivi, Vanda e Raquel, pelo incentivo e por “segurarem as pontas” na minha ausência.

“Àz amigas” do BBBR pela amizade, apoio incondicional e por sempre se lembrarem de mim, mesmo com a minha alta taxa de absenteísmo nesses dois anos.

Deus, gratidão por me proporcionar tudo isso.

Acredite no fluxo.  
O que hoje é pesado, amanhã é  
aprendizado...  
Tudo passa.

(Geronimo Theml)

## RESUMO

**Introdução:** Apesar da redução das taxas brasileiras de mortalidade infantil e mortalidade na infância, bebês ainda morrem por causas evitáveis, com diferenças entre regiões e estados brasileiros. A taxa de mortalidade infantil em Curitiba, capital do Paraná, é considerada baixa (8,2/1.000 nascidos vivos em 2018), o que torna a sua redução um desafio, exigindo concentração dos esforços nos óbitos por causas evitáveis. **Objetivo:** Identificar padrões de óbito de zero a dois anos de idade em coorte histórica do município de Curitiba. **Método:** Coorte histórica reconstruída a partir do Sistema de Informações sobre Nascidos Vivos e Sistema de Informações sobre Mortalidade. A população compreendeu nascidos vivos de 1999 a 2017, residentes em Curitiba, e óbitos de zero a dois anos de idade. Realizou-se o processo *Knowledge Discovery in Database* e suas etapas – no pré-processamento, ocorreram a consistência e *linkage* dos registros; na mineração de dados, executaram-se as tarefas de classificação e de descoberta de regras de associação, considerando como desfecho “evitabilidade do óbito”; e, no pós-processamento, utilizaram-se as estratégias de transcrição de regras, generalização, descoberta de regras de exceção e filtro de regras de associação. Realizou-se teste estatístico de qui-quadrado e Kruskal-Wallis, adotando nível de significância de  $p \leq 0,05$  e intervalo de confiança 95%. **Resultados:** Corroborou-se a relação do óbito com malformação congênita ( $p < 0,001$ ), Apgar no primeiro minuto ( $p < 0,001$ ), Apgar no quinto minuto ( $p < 0,001$ ) e peso ao nascer ( $p < 0,001$ ). Os valores de corte de peso ao nascer relacionados ao óbito aproximaram-se de 1.500g. Outros valores de corte encontrados foram sete para Apgar no quinto minuto, entre 28 e 31 semanas para duração da gestação e um mês completo para idade da criança. Evidenciou-se que “classificação de risco ao nascer” ( $p < 0,001$ ) e “idade da criança em meses completos” ( $p < 0,001$ ) possuem relação com o óbito evitável, especialmente quando associados às variáveis “DS de residência” ( $p = 0,024$ ) e “vinculação SUS” ( $p < 0,001$ ). **Conclusão:** A abordagem mista, compreendendo algoritmos, estatística descritiva e testes estatísticos, permitiu corroborar fatores já conhecidos, identificar diferentes valores de corte e descobrir variáveis de governabilidade do gestor com potencial de contribuição para enfrentamento do problema. Destacaram-se valor de corte de “peso ao nascer” abaixo do critério utilizado para risco, período crítico de óbito além do período neonatal e “classificação de risco ao nascer” relacionado ao óbito evitável por ação da saúde, sobretudo quando associado às condições socioeconômicas e ambientais – “DS de residência” e “vinculação SUS”. Suscita-se a necessidade de reavaliação dos critérios de risco ao nascer atualmente utilizados, com sugestão de inclusão de critérios que representem condições socioeconômicas e ambientais, como o índice de vulnerabilidade das áreas de abrangência das unidades municipais de saúde. A qualidade dos registros e a limitação do *software* utilizado podem ter gerado ruídos no conjunto de dados. Sugerem-se a análise das regras, não possíveis de serem esgotadas neste estudo, e a agregação de dados de outros sistemas de informação.

**Palavras-chave:** Mortalidade infantil. Mortalidade da criança. Saúde materno-infantil. Causas de morte. Determinantes sociais da saúde.



## ABSTRACT

**Introduction:** In Brazil, despite reduction of Infant Mortality and Childhood Mortality Rates, babies still die from preventable causes, given differences among Brazilian regions and states. In Curitiba, Infant Mortality Rate (8.2/1,000 live births in 2018) is considered low and its reduction is challenging, which requires greater efforts to avoid deaths from preventable causes. **Objective:** Identify death patterns in children up to two years old in a historical cohort in Curitiba. **Method:** Historical cohort reconstructed from Live Birth Information System (SINASC) and Mortality Information System (SIM). Population was composed of live births from 1999 to 2017 and deaths up to two years old of Curitiba resident individuals. The Knowledge Discovery in Database process and its steps were performed - consistency and linkage of records were carried out along pre-processing; Classification and Association Rules Discovery tasks were performed in data mining, considering "avoidability of death" as the outcome, in post-processing phase the strategies applied were rule transcription, generalization, discovery of exception rules and association rules filter. Chi-square and Kruskal Wallis statistical tests were undertaken at a significance level of  $p \leq 0.05$  and a confidence interval of 95%. **Results:** Findings supported the relationship between death and congenital malformation ( $p < 0.001$ ), Apgar score at 1 min ( $p < 0.001$ ), Apgar score at 5 min ( $p < 0.001$ ) and birthweight ( $p < 0.001$ ). Birthweight cut-off values in association with death approached 1,500g. Other cut-off values encountered were seven for Apgar at 5 minutes, over 28 to 31 weeks of pregnancy and reaching one month of age. It was evidenced that "risk rating at birth" ( $p < 0.001$ ) and "age of child in completed months" ( $p < 0.001$ ) are related to preventable death, especially when associated with the variables "DS of residence" ( $p = 0.024$ ) and "SUS linkage" ( $p < 0.001$ ). **Conclusion:** The mixed approach consisting of algorithms, descriptive statistics and statistical tests endorsed known factors, by identifying different cut-off values and discovering governable variables in order to improve manager's confrontation against the issue. It was highlighted that birthweight cut-off value under the criteria adopted for risk, critical period of death beyond neonatal period and "birth risk classification" were associated to preventable death, especially when related to socioeconomic and environmental conditions - "DS of residence" and "SUS linkage". In face of the need to reevaluate the currently used birth risk criteria, the inclusion of socioeconomic and environmental conditions as criteria, such as Vulnerability Index of Areas Covered by Municipal Health Units (IVAB) is implied. Software limitation may have generated noise in data set. We suggest screening the rules that could not be exhausted in this study and add data from other information systems.

**Keywords:** Infant mortality. Child mortality. Maternal and child health. Cause of death. Social determinants of health.

## LISTA DE FIGURAS

Figura 1 – Modelo dos DSS .....	23
Figura 2 – Representação dos SISs utilizados na pesquisa em relação às fases dos primeiros 1.000 dias de vida.....	26
Figura 3 – KDD e etapas principais.....	33
Figura 4 – Tela do <i>software</i> WEKA, com sinalização de alguns elementos da tarefa de classificação. ....	34
Figura 5 – Segmento de árvore de decisão com sinalização do nó-raiz, nós filhos e nó-folha. ....	35
Figura 6 – Segmento de árvore de decisão com sinalização de uma ramificação, cobertura e erro.....	35
Figura 7 – Exemplos de regras descobertas na tarefa de associação, contendo desfecho de óbito. ....	36
Figura 8 – Formulação matemática de suporte e confiança da tarefa de descoberta de regras de associação (Regra $A \rightarrow C$ ).....	37
Figura 9 – Exemplos de regras gerais e de exceção após submissão do conjunto de regras a um algoritmo DRE.....	38
Figura 10 – Etapas da pesquisa.....	43
Figura 11 – Exemplo de transcrição de árvore de decisão em regras de linguagem natural. ....	47
Figura 12 – Exemplo de transcrição de regras gerais e de exceção em regras de linguagem natural.....	48
Figura 13 – Resultados gerais do <i>linkage</i> . ....	52
Figura 14 – Variáveis encontradas no experimento 1 por ano, com representação da força da relação ao óbito pela numeração em ordem crescente do nó-raiz ao nó-folha. ....	57
Figura 15 – Valores de corte de peso ao nascer encontrados em cada árvore de decisão do experimento 1 por ano, relacionados aos conceitos de baixo peso ao nascer. ....	58
Figura 16 – Variáveis encontradas no experimento 2 por ano, com representação da força da associação ao óbito pela numeração em ordem crescente do nó-raiz ao nó-folha. ....	59
Figura 17 – Regras encontradas no experimento 3. ....	60

Figura 18 – Regra geral 2 e respectivas regras de exceção selecionadas no pós-processamento do experimento 4. ....	61
Figura 19 – Regra geral 4 e respectivas regras de exceção selecionadas no pós-processamento do experimento 4. ....	62
Figura 20 – TMI's por DS e ano de 2007 a 2016 (preliminar), no município de Curitiba. ....	70
Figura 21 – <i>Script</i> do <i>linkage</i> realizado no R.® .....	143
Figura 22 – Resultados detalhados do <i>linkage</i> . ....	144
Figura 23 – Árvore de decisão do ano de 2000 do experimento 1. ....	145
Figura 24 – Árvore de decisão do ano de 2001 do experimento 1. ....	145
Figura 25 – Árvore de decisão do ano de 2002 do experimento 1. ....	149
Figura 26 – Árvore de decisão do ano de 2003 do experimento 1. ....	151
Figura 27 – Árvore de decisão do ano de 2004 do experimento 1. ....	152
Figura 28 – Árvore de decisão do ano de 2005 do experimento 1. ....	152
Figura 29 – Árvore de decisão do ano de 2006 do experimento 1. ....	153
Figura 30 - Árvore de decisão do ano de 2007 do experimento 1. ....	154
Figura 31 – Árvore de decisão do ano de 2008 do experimento 1. ....	155
Figura 32 – Árvore de decisão do ano de 2009 do experimento 1. ....	155
Figura 33 – Árvore de decisão do ano de 2010 do experimento 1. ....	156
Figura 34 – Árvore de decisão do ano de 2011 do experimento 1. ....	157
Figura 35 – Árvore de decisão do ano de 2012 do experimento 1. ....	158
Figura 36 – Árvore de decisão do ano de 2013 do experimento 1. ....	159
Figura 37 – Árvore de decisão do ano de 2014 do experimento 1. ....	160
Figura 38 – Árvore de decisão do ano de 2015 do experimento 1. ....	161
Figura 39 – Árvore de decisão do ano de 2016 do experimento 1. ....	162
Figura 40 – Árvore de decisão do ano de 2017 do experimento 1. ....	163
Figura 41 – Árvore de decisão do ano de 2002 do experimento 2. ....	164
Figura 42 – Árvore de decisão do ano de 2003 do experimento 2. ....	164
Figura 43 – Árvore de decisão do ano de 2004 do experimento 2. ....	164
Figura 44 – Árvore de decisão do ano de 2005 do experimento 2. ....	164
Figura 45 – Árvore de decisão do ano de 2006 do experimento 2. ....	164
Figura 46 – Árvore de decisão do ano de 2008 do experimento 2. ....	165
Figura 47 – Árvore de decisão do ano de 2009 do experimento 2. ....	165
Figura 48 – Árvore de decisão do ano de 2010 do experimento 2. ....	165

Figura 49 – Árvore de decisão do ano de 2011 do experimento 2.....	165
Figura 50 – Árvore de decisão do ano de 2012 do experimento 2.....	165
Figura 51 – Árvore de decisão do ano de 2014 do experimento 2.....	165
Figura 52 – Árvore de decisão do ano de 2015 do experimento 2.....	166
Figura 53 – Árvore de decisão do ano de 2016 do experimento 2.....	166
Figura 54 – Árvore de decisão do ano de 2017 do experimento 2.....	167

## LISTA DE QUADROS

Quadro 1 – Critérios de risco considerados pelo município de Curitiba para classificação de risco ao nascer.....	29
Quadro 2 – Atividades de pré-processamento e respectiva descrição.....	33
Quadro 3 – Mapeamento entre referencial teórico e encaminhamentos metodológicos. ....	39
Quadro 4 – Atividades de pré-processamento realizadas e exemplos.....	44
Quadro 5 – Experimentos de mineração de dados e características dos subconjuntos utilizadas (variáveis SINASC, variáveis SIM, registros e disposição dos anos). ....	46
Quadro 6 – Variáveis encontradas mais fortemente relacionadas ao desfecho, respectivas camadas do modelo dos DSS, experimento em que foram encontradas, teste estatístico e nível de significância. ....	65
Quadro 7 – Variáveis do SIM utilizadas nos experimentos de KDD, com respectiva descrição, valores, camada aproximada dos DSS e experimento utilizado .....	131
Quadro 8 – Variáveis do SINASC utilizadas nos experimentos de KDD, com respectiva descrição, valores, camada aproximada dos DSS e experimento utilizado .....	132
Quadro 9 – Variáveis SIM não utilizadas nos experimentos de KDD, com respectiva descrição e justificativa. ....	137
Quadro 10 – Variáveis SINASC não utilizadas nos experimentos de KDD, com respectiva descrição e justificativa. ....	140

## LISTA DE TABELAS

Tabela 1 – Número de registros por ano e SIS, com estratificação dos registros do SIM por grupo de evitabilidade do óbito. ....	42
Tabela 2 – Perfil da população estudada do conjunto de dados SIM-SINASC final, com variáveis e respectiva frequência absoluta e relativa. ....	52
Tabela 3 – Número de registros no conjunto de dados SIM-SINASC final por ocorrência do desfecho (óbito e não óbito) por ano, com estratificação dos óbitos por grupo de evitabilidade. ....	55
Tabela 4 – Frequência absoluta e relativa dos nascidos vivos entre 2008 e 2017 que foram a óbito, por DS e classificação de risco ao nascer. ....	63
Tabela 5 – Frequência absoluta e relativa dos nascidos vivos entre 2006 e 2017, por vinculação ao SUS e ocorrência do óbito. ....	63
Tabela 6 – Frequência absoluta de todos os grupos de óbito e frequência relativa dos grupos de óbito em relação a todos os nascidos vivos e nascidos vivos que foram a óbito. ....	64

## LISTA DE ABREVIATURAS E SIGLAS

AIG	Adequado para a Idade Gestacional
CBO	Código Brasileiro de Ocupação
CE	Centro de Epidemiologia
CID-10	Classificação Internacional de Doenças e Problemas Relacionados à Saúde
CNES	Cadastro Nacional de Estabelecimento em Saúde
COOSMIC	Coorte de Saúde Materno-infantil de Curitiba
CEP	Comitê de Ética em Pesquisa
DNV	Declaração de Nascido Vivo
DO	Declaração de Óbito
DRE	Descobridor de Regras de Exceção
DS	Distrito Sanitário
DSS	Determinantes Sociais da Saúde
GIG	Grande para a Idade Gestacional
IBGE	Instituto Brasileiro de Geografia e Estatística
IDHM	Índice de Desenvolvimento Humano Municipal
IVAB	Índice de Vulnerabilidade das Áreas de Abrangência das Unidades Municipais de Saúde
IVF-PR	Índice de Vulnerabilidade das Famílias do Paraná
KDD	<i>Knowledge Discovery in Database</i>
MD	Mal Definidas
MS	Ministério da Saúde
NCO	Não Classificadas/Outras
NE	Não Evitáveis
Pibic	Projeto Institucional de Bolsas de Iniciação Científica
PIG	Pequeno para a Idade Gestacional
PUCPR	Pontifícia Universidade Católica do Paraná
RAAP	Redutíveis por Adequada Atenção ao Parto
RACG	Redutíveis por Adequado Controle na Gravidez
RAPDTP	Redutíveis por Ações de Prevenção, Diagnóstico e Tratamento Precoces
RAS	Rede de Atenção à Saúde

RI	Redutíveis por Imunoprevenção
RN	Recém-Nascido
RPOS	Redutíveis por Parcerias com Outros Setores
Seade	Sistema Estadual de Análise de Dados
SIM	Sistema de Informações sobre Mortalidade
SINASC	Sistema de Informações sobre Nascidos Vivos
SIS	Sistema de Informação em Saúde
SMS	Secretaria Municipal de Saúde
SUS	Sistema Único de Saúde
TMI	Taxa de Mortalidade Infantil
UMS	Unidade Municipal de Saúde
Unicef	United Nations Children's Fund
WEKA	<i>Waikato Environment for Knowledge Analysis</i>



## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>18</b>
1.1	OBJETIVOS.....	20
1.1.1	<b>Objetivo geral</b> .....	<b>20</b>
1.1.2	<b>Objetivos específicos</b> .....	<b>21</b>
1.2	CONTRIBUIÇÃO CIENTÍFICA E SOCIAL .....	21
<b>2</b>	<b>REFERENCIAL TEÓRICO</b> .....	<b>22</b>
2.1	PRIMEIROS 1.000 DIAS DE VIDA E DETERMINANTES SOCIAIS DA SAÚDE 22	
2.2	SISTEMAS DE INFORMAÇÃO EM SAÚDE .....	25
2.2.1	<b>Sistema de informações sobre nascidos vivos</b> .....	<b>26</b>
2.2.2	<b>Sistema de informações sobre mortalidade</b> .....	<b>30</b>
2.3	KNOWLEDGE DISCOVERY IN DATABASE .....	32
2.3.1	<b>Pré-processamento</b> .....	<b>33</b>
2.3.2	<b>Mineração de dados</b> .....	<b>34</b>
2.3.3	<b>Pós-processamento</b> .....	<b>37</b>
<b>3</b>	<b>MAPEAMENTO ENTRE O REFERENCIAL TEÓRICO E OS ENCAMINHAMENTOS METODOLÓGICOS</b> .....	<b>39</b>
<b>4</b>	<b>ENCAMINHAMENTOS METODOLÓGICOS</b> .....	<b>41</b>
4.1	NATUREZA, POPULAÇÃO E CENÁRIO DA PESQUISA .....	41
4.2	CARACTERÍSTICA DOS CONJUNTOS DE DADOS .....	42
4.3	ETAPAS DA PESQUISA .....	43
4.3.1	<b>Etapa 1 – revisão integrativa</b> .....	<b>43</b>
4.3.2	<b>Etapa 2 – KDD</b> .....	<b>43</b>
4.3.3	<b>Etapa 3 – análise dos resultados</b> .....	<b>49</b>
<b>5</b>	<b>RESULTADOS</b> .....	<b>50</b>
5.1	REVISÃO INTEGRATIVA .....	50
5.2	PRÉ-PROCESSAMENTO .....	51
5.3	MINERAÇÃO DE DADOS E PÓS-PROCESSAMENTO.....	56
<b>6</b>	<b>DISCUSSÃO</b> .....	<b>66</b>
6.1	QUAIS VARIÁVEIS DE GOVERNABILIDADE DIRETA DO GESTOR TÊM POTENCIAL DE SUBSIDIAR O DIRECIONAMENTO DE AÇÕES PARA REDUÇÃO	

DOS ÓBITOS EVITÁVEIS DE ZERO A DOIS ANOS DE IDADE NO MUNICÍPIO DE CURITIBA?.....	66
6.2 LIMITAÇÕES.....	71
6.3 PERSPECTIVAS DE TRABALHOS FUTUROS .....	74
<b>7 CONSIDERAÇÕES FINAIS .....</b>	<b>76</b>
<b>APÊNDICE A – ARTIGO DE REVISÃO INTEGRATIVA SOBRE VARIÁVEIS ASSOCIADAS À MORTALIDADE MATERNO-INFANTIL.....</b>	<b>87</b>
<b>APÊNDICE B – ARTIGO DE REVISÃO INTEGRATIVA SOBRE MINERAÇÃO DE DADOS NO APOIO À GESTÃO EM SAÚDE MATERNO-INFANTIL .....</b>	<b>106</b>
<b>APÊNDICE C – VARIÁVEIS UTILIZADAS NOS EXPERIMENTOS DE KDD, COM RESPECTIVA DESCRIÇÃO, VALORES, CAMADA APROXIMADA DOS DSS E EXPERIMENTO UTILIZADO, POR SIS.....</b>	<b>131</b>
<b>APÊNDICE D – VARIÁVEIS NÃO UTILIZADAS NOS EXPERIMENTOS DE KDD, COM RESPECTIVA DESCRIÇÃO E JUSTIFICATIVA, POR SIS .....</b>	<b>137</b>
<b>APÊNDICE E – LINKAGE – DESCRIÇÃO DETALHADA E SCRIPT REALIZADO NO R .....</b>	<b>142</b>
<b>APÊNDICE F – ÁRVORES DE DECISÃO DO EXPERIMENTO 1, POR ANO.....</b>	<b>145</b>
<b>APÊNDICE G – ÁRVORES DE DECISÃO DO EXPERIMENTO 2.....</b>	<b>164</b>
<b>ANEXO A – DNV .....</b>	<b>168</b>
<b>ANEXO B – DO.....</b>	<b>169</b>
<b>ANEXO C – PARECER CONSUBSTANCIADO DO CEP DA PUCPR .....</b>	<b>170</b>
<b>ANEXO D – PARECER CONSUBSTANCIADO DO CEP DA SMS DE CURITIBA.....</b>	<b>175</b>

## 1 INTRODUÇÃO

Os primeiros 1.000 dias de vida compreendem o período entre a concepção e os dois anos de idade (270 dias de gestação mais 730 dias entre zero e dois anos) (BLACK et al., 2008). O crescimento, desenvolvimento e maturação de órgãos e sistemas nesse período estabelecem padrões metabólicos, cognitivos e emocionais com repercussões ao longo da vida. A estatura, desenvolvimento cognitivo, motor e socioemocional, aprendizagem, *performance* escolar, capacidade laboral e produtividade, doenças crônicas e morbimortalidade na infância são impactados por situações ocorridas nesse período (ORGANIZAÇÃO MUNDIAL DA SAÚDE; COMISSÃO PARA OS DETERMINANTES SOCIAIS DA SAÚDE, 2010; BLACK et al., 2013; CUNHA; LEITE; ALMEIDA, 2015; VICTORA et al., 2015; MAMELI; MAZZANTINI; ZUCCOTTI, 2016; MOORE et al., 2017).

Escolaridade e idade materna, peso ao nascer, malformação congênita, Apgar no primeiro e quinto minuto, acesso a serviços de saúde e condições socioeconômicas e ambientais são fatores de risco à mortalidade infantil e demonstram a relação entre as condições de crescimento e desenvolvimento da gestação e a sobrevivência da criança para além desse período (BRASIL, 2012a; TEIXEIRA et al., 2016; KROPIWIEC; FRANCO; AMARAL, 2017). Os óbitos nos primeiros 1.000 dias de vida compreendem os óbitos fetais (não compreendidos neste estudo), de menores de um ano e de um a dois anos de idade.

Os óbitos podem ser expressos pela taxa de mortalidade, que indica a quantidade de óbitos por 1.000 nascidos vivos. Enquanto a Taxa de Mortalidade Infantil (TMI) representa os óbitos de menores de um ano (BRASIL, 2009a), os óbitos de um a dois anos não possuem indicador específico, estando contemplados na taxa de mortalidade na infância ou taxa de mortalidade em menores de cinco anos, que compreende óbitos de zero a cinco anos de idade (BRASIL, 2012b). A mortalidade na infância no mundo reduziu pela metade de 1990 (taxa de 93/1.000 nascidos vivos) a 2016 (41/1.000 nascidos vivos), porém com diferenças entre países desenvolvidos e subdesenvolvidos - nestes, a chance de óbito é 20 a 25 vezes maior que aqueles. Em 2017, a taxa de mortalidade na infância no Brasil foi

de 13,9/1.000 nascidos vivos, e em Curitiba, 9,8/1.000 nascidos vivos, a menor taxa dentre as capitais brasileiras<sup>1</sup>.

A TMI reflete as condições de desenvolvimento socioeconômico e infraestrutura ambiental, bem como o acesso e a qualidade da atenção à saúde materno-infantil (ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE, 2008). As menores taxas em 2016 foram inferiores a 2/1.000 nascidos vivos, identificadas em países desenvolvidos como Japão, Islândia e Singapura. A África Subsaariana concentrou as maiores taxas, com variação de 35,1 a 45,6/1.000 nascidos vivos. Em 2016 a TMI brasileira foi de 14,0/1.000 nascidos vivos (BRASIL, 2019), maior que de outros países da América do Sul (Equador – 11,2; Paraguai – 11,1; Peru – 7,2; Argentina – 6,2) (UNITED NATIONS CHILDREN'S FUND, 2018).

Essa diferença da TMI entre os países desenvolvidos e em desenvolvimento também pode ser notada entre regiões do Brasil, reforçando situações de iniquidade em saúde. Em 2017, a TMI brasileira reduziu para 13,4/1.000 nascidos vivos. As regiões Norte e Nordeste obtiveram as maiores taxas – 17,3 e 15,8/1.000 nascidos vivos, respectivamente –, enquanto as regiões Sudeste e Sul apresentaram as menores – respectivamente, 11,7 e 10,1/1.000 nascidos vivos. Dentre os estados brasileiros e Distrito Federal, Amapá obteve a maior TMI (23,0) e Santa Catarina, a menor (9,9). A TMI do Paraná (10,4) ficou acima da do Rio Grande do Sul (10,1) e Santa Catarina (BRASIL, 2019). Em Curitiba, a TMI foi de 16,6/1.000 nascidos vivos no ano de 1998 e, em 2018, foi de 8,2/1.000 nascidos vivos, apresentando tendência de queda (CURITIBA, 2019a).

A redução da TMI no município é reflexo de melhores condições socioeconômicas e de esforços na atenção à saúde da gestante, como a captação precoce para início do pré-natal (84% das gestantes iniciam o acompanhamento antes do quarto mês de gestação), o número médio de consultas do pré-natal (7,93 consultas por gestante) e a redução da gravidez na adolescência (de 19,3%, em 1999, para 14,2%, em 2010) (CURITIBA, 2018a). A TMI em Curitiba é considerada baixa, o que exige maiores esforços para sua redução, com concentração para os óbitos por causas evitáveis (CURITIBA, 2018a), que representaram 67% dos óbitos

---

<sup>1</sup> taxa calculada pela autora, pelo método direto, obtida pelo quociente de óbitos de menores de menos de cinco anos de idade e nascidos vivos, multiplicado por 1.000. Dados consultados no TabNet DATASUS, disponível em <http://www2.datasus.gov.br/DATASUS/index.php?area=0205>, obtendo-se os seguintes dados de residentes em Curitiba no ano de 2017: 222 óbitos de crianças menores de cinco anos de idade e 22.745 nascidos vivos.

de menores de um ano entre 1999 e 2017<sup>2</sup>. Segundo a United Nations Children's Fund (2017, 2018), apesar do progresso brasileiro em relação à mortalidade infantil, bebês ainda morrem por causas evitáveis, sendo um desafio a sua redução.

Identificar e controlar as influências no desenvolvimento da criança nos primeiros 1.000 dias de vida pode auxiliar o gestor no enfrentamento do problema, uma vez que as condições em que as pessoas nascem, crescem, vivem, trabalham e envelhecem aumentam ou diminuem sua vulnerabilidade frente a situações de saúde (ORGANIZAÇÃO MUNDIAL DA SAÚDE; COMISSÃO PARA OS DETERMINANTES SOCIAIS DA SAÚDE, 2010; MOORE et al., 2017). A identificação de influências relacionadas a um desfecho pode se dar pela mineração de dados, uma das etapas do processo KDD (*Knowledge Discovery in Database*<sup>3</sup>). A partir de um conjunto de dados, o KDD permite a extração de padrões que, se válidos, novos, potencialmente úteis e compreensíveis, podem complementar o conhecimento existente (FAYYAD et al., 1996).

Considerando o desafio da redução da mortalidade infantil, especialmente os óbitos evitáveis, e a prioridade de garantia de atenção a todas as crianças nos primeiros dois anos de vida estabelecida no Plano Municipal de Saúde 2018-2021 de Curitiba (CURITIBA, 2018a), objetivou-se descobrir os padrões<sup>4</sup> de óbito de zero a dois anos de idade em uma coorte histórica, com identificação das variáveis relacionadas à sua evitabilidade. Assim, a questão norteadora desta pesquisa foi: quais variáveis de governabilidade direta do gestor têm potencial de subsidiar o direcionamento de ações para redução dos óbitos evitáveis de zero a dois anos de idade no município de Curitiba?

## 1.1 OBJETIVOS

### 1.1.1 Objetivo geral

Descobrir padrões de óbito de zero a dois anos de idade em coorte histórica do município de Curitiba.

---

<sup>2</sup> Proporção calculada pela autora, com dados consultados no TabNet DATASUS, disponível em: <http://www2.datasus.gov.br/DATASUS/index.php?area=02>.

<sup>3</sup> Extração de conhecimento em base de dados

<sup>4</sup> Entende-se por padrão uma regra composta por uma combinação de variáveis e valores levando a determinado desfecho.

### 1.1.2 Objetivos específicos

- a) Identificar variáveis associadas ao óbito materno-infantil por meio de revisão integrativa;
- b) Identificar o potencial da mineração de dados para apoio à gestão em saúde materno-infantil por meio de revisão integrativa;
- c) Identificar novas variáveis de governabilidade direta da gestão, relacionadas à evitabilidade do óbito entre zero e dois anos de idade, por meio da mineração de dados.

### 1.2 CONTRIBUIÇÃO CIENTÍFICA E SOCIAL

Contribui-se cientificamente pela identificação de variáveis relacionadas ao óbito distintas do usualmente analisadas e consolidadas na literatura.

Em relação à contribuição social, os resultados podem subsidiar os gestores na tomada de decisão com vistas à redução de iniquidades em saúde e à melhora da qualidade da atenção à saúde materno-infantil.

## 2 REFERENCIAL TEÓRICO

Nesta seção, discute-se a literatura sobre os primeiros 1.000 dias de vida e a importância do cuidado nessa fase, relacionando-os com o modelo dos Determinantes Sociais da Saúde (DSS). Posteriormente, são abordados o Sistema de Informações sobre Nascidos Vivos (SINASC) e o Sistema de Informações sobre Mortalidade (SIM). Por último, descrevem-se o processo de *Knowledge Discovery in Database* (KDD) e suas três etapas – pré-processamento, mineração de dados e pós-processamento.

### 2.1 PRIMEIROS 1.000 DIAS DE VIDA E DETERMINANTES SOCIAIS DA SAÚDE

Os primeiros 1.000 dias de vida correspondem a 270 dias de gestação mais 730 dias de zero dois anos de idade (BLACK et al., 2008). Nesse período, há um rápido crescimento, desenvolvimento e maturação de órgãos e sistemas, com estabelecimento de padrões metabólicos, cognitivos e emocionais. Situações ocorridas nessa fase desempenham papel fundamental no processo de formação e desenvolvimento da criança, com impactos na vida adulta e terceira idade (CUNHA; LEITE; ALMEIDA, 2015; MOORE et al., 2017). Criminalidade, analfabetismo e problemas de saúde pública, como obesidade, doenças cardíacas e problemas de saúde mental, têm sua origem nos primeiros anos de vida (ORGANIZAÇÃO MUNDIAL DA SAÚDE; COMISSÃO PARA OS DETERMINANTES SOCIAIS DA SAÚDE, 2010; VICTORA et al., 2015; MOORE et al., 2017).

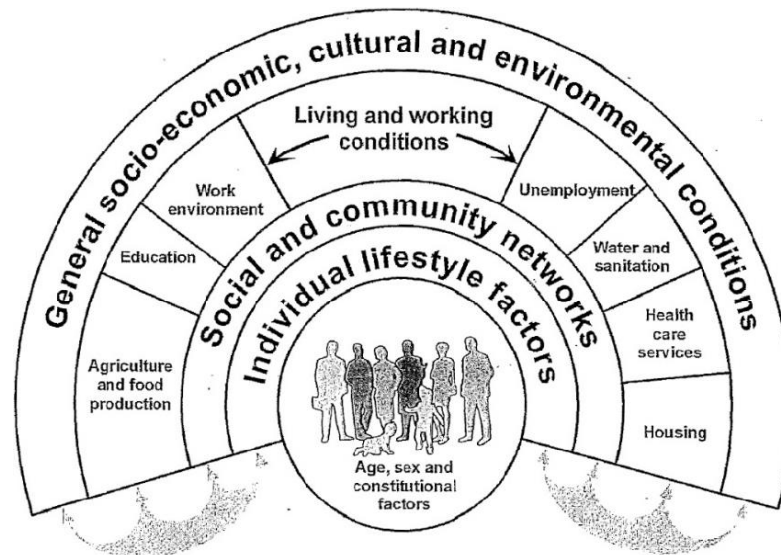
O estabelecimento de ações favorecendo a nutrição e o desenvolvimento fetal e infantil adequados resultam em benefícios na saúde ao longo da vida: redução da morbidade e mortalidade na infância, da obesidade e de doenças crônicas; aumento da estatura, do desenvolvimento cognitivo, motor e socioemocional, da *performance* escolar, da capacidade de aprendizagem e da capacidade de trabalho e produtividade (ORGANIZAÇÃO MUNDIAL DA SAÚDE; COMISSÃO PARA OS DETERMINANTES SOCIAIS DA SAÚDE, 2010; BLACK et al., 2013; CUNHA; LEITE; ALMEIDA, 2015; VICTORA et al., 2015; MAMELI; MAZZANTINI; ZUCCOTTI, 2016; MOORE et al., 2017).

O crescimento e desenvolvimento infantil nos primeiros 1.000 dias de vida são afetados pela nutrição da mãe, disponibilidade de nutrientes, segurança alimentar,

recursos dos cuidadores, condições socioeconômicas e ambientais, entre outros fatores (ORGANIZAÇÃO MUNDIAL DA SAÚDE; COMISSÃO PARA OS DETERMINANTES SOCIAIS DA SAÚDE, 2010; MOORE et al., 2017). Evidencia-se a relação entre as influências individuais e coletivas, em que as condições gerais (socioeconômicas e ambientais) influenciam e modulam as individuais (nutrição da mãe, recursos dos cuidadores, escolhas individuais etc.). A relação entre essas condições pode ser representada pelo modelo dos DSS proposto por Dahlgren e Whitehead (1991).

O modelo dos DSS representa, em cinco camadas concêntricas, as principais influências na saúde e remete aos diferentes graus de governabilidade do indivíduo sobre sua saúde e respectivos níveis de intervenção por parte dos gestores (Figura 1). A camada mais interna, de características individuais, como raça/cor, idade, sexo e hereditariedade, influencia a saúde materno-infantil, porém elas são consideradas características fixas, sobre as quais há limitada governabilidade e possibilidade de mudança.

Figura 1 – Modelo dos DSS



Fonte: Dahlgren e Whitehead (1991, p. 11).

Estilo de vida vem em seguida e diz respeito às escolhas e comportamentos dos indivíduos. Deve-se considerar que, nesse período, as crianças são dependentes e sofrem os resultados de escolhas e estímulos realizados por terceiros (CUNHA; LEITE; ALMEIDA, 2015); portanto, escolhas saudáveis por parte



de pais e cuidadores aumentam as chances de desenvolvimento saudável das crianças (MOORE et al., 2017). Nesse sentido, gestores e profissionais em saúde devem levar em consideração o poder de escolha do indivíduo e as ações de educação em saúde devem influenciar o estilo de vida e mudança de atitudes (DAHLGREN; WHITEHEAD, 1991).

As escolhas individuais são influenciadas pelas redes sociais e comunitárias, representadas pela próxima camada. O suporte da família, amigos, vizinhos e comunidade também influencia o enfrentamento ou prevenção de doenças. Cunha, Leite e Almeida (2015) e Moore et al. (2017) destacam a importância das relações sociais e formação de vínculos entre crianças e suas famílias para o bom crescimento e desenvolvimento infantil. O suporte social à gestante reduz o estresse, depressão, comportamentos de risco e as chances de malcuidado da criança após o nascimento. Reconhecer a força do grupo no enfrentamento de problemas de saúde e na prevenção de doenças deve ser premissa presente nas ações dos gestores (DAHLGREN; WHITEHEAD, 1991).

A próxima camada representa as condições de vida e trabalho, que inclui o acesso a serviços sociais e de saúde, além de moradia, emprego, educação, acesso a alimentos e ambientes saudáveis. A moradia adequada promove segurança e autonomia à família, com impactos na saúde e bem-estar da família e no desenvolvimento físico, cognitivo e emocional da criança (MOORE et al., 2017). O acesso da gestante a alimentos saudáveis é importante para o desenvolvimento fetal e infantil (SAVE THE CHILDREN, 2012). Ações de educação em saúde que incentivem mudanças de comportamentos são mais efetivas na medida em que o ambiente favorece a opção por escolhas mais saudáveis. As intervenções relacionadas a essa camada são de governabilidade direta do poder público e devem se dar por estratégias e políticas públicas intersetoriais, pois envolvem outros setores, além da saúde (DAHLGREN; WHITEHEAD, 1991).

A camada mais externa, de condições socioeconômicas, culturais e ambientais gerais, influencia todas as outras. Nos países de baixa renda, as crianças possuem maior risco de desnutrição crônica na infância (BLACK et al., 2013) e a desnutrição materna está vinculada ao acesso limitado a alimentos e serviços de saúde, condições precárias de trabalho e submissão feminina por costumes e tradições culturais (SAVE THE CHILDREN, 2012). Quanto mais baixa for a situação social da criança e família, piores serão os desfechos na sua saúde e

bem-estar, acentuando as iniquidades em saúde (OMS; CDSS, 2010; MOORE et al., 2017). As intervenções realizadas nesse nível devem promover mudanças estruturais por meio de estratégias econômicas internas (como políticas sobre imposto) e/ou acordos ambientais internacionais (DAHLGREN; WHITEHEAD, 1991).

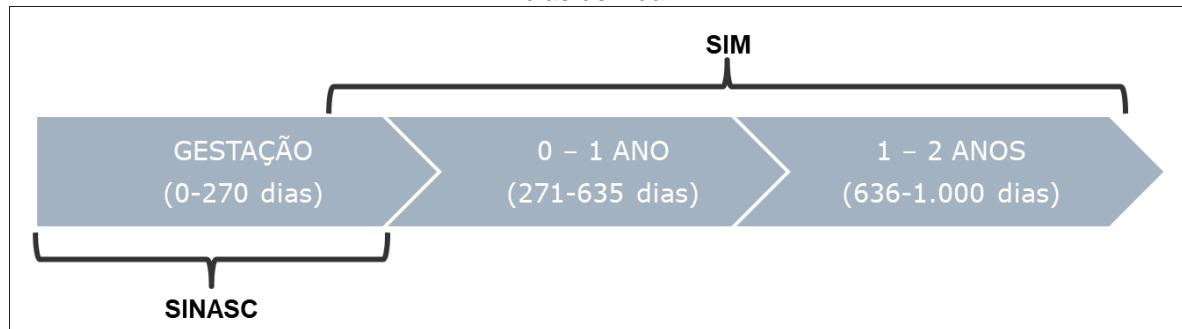
Apesar de as influências estarem dispostas em camadas, é forte a relação entre elas, sendo fundamental que as intervenções sejam transversais, de maneira que atinjam o nível individual e coletivo, com o intuito de reduzir iniquidades em saúde, além de promover um ótimo crescimento e desenvolvimento infantil (ORGANIZAÇÃO MUNDIAL DA SAÚDE; COMISSÃO PARA OS DETERMINANTES SOCIAIS DA SAÚDE, 2010; CUNHA; LEITE; ALMEIDA, 2015).

## 2.2 SISTEMAS DE INFORMAÇÃO EM SAÚDE

Os Sistemas de Informação em Saúde (SISs) são conjuntos de elementos organizados que possibilitam a coleta, armazenamento e processamento de dados. A análise dos dados proporciona informações para melhor compreensão da situação de saúde, subsidiando planejamento, organização e avaliação dos serviços de saúde. Proporciona informações para apoiar o processo decisório dos profissionais envolvidos na atenção em saúde, da decisão clínica local à formulação de políticas públicas (BRASIL, 2008; COELI et al., 2009; MARIN, 2010).

De acordo com Coeli et al. (2009), os SISs geram bancos de dados que podem ser classificados em epidemiológicos, administrativos e clínicos. Os epidemiológicos incluem eventos vitais, doenças e agravos de notificação compulsória e os administrativos, informações contábeis e de produção dos serviços (procedimentos e atendimentos realizados). Ambos são usados para fins de vigilância, avaliação e pesquisa. O banco de dados clínicos armazena informações clínicas, de natureza multimídia (exames de imagem) ou não (resultados laboratoriais), que são geradas ao longo do cuidado do usuário no seu curso de vida. Os SISs escolhidos para esta pesquisa abrangem aqueles com informações sobre os primeiros 1.000 dias de vida e são de natureza epidemiológica (Figura 2).

Figura 2 – Representação dos SISs utilizados na pesquisa em relação às fases dos primeiros 1.000 dias de vida.



Fonte: A autora (2020).

### 2.2.1 Sistema de informações sobre nascidos vivos

O SINASC foi implantado pelo Ministério da Saúde (MS) em 1990, com o objetivo de coletar dados sobre os nascimentos em todo o território nacional, a partir dos quais é possível extrair informações sobre os nascidos vivos de forma a subsidiar intervenções relacionadas à saúde materno-infantil e planejamento de gestão dos serviços de saúde (BRASIL, 2011a, 2018).

Para todos os nascimentos que apresentem sinais vitais, deve ser preenchida a Declaração de Nascido Vivo (DNV), disponibilizada no Anexo A (BRASIL, 2011a), que deve ser digitada no SINASC. A DNV é individual e padronizada nacionalmente, apresentando numeração única atribuída pelo MS para controle de sua distribuição. É emitida em três vias: uma encaminhada à Secretaria Municipal de Saúde (SMS) do município de ocorrência do nascimento pelo estabelecimento ou profissional que realizou o parto, outra entregue ao pai/responsável legal para obtenção de Certidão de Nascimento junto ao Cartório de Registro Civil e uma arquivada no prontuário da mãe pelo estabelecimento de saúde ou profissional de saúde que realizou o parto. A digitação da DNV no SINASC é de responsabilidade da SMS do município de ocorrência do nascimento (BRASIL, 2011a) e, em Curitiba, realizada pelo Centro de Epidemiologia (CE) da SMS de Curitiba.

A DNV passou por modificações e as variáveis dos oito blocos de dados da atual versão são identificadas no SINASC desde 2013 (BRASIL, 2009b, 2011b).

**Bloco I – Identificação do Recém-Nascido (RN).** Dados inerentes ao RN, incluindo número da DNV, nome, data e hora do nascimento, sexo ao nascer, peso ao nascer, índice de Apgar no primeiro e quinto minuto e malformação congênita.

Número de DNV, data de nascimento e sexo ao nascer são dados importantes, pois possibilitam a busca do RN em outras bases de dados.

O peso ao nascer, quando inferior a 2.500 g, é considerado RN de baixo peso; abaixo de 1.500 g, muito baixo peso, e, quando menos de 1.000 g, extremo baixo peso (UNITED NATIONS CHILDREN'S FUND; WORLD HEALTH ORGANIZATION, 2004). O sexo do RN e peso ao nascer são considerados em conjunto com semanas de gestação (bloco V) para avaliar o crescimento do RN.

O índice de Apgar é o resultado do somatório de cinco sinais avaliados na criança ao nascer: frequência cardíaca, respiração, tônus muscular, cor da pele e presença de reflexos (APGAR, 1953, 1958). Essa avaliação é realizada no primeiro e quinto minuto após o nascimento e seu valor representa as dificuldades e condições do nascimento. Quando de 0 a 2, indica dificuldade grave; de 3 a 4, dificuldade moderada; de 5 a 7, leve dificuldade; e, de 8 a 10, que o bebê nasceu em ótimas condições. Apgar acima de 7 no quinto minuto está relacionado ao bom prognóstico neonatal, porém, de acordo com Jeganathan et al. (2017), existem poucos estudos relacionados à recuperação do Apgar entre o primeiro e o quinto minuto. Além das condições do RN, esse intervalo pode indicar informações sobre a assistência prestada.

A malformação congênita (ou anomalia) compreende qualquer malformação física ou funcional desenvolvida durante a fase intrauterina, sendo condição frequente relacionada ao óbito, principalmente ao neonatal (WORLD HEALTH ORGANIZATION, 2015), sua detecção no momento do nascimento deve ser assinalada com preenchimento do bloco VI, específico de malformação congênita.

**Bloco II – Local da ocorrência.** Dados sobre ocorrência do parto em hospital, outros estabelecimentos de saúde, domicílio ou outro. Quando estabelecimento de saúde, registra-se o número no Cadastro Nacional de Estabelecimento em Saúde (CNES), por meio do qual é possível identificar se a gestante é vinculada ao Sistema Único de Saúde (SUS). Salienta-se que essa possibilidade se deve à organização da Rede de Atenção à Saúde (RAS) da gestante em Curitiba, uma vez que as maternidades do SUS no município não atendem a outras formas de convênio. Este bloco contém ainda o endereço, município e unidade da federação onde ocorreu o nascimento.

**Bloco III – Mãe.** Neste bloco, constam dados pessoais da mãe do RN (nome, cartão SUS, data de nascimento, residência) e características sociodemográficas

(escolaridade, ocupação, naturalidade, situação conjugal, raça/cor). O nome da mãe possibilita a integração com outras bases e as características sociodemográficas são importantes para inferir as condições nas quais provavelmente ocorrerão os primeiros 1.000 dias de vida. Pelo endereço, é possível identificar o Distrito Sanitário (DS) e Unidade Municipal de Saúde (UMS) responsáveis pelo território onde moram a gestante e RN.

**Bloco IV – Pai.** Dados de identificação paternos: nome e idade.

**Bloco V – Gestação e parto.** Dados sobre as gestações anteriores (quantidade de gestação anterior, número de cesárea e de partos vaginais, número de nascidos vivos e perdas fetais/aborto), gestação atual (número de semanas e método utilizado para estimar as semanas, início do pré-natal e número de consultas de pré-natal) e parto (apresentação, indução de trabalho de parto, tipo de parto, ocorrência de cesárea antes do trabalho de parto e profissional que assistiu o parto).

O número de semanas em conjunto com sexo e peso ao nascer (bloco I), permite avaliar o tamanho do RN no momento do nascimento, identificando-o em Pequeno para a Idade Gestacional (PIG) quando percentil abaixo de 10, Adequado para a Idade Gestacional (AIG) quando percentil de 10 a 90 e Grande para a Idade Gestacional (GIG) quando percentil acima de 90. A classificação de RNs PIG e GIG pode sinalizar problemas no crescimento e desenvolvimento na gestação, além de fornecer informação que pode ser utilizada para orientar a atenção, o cuidado e o direcionamento dos recursos em saúde.

Dentre as curvas de crescimento existentes, foram utilizadas a proposta por Fenton e a Intergrowth. A primeira foi elaborada a partir de metanálise de estudos realizados na Alemanha, Austrália, Canadá, Escócia, Estados Unidos e Itália entre 1987 e 2012 (FENTON; KIM, 2013). A segunda foi obtida por amostras representativas de Pelotas (Brasil), Turin (Itália), Muscat (Oman), Oxford (Reino Unido), Seattle (Estados Unidos), Shunyi County (China), Nagpur (Índia) e Parklands (Quênia) e levou em consideração que RNs prematuros provavelmente sofreram interferências na gestação, obtendo-se uma curva para RN com idade inferior a 33 semanas (PAPAGEORGHIU et al., 2014) e outra para RN entre 33 e 42 semanas (VILLAR et al., 2014).

**Bloco VI – Anomalia congênita (malformação).** De preenchimento obrigatório quando há sinalização no bloco I de detecção de malformação congênita no nascimento. Campo descritivo, codificado pelo CE da SMS de Curitiba, conforme

Classificação Internacional de Doenças e Problemas Relacionados à Saúde (CID-10) para digitação no SINASC.

**Bloco VII – Preenchimento.** Bloco de identificação do profissional responsável pelo preenchimento da DNV, contendo nome, função e documento de identificação.

**Bloco VIII – Cartório.** Bloco de preenchimento exclusivo do cartório, com dados de onde foi registrado o nascimento.

A partir da DNV, o CE da SMS de Curitiba identifica os RNs de risco ao nascer e preconiza a vigilância prioritária desse grupo (CURITIBA, 2019b). Considera-se RN de risco ao nascer aquele exposto a situações mais propensas a desfechos indesejáveis. A partir dessas situações, o MS sugere como critérios para a classificação de risco (BRASIL, 2014a): baixo nível socioeconômico, história de morte de criança menor de cinco anos na família, criança explicitamente indesejada, idade materna menor que 20 anos, duração da gestação menor que 37 semanas, peso ao nascer menor que 2.500 g e escolaridade da mãe menor que oito anos de estudo (BRASIL, 2014a). A definição dos critérios de risco ao nascer em Curitiba baseou-se em estudos nacionais e internacionais, e não contempla todos os critérios sugeridos pelo MS. Os critérios considerados em Curitiba são agrupados em critérios de risco isolado ou de risco associado. Para o RN ser classificado como de risco ao nascer, é necessária a identificação de pelo menos um fator isolado ou dois ou mais fatores associados, conforme Quadro 1 (CURITIBA, 2019b).

Quadro 1 – Critérios de risco considerados pelo município de Curitiba para classificação de risco ao nascer

Risco isolado	Risco associado
Apgar no quinto minuto < 7	Idade materna entre 16 e 19 anos
Idade gestacional < 37 semanas	Instrução materna < 1º grau
Peso ao nascer < 2.500 g	Filhos vivos de gestações anteriores > 2
Nº consultas de pré-natal < 4	Filhos mortos de gestações anteriores > 0
Idade materna < 16 anos	

Fonte: Adaptado de Curitiba (2019b).

Salienta-se que não existe um campo específico na DNV para informação do risco ao nascer, porém o CE da SMS de Curitiba codifica esse risco com o código do bairro de residência, que possui de quatro a cinco dígitos, em que o primeiro indica o

risco ao nascer se número ímpar, os dois seguintes, o código do bairro e os últimos, o código do DS. Por exemplo, os códigos 3759 e 2759 referem-se a RN residente no bairro CIC (X75X), DS CIC (XXX9), sendo o primeiro RN de risco ao nascer (3XXX) e o segundo, de não risco ao nascer (2XXX).

### **2.2.2 Sistema de informações sobre mortalidade**

O SIM foi desenvolvido pelo MS em 1975 e unificou mais de 40 modelos de instrumentos de coleta de dados sobre mortalidade utilizados no país. Foi informatizado em 1979, sendo considerado importante ferramenta de gestão, pois as informações possíveis de extração a partir dos dados subsidiam o estabelecimento de prioridades para as políticas públicas por meio de estatísticas de mortalidade, além de permitir a comparação do perfil epidemiológico brasileiro com o de outros países (BRASIL, 2011b, 2019).

O documento para registro dos óbitos é a Declaração de Óbito (DO), padronizada em território nacional em 1976. Possui sequência numérica única para controle de emissão e distribuição, e seu preenchimento é de responsabilidade ética e jurídica do profissional médico nas seguintes condições (BRASIL, 2011b, p. 8):

Em todo óbito por causa natural ou por causa acidental e/ou violenta;  
No óbito fetal, se a gestação teve duração igual ou superior a 20 semanas, ou o feto teve peso corporal igual ou superior a 500 gramas e/ou estatura igual ou superior a 25 centímetros;  
No óbito não fetal, quando a criança nascer viva e morrer logo após o nascimento, independentemente da duração da gestação, do peso do recém-nascido e do tempo que tenha permanecido vivo.

A DO (Anexo B) possui três vias: uma enviada para a SMS do município de ocorrência do óbito, outra utilizada para obtenção da Certidão de Óbito junto ao Cartório de Registro Civil e uma cujo destino varia de acordo com as características do óbito e local de ocorrência (BRASIL, 2011b). Sua digitação no SIM é de responsabilidade da SMS do município de ocorrência, sendo composta por nove blocos.

**Bloco I – Identificação.** Contém dados de identificação e sociodemográficos do falecido. A data de nascimento, nome e nome da mãe são importantes para integração com outras bases de dados. Os dados demográficos referem-se à idade, sexo, raça/cor, situação conjugal, escolaridade e ocupação.

**Bloco II – Residência.** Dados do local de residência do falecido, permitindo inferir as condições socioambientais em que vivia.

**Bloco III – Ocorrência.** Dados sobre local de ocorrência do óbito: em hospital, outros estabelecimentos de saúde, domicílio, via pública ou outro. Quando estabelecimento de saúde, identifica-se pelo seu número no CNES. O endereço da ocorrência só é preenchido quando o óbito ocorre em via pública. Constam ainda o município e a unidade da federação onde ocorreu o óbito.

**Bloco IV – Óbito fetal ou menor de um ano.** Bloco exclusivo para registro desses tipos de óbito. Contém dados sobre a mãe (escolaridade, Código Brasileiro de Ocupação – CBO, número de filhos tidos e fetais e abortos), gestação e parto (semanas de gestação, tipo de gravidez e parto, peso ao nascer e número da DNV). O número da DNV como variável na DO permite a identificação do respectivo registro de nascimento no SINASC.

**Bloco V – Condições e causas do óbito.** Destinado à qualificação das condições e causas que provocaram o óbito. Das condições, identificam-se o óbito materno, recebimento de assistência médica e confirmação de diagnóstico por necropsia. Como causas, identificam-se todas as doenças, estados mórbidos ou lesões que produziram ou contribuíram para a morte, sendo a causa básica considerada a doença ou lesão que deu início ao processo que culminou na morte (BRASIL, 2011b).

A partir da DO, inicia-se a investigação das causas do óbito, fundamental para direcionamento de ações<sup>5</sup>. Nesse sentido, ressalta-se a importância dos Comitês de Prevenção do Óbito Infantil e Fetal, que, além de investigar as causas, analisam a sua evitabilidade, com o propósito de prevenir mortes similares no futuro (BRASIL, 2009b). Após a investigação, o CE da SMS de Curitiba codifica as causas de acordo com o CID-10, por meio do qual é possível inferir a evitabilidade do óbito.

Existem diferentes classificações de evitabilidade do óbito. São recomendadas pelo MS: classificação de evitabilidade da Fundação Sistema Estadual de Análise de Dados (Seade), lista de mortes evitáveis por intervenções do SUS e classificação de Wigglesworth expandida (BRASIL, 2009b). Dentre essas, a classificação da Fundação Seade é capaz de distinguir os óbitos evitáveis de

---

<sup>5</sup> Em Curitiba, 100% dos óbitos infantis são investigados e analisados, sendo a causa básica definida após encerramento da investigação.



maneira mais precisa (DIAS; SANTOS NETO; ANDRADE, 2017) e possui oito grupos de evitabilidade do óbito: Redutíveis por Imunoprevenção (RI), Redutíveis por Adequado Controle na Gravidez (RACG), Redutíveis por Adequada Atenção ao Parto (RAAP), Redutíveis por Ações de Prevenção, Diagnóstico e Tratamento Precoces (RAPDTP), Redutíveis por Parcerias com Outros Setores (RPOS), Não Evitáveis (NE), Mal Definidas (MD) e Não Classificadas/Outras (NCO) (ORTIZ, 2000 *apud* BRASIL, 2009b).

**Bloco VI – Médico.** Dados de identificação do médico que atestou o óbito (nome, número no Conselho Regional de Medicina, meio de contato, data e assinatura).

**Bloco VII – Causas externas.** De caráter epidemiológico, complementa o bloco V quando é óbito provavelmente não natural: acidente, suicídio e homicídio. Contém dados do local de ocorrência, se foi acidente de trabalho e campo para descrição sumária do evento.

**Bloco VIII – Cartório.** Bloco de preenchimento exclusivo do cartório, com dados sobre onde foi registrado o óbito.

**Bloco IX – Localidade sem médico.** Preenchido pelo Cartório de Registro Civil no caso de óbito ocorrido em localidade sem médico para atestá-lo.

### 2.3 KNOWLEDGE DISCOVERY IN DATABASE<sup>6</sup>

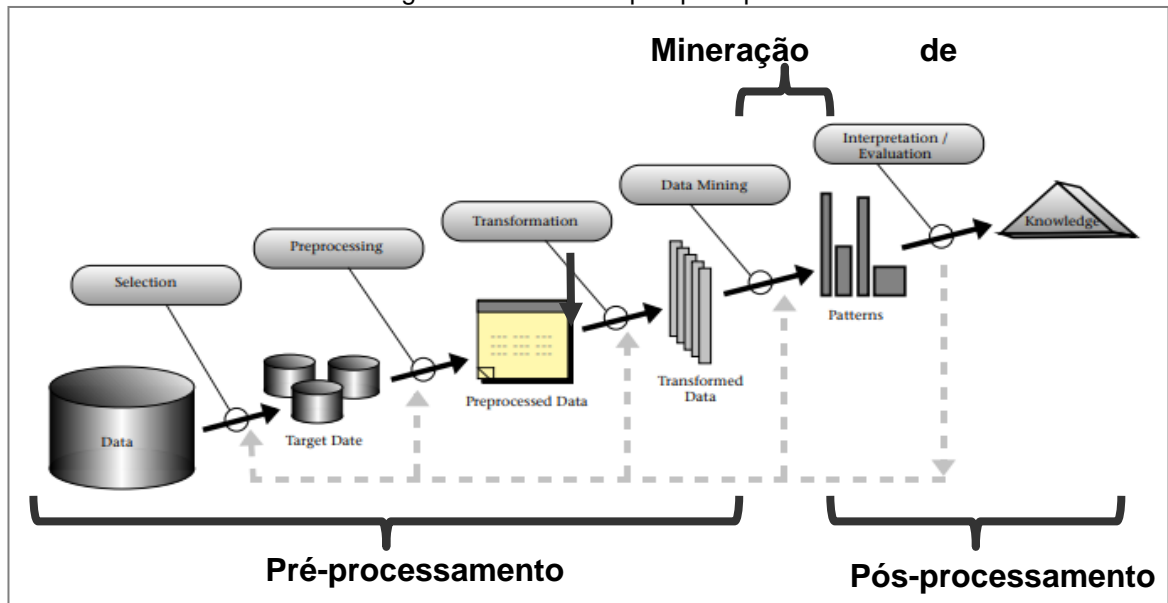
O KDD é um processo que permite a extração de padrões<sup>7</sup> de um conjunto de dados (FAYYAD et al., 1996) sem a necessidade de conhecimento prévio de qualquer dependência entre as variáveis (WITTEN; FRANK, 2005). Esses padrões, quando válidos, novos, potencialmente úteis e compreensíveis, podem complementar o conhecimento existente (FAYYAD et al., 1996). Por essas características, entende-se: ser válido – necessário certo grau de certeza; ser novo – deve acrescentar elementos ao que já se conhece; ser útil – deve agregar valor e ser passível de ser incorporado na prática; e ser compreensível – deve ser em linguagem que o ser humano entenda (REZENDE et al., 2003). Para a descoberta dos padrões, o KDD compreende as etapas de pré-processamento, mineração de dados e pós-processamento (Figura 3).

---

<sup>6</sup> Descoberta de conhecimento em base de dados

<sup>7</sup> Refere-se como padrão um conjunto de variáveis e valores relacionados a determinado desfecho.

Figura 3 – KDD e etapas principais.



Fonte: Fayyad et al. (1996, p. 41).

### 2.3.1 Pré-processamento

Na etapa de pré-processamento, realizam-se a captação, organização e tratamento dos dados, com o objetivo de preparar e qualificar o conjunto de dados para maior probabilidade de descoberta de padrões válidos, novos e úteis para complementar o conhecimento existente. Segundo Goldschmidt, Passos e Bezerra (2015), esta etapa compreende diferentes atividades, apresentadas no Quadro 2.

Quadro 2 – Atividades de pré-processamento e respectiva descrição.

Atividade	Descrição
Seleção de dados	Seleção de subconjunto por meio da identificação dos registros e variáveis que serão trabalhados. Duas heurísticas: eliminar variáveis com valores constantes ou variáveis com valores de identificação.
Limpeza	Tratamento dos dados para assegurar sua qualidade (completude, veracidade e integridade).
Codificação	Transformar os valores em nominais ou numéricos, de forma a atender às necessidades dos algoritmos, visando a obter a melhor representação do conhecimento.
Enriquecimento	Agregar dados oriundos de outros sistemas de informação, com o objetivo de aumentar as possibilidades de descoberta de padrões válidos, novos e potencialmente úteis.
Construção	Criação de novas variáveis a partir daquelas já existentes.
Correção de prevalência	Correção de eventual desequilíbrio na distribuição dos registros – ocorrência maior de um valor que dos outros –, por meio de amostragem estratificada ou replicação aleatória de registros.

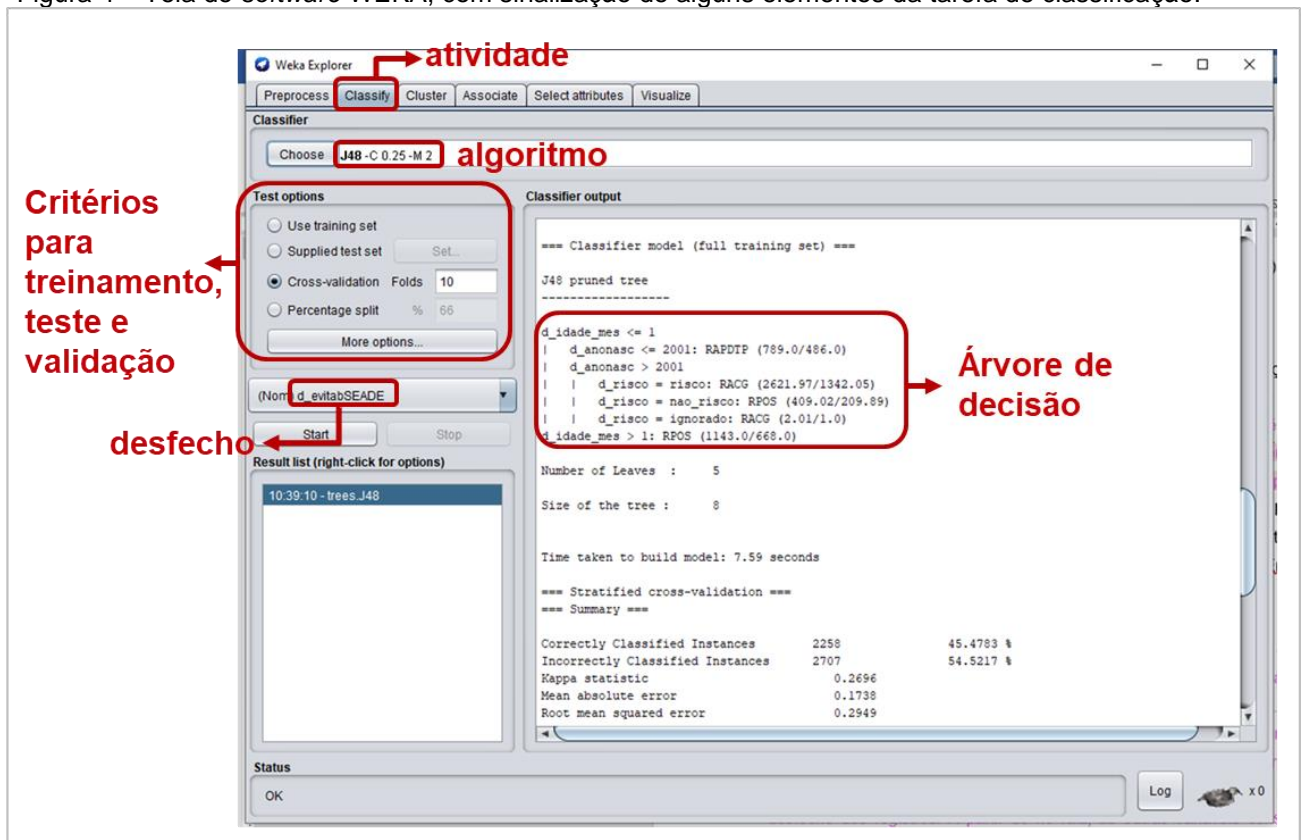
Fonte: Adaptado de Goldschmidt, Passos e Bezerra (2015).

### 2.3.2 Mineração de dados

Na etapa de mineração de dados, os algoritmos descobrem padrões nos dados pré-processados. Esses algoritmos são agrupados em tarefas conforme o problema proposto, como, por exemplo, classificação e descoberta de regras de associação.

**Classificação.** O objetivo desta tarefa é a predição de um desfecho. O algoritmo separa o conjunto de dados em conjunto para treinamento, teste e validação e constrói uma função que permita prever a probabilidade da ocorrência do desfecho (GOLDSCHMIDT; PASSOS; BEZERRA, 2015). Dos algoritmos disponíveis para essa tarefa, destaca-se o C4.5 (QUINLAN, 1993), implementado pelo J48 no *Waikato Environment for Knowledge Analysis* (WEKA) (HALL et al., 2009). A partir da Figura 4, é possível identificar os elementos da tarefa de classificação a partir de uma tela do WEKA.

Figura 4 – Tela do software WEKA, com sinalização de alguns elementos da tarefa de classificação.

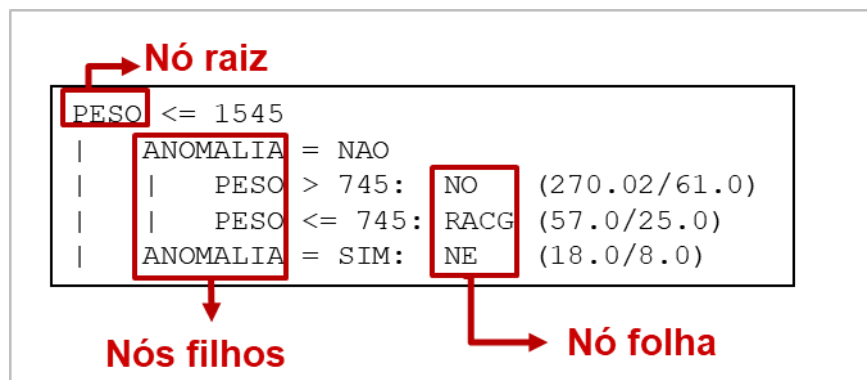


Fonte: A autora (2020).

Uma das formas de representação do resultado da tarefa de classificação é a árvore de decisão, que possibilita a identificação das variáveis e valores

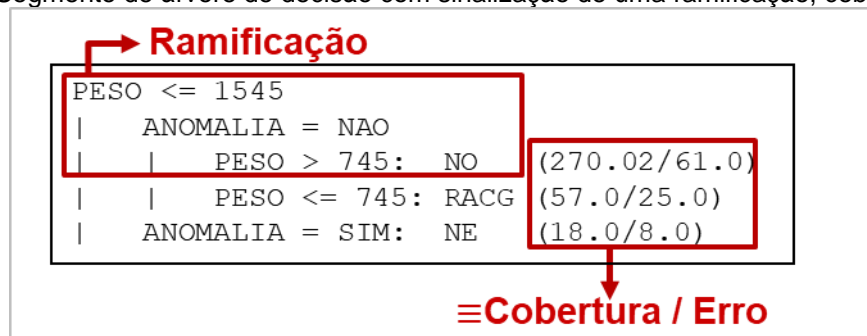
considerados pelo algoritmo para prever o desfecho. Na árvore de decisão, cada variável é denominada nó, sendo a primeira o nó-raiz; as próximas, os nós filhos; e o desfecho, o nó-folha (Figura 5). Cada sequência de nós-raiz, filhos e folha é denominada ramificação e representa um padrão no conjunto de dados ou regra. No fim de cada regra, os valores entre parênteses representam, na sequência, a cobertura e o erro estimados (Figura 6). O primeiro refere-se ao número estimado de registros que obedecem à regra e o segundo, ao número estimado de registros que fogem à regra, ou seja, apresentam mesmos nós-raiz e filhos, porém diferente nó-folha (desfecho) (KOHAVI; QUINLAN, 1999).

Figura 5 – Segmento de árvore de decisão com sinalização do nó-raiz, nós filhos e nó-folha.



Fonte: A autora (2020).

Figura 6 – Segmento de árvore de decisão com sinalização de uma ramificação, cobertura e erro.



Fonte: A autora (2020).

Na Figura 6, a regra (ramificação) sinalizada demonstra que, quando peso ao nascer menor ou igual a 1.545 g e maior que 745 g, sem presença de anomalia, não há ocorrência de óbito (variável desfecho de valor "NO"). Estima-se que isso ocorra em 270.02 registros (cobertura), porém, em 61 registros (erro), apesar dos mesmos nós-raiz e filhos, há a ocorrência do óbito. Os erros podem ser considerados exceções e suas regras tendem a ser surpreendentes, pois representam uma

contradição à regra de senso comum (MILANI; CARVALHO, 2013). As exceções não são possíveis de serem qualificadas pela classificação, sendo necessária a tarefa de descoberta de regras de associação.

**Descoberta de regras de associação.** Enquanto na classificação ocorre a predição de um desfecho, na associação, há descrição das combinações frequentes entre todos os itens, com cada item representando uma variável e seu respectivo valor. As regras são representadas no formato  $A \rightarrow C$ , lidas como “se [A], então [C]”, em que [A] representa um conjunto de itens de dados antecedentes, associados à [C], conjunto de itens de dados consequentes (Figura 7). Destaca-se como algoritmo o Apriori (AGRAWAL; SRIKANT, 1994) e o potencial desta tarefa é identificar, no pós-processamento, as situações de exceção, ou seja, o item de dado que, quando presente no antecedente, altera o consequente.

Figura 7 – Exemplos de regras descobertas na tarefa de associação, contendo desfecho de óbito.

SE Gestação anterior = 01, ENTÃO óbito = NÃO (32.2%, 99.3%)
SE Gestação anterior = 01 E Duração gestação = 22 a 27 semanas, ENTÃO óbito = SIM (0.1%, 62.5%)
SE Gestação anterior = 02, ENTÃO óbito = NÃO (14.3%, 99.0%)
SE Gestação anterior = 02, e Duração gestação = 22 a 27 semanas ENTÃO óbito = SIM (0.1%, 69.2%)
SE Apresentação = pélvica/podálica, ENTÃO óbito = NÃO (3.8%, 96.1%)
SE Apresentação = pélvica/podálica, E Duração gestação = 22 a 27 semanas ENTÃO óbito = SIM (0.1%, 66.7%)

Fonte: A autora (2018).

No fim de cada regra de associação descoberta, existem valores entre parênteses que representam sequencialmente o suporte e a confiança (Figura 7). O suporte refere-se à ocorrência do antecedente e a confiança, à ocorrência do consequente dada a ocorrência do antecedente (AGRAWAL; IMIELINSKI; SWAMI, 1993; BORGELT, 2010) (Figura 8). Como exemplo, lê-se na primeira regra da Figura 7 que, em 32,2% dos registros no conjunto de dados, a quantidade de gestação anterior foi 01, dos quais 99,3% não foram a óbito.

Figura 8 – Formulação matemática de suporte e confiança da tarefa de descoberta de regras de associação (Regra  $A \rightarrow C$ ).

$$\text{SUPORTE} = \frac{\text{número de registros em que [A] ocorre}}{\text{número total de registros}}$$

$$\text{CONFIANÇA} = \frac{\text{número de registros em que [A] e [C] ocorrem}}{\text{número de registros em que [A] ocorre}}$$

Fonte: Adaptado de Agrawal, Imielinski e Swami (1993) e Borgelt (2010).

### 2.3.3 Pós-processamento

No pós-processamento, as regras descobertas são analisadas no intuito de identificar sua contribuição à solução do problema estudado. É nesta etapa, por exemplo, que regras óbvias e já consolidadas podem ser desconsideradas. As estratégias são selecionadas de acordo com a tarefa realizada. Para a classificação, a estratégia de generalização considera que o nó-raiz e primeiros nós filhos representam as variáveis mais fortemente relacionadas ao desfecho, pois são as primeiras a ser consideradas pelo algoritmo para sua predição (MILANI; CARVALHO, 2013; TEIXEIRA; COLMANETTI; CARVALHO, 2015).

Outra estratégia de pós-processamento da tarefa de classificação é a transcrição da árvore de decisão em regras para facilitar a compreensão. Considera-se que o caminho do nó-raiz até o último nó-filho é o antecedente da regra e o nó-folha, o consequente (MILANI; CARVALHO, 2013). Na Figura 6, a transcrição da primeira regra seria:

**SE** peso ao nascer menor ou igual a 1.545 g

**E** anomalia igual a NÃO

**E** peso ao nascer maior que 745 g,

**ENTÃO** óbito igual a NÃO.

Essa transcrição facilita a compreensão da regra e a qualificação de suas exceções, realizada no pós-processamento das regras de associação descobertas.

Para a identificação de regras de exceção, pressupõe-se que, a partir de uma regra geral, as regras de exceção contêm um item de dado (variável e respectivo valor) que altera o desfecho da regra geral. Submete-se o conjunto de regras de associação descobertas a um algoritmo “Descobridor de Regras de Exceção” (DRE), que lista as regras gerais e suas respectivas exceções (MILANI; CARVALHO, 2013).

A partir das regras gerais e respectivas exceções, utiliza-se a estratégia do filtro de regras de associação para avaliar e selecionar ou descartar as regras conforme critérios estabelecidos para cada caso.

Para demonstrar o potencial da descoberta de uma regra de exceção, na Figura 9 estão listadas as regras da Figura 8 após submissão ao DRE. É possível identificar que “duração da gestação igual a 22 a 27 semanas” (em destaque) altera o desfecho do óbito de NÃO para SIM em mais de uma regra geral, sugerindo que esse item de dado é um fator de risco para a ocorrência do óbito.

Figura 9 – Exemplos de regras gerais e de exceção após submissão do conjunto de regras a um algoritmo DRE.

<p>Regra geral 1: SE Gestação anterior = 01, ENTÃO óbito = NÃO (32.2%, 99.3%) Regra de exceção 1.1: SE Gestação anterior = 01 E <b>Duração gestação = 22 a 27 semanas</b>, ENTÃO óbito = SIM (0,1%, 62.5%)</p> <p>Regra geral 2: SE Gestação anterior = 02, ENTÃO óbito = NÃO (14.3%, 99.0%) Regra de exceção 2.1: SE Gestação anterior = 02, e <b>Duração gestação = 22 a 27 semanas</b> ENTÃO óbito = SIM (0.1%, 69.2%)</p> <p>Regra geral 3: SE Apresentação = pélvica/podálica, ENTÃO óbito = NÃO (3.8%, 96.1%) Regra de exceção 3.1:</p>
---

Fonte: A autora (2018).

### 3 MAPEAMENTO ENTRE O REFERENCIAL TEÓRICO E OS ENCAMINHAMENTOS METODOLÓGICOS

No Quadro 3, são apresentadas as referências que apoiaram a definição dos encaminhamentos metodológicos.

Quadro 3 – Mapeamento entre referencial teórico e encaminhamentos metodológicos.

Referência	Conceito	Contribuição
Black et al. (2008)	Primeiros 1.000 dias de vida	Recorte etário adotado para a população a ser pesquisada
Dahlgren e Whitehead (1991)	DSS	Modelo utilizado para inferir utilidade dos padrões descobertos de acordo com a governabilidade do setor de saúde
Fayyad et al. (1996)	KDD permite a extração de padrões de um conjunto de dados e, quando válidos, novos, potencialmente úteis e compreensíveis, podem complementar o conhecimento existente	Processo adotado para a descoberta de padrões para complementação do conhecimento existente
Witten e Frank (2005)	Não há necessidade de conhecimento prévio de alguma dependência entre as variáveis para a extração de conhecimento pelo KDD	Uso do KDD em coorte histórica para encontrar padrões relacionados ao óbito e novas variáveis
Quinlan (1993) Hall et al. (2009)	Algoritmo C4.5 implementado pelo J48 no WEKA	Algoritmos adotados para a etapa de mineração de dados
Agrawal e Srikant (1994)	Apriori	
Kohavi e Quinlan (1999)	Algoritmo C4.5 – cobertura e erro estimado	Erro estimado considerado exceção
Agrawal, Imielinski e Swami (1993) Borgelt (2010)	Suporte e confiança	Adoção de valores para identificação de regras de exceção representativas
Teixeira, Colmanetti e Carvalho (2015)	Estratégia de pós-processamento de árvores de decisão baseada na generalização	Estratégia adotada na etapa de pós-processamento da árvore de decisão (tarefa de classificação), considerando que os primeiros nós são os mais fortemente relacionados ao desfecho
Milani e Carvalho (2013)	Estratégia de pós-processamento de descoberta de regras de exceção	Estratégia adotada na etapa de pós-processamento das regras de associação para qualificação das regras de exceção
Fenton e Kim (2013)	Tamanho do feto/RN na curva de crescimento Fenton, levando em consideração peso ao nascer e semana gestacional	Referências utilizadas para construção de variáveis de tamanho do feto/RN na curva de crescimento em PIG, AIG e GIG.
Papageorghiou et al. (2014) Villar et al. (2014)	Tamanho do feto/RN na curva de crescimento Intergrowth, levando em consideração peso ao nascer e semana gestacional	Foram construídas duas variáveis, uma considerando curva de crescimento Fenton e outra considerando curva Intergrowth.



Jeganathan et al. (2017)	Existência de poucos estudos sobre a recuperação do Apgar entre o primeiro e o quinto minuto. Além das condições do RN, esse intervalo pode indicar informações sobre a assistência prestada	Construção de variável de recuperação de Apgar, considerando os valores no primeiro e quinto minuto e a diferença entre eles
Ortiz (2000 <i>apud</i> BRASIL, 2009b)	Classificação de evitabilidade do óbito da Fundação Seade	Identificação da causa do óbito e inferência do nível de governabilidade do gestor da saúde sobre ela
Dias, Santos Neto e Andrade (2017)	Classificação de evitabilidade do óbito da Fundação Seade é capaz de distinguir os óbitos evitáveis de maneira mais precisa	Escolha dessa classificação para evitabilidade
United Nations Children's Fund; World Health Organization (2004)	Conceito de baixo peso ao nascer (baixo peso menor que 2.500 g; muito baixo peso menor que 1.500 g; e extremo baixo peso menor que 1.000 g).	Avaliação do valor de corte de peso ao nascer

Fonte: A autora (2020).

## 4 ENCAMINHAMENTOS METODOLÓGICOS

### 4.1 NATUREZA, POPULAÇÃO E CENÁRIO DA PESQUISA

Trata-se de uma coorte histórica devido à reconstrução do acompanhamento de grande número de indivíduos por meio de dados coletados antes do início da pesquisa (AQUINO; BARRETO; SZKLO, 2014). Na coorte, indivíduos expostos e não expostos são acompanhados até certo momento com identificação da ocorrência ou não de um desfecho, oportunizando o conhecimento sobre os riscos em saúde, identificação de doenças e pistas para a causalidade (ALMEIDA FILHO; BARRETO, 2014).

Este estudo faz parte de uma pesquisa maior – Coorte Saúde Materno-Infantil de Curitiba (COOSMIC) –, porém seu escopo limita-se a trabalhar com o conjunto de dados oriundo do SIM e SINASC. O COOSMIC é realizado pela Pontifícia Universidade Católica do Paraná (PUCPR), em parceria com a SMS de Curitiba, e possui aprovação do Comitê de Ética em Pesquisa (CEP) de ambas as instituições, sob Parecer nº 2.672.835 (Anexo C) e Parecer nº 2.728.771 (Anexo D), respectivamente.

A população compreende os nascidos vivos de 1999 a 2017, de mães residentes em Curitiba. Para sua reconstrução, foram utilizadas bases de dados do SIM e SINASC. O início do recorte temporal foi escolhido pela implantação da Rede Mãe Curitibana Vale a Vida, em 1999.

Curitiba é a capital do Paraná, com população estimada para 2019 de 1.933.105 pessoas (INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA, 2019). Sua área de 435,036 km<sup>2</sup> engloba 75 bairros e Índice de Desenvolvimento Humano Municipal (IDHM) de 0,823 (CURITIBA, 2018b) – faixa de IDHM considerada muito alta. A RAS em Curitiba conta com 110 Unidades Básicas de Saúde, nove Unidades de Pronto Atendimento, seis Centros de Especialidades Médicas, três Centros de Especialidades Odontológicas, entre outros equipamentos. Essa rede está organizada em dez Distritos Sanitários: Bairro Novo (BN), Boa Vista (BV), Boqueirão (BQ), Cajuru (CJ), Cidade Industrial de Curitiba (CIC), Matriz (MZ), Pinheirinho (PN), Portão (PR), Santa Felicidade (SF) e Tatuquara (TQ).

## 4.2 CARACTERÍSTICA DOS CONJUNTOS DE DADOS

A partir da Tabela 1, é possível identificar a quantidade de registros por ano e SIS antes do *linkage*. À direita da coluna de registros do SIM, está a respectiva estratificação pelo grupo de evitabilidade (reduzibilidade) do óbito conforme classificação da Fundação Seade: mal definido (MD), não classificado/outro (NCO), não evitável (NE), evitável por adequada atenção ao parto (RAAP), evitável por adequado controle na gravidez (RACG), evitável por ações de prevenção, diagnóstico e tratamento precoces (RAPDTP), evitável por imunoprevenção (RI) e evitável por parcerias com outros setores (RPOS).

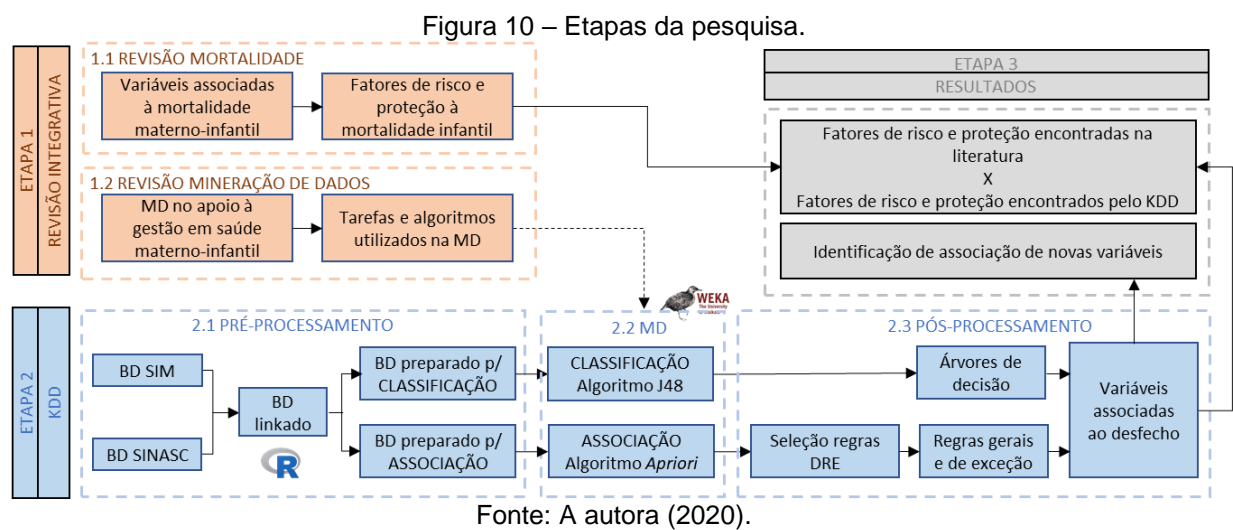
Tabela 1 – Número de registros por ano e SIS antes do *linkage*, com estratificação dos registros do SIM por grupo de evitabilidade do óbito.

Ano de registro	SINASC	SIM	Grupo de evitabilidade do óbito							
			Mal Definido	Não classificado / Outro	Não Evitável	Redutível por Adequada Atenção ao Parto	Redutível por Adequado Controle na Gravidez	Redutível por Ações de Prevenção, Diagnóstico e Tratamento Precoces	Redutível por Imunoprevenção	Redutível por Parcerias com Outros Setores
1999	29.520	471	22		106	30	28	170	2	113
2000	29.312	466	9		102	27	33	188	3	104
2001	27.090	405	6		90	35	19	165	1	89
2002	26.325	329	2		56	37	88	83	3	60
2003	24.803	330	2		62	37	97	54	1	77
2004	25.123	304			46	37	84	60	1	76
2005	24.442	308	5		46	31	92	55		79
2006	24.331	277	1		48	30	79	48		71
2007	23.995	270	2		35	27	92	43	1	70
2008	24.996	261	2		41	25	99	41		53
2009	24.555	240		1	33	23	86	29		68
2010	25.028	245			33	20	100	34		58
2011	24.872	239	3		40	23	81	34	2	56
2012	24.701	257	1		50	25	90	35	3	53
2013	24.934	233		1	36	25	80	28	2	61
2014	24.815	198			27	27	71	20		53
2015	24.506	230			34	32	88	27		49
2016	23.217	221		2	28	17	79	34	1	60
2017	22.744	198	2		39	28	78	12		39
<b>Total</b>	<b>479.309</b>	<b>5.482</b>	<b>57</b>	<b>4</b>	<b>952</b>	<b>536</b>	<b>1.464</b>	<b>1.160</b>	<b>20</b>	<b>1.289</b>

Fonte: A autora (2020).

### 4.3 ETAPAS DA PESQUISA

A pesquisa foi desenvolvida em três etapas (Figura 10), nas quais se contou com o apoio dos estudantes do Projeto Institucional de Bolsas de Iniciação Científica (Pibic) 2018-2019 do curso de Medicina, do 2º grau (Pibic Júnior) e do curso de Ciência da Computação para a realização, respectivamente, da revisão integrativa de mortalidade, dos testes de *linkage* e dos testes de mineração de dados. Os testes realizados auxiliaram na modelagem dos encaminhamentos.



#### 4.3.1 Etapa 1 – revisão integrativa

A primeira etapa compreendeu duas revisões integrativas, cada uma correspondente a um objetivo específico. Para cumprir o objetivo específico de “identificar variáveis associados ao óbito materno-infantil por meio de revisão integrativa”, realizou-se em conjunto com estudante Pibic de Medicina (2018-2019) o artigo sobre variáveis associadas à mortalidade materno-infantil que se encontra no Apêndice A. Para o cumprimento do objetivo específico de “identificar o potencial da mineração de dados para apoio à gestão em saúde materno-infantil por meio de revisão integrativa”, realizou-se o artigo Mineração de dados no apoio à gestão em saúde materno-infantil: uma revisão integrativa, que se encontra do Apêndice B.

#### 4.3.2 Etapa 2 – KDD

**Pré-processamento.** No Excel®, realizaram-se a análise e consistência das bases de dados conforme as atividades de pré-processamento descritas no Quadro

4. Vale ressaltar que, nesta etapa, foram criadas novas variáveis a partir de outras (Apêndice C), dentre elas, a correspondente ao desfecho estudado “evitabilidade do óbito” a partir da causa básica de óbito. Optou-se por trabalhar com grupo de evitabilidade do óbito, no intuito de apontar a governabilidade da gestão, e com a classificação Seade, por ser uma das classificações recomendadas pelo MS para análise de evitabilidade do óbito infantil e fetal (BRASIL, 2009b) com a capacidade de classificar a maior quantidade de óbitos em evitáveis ou não evitáveis (DIAS; SANTOS NETO; ANDRADE, 2017).

Durante o pré-processamento, foram realizadas reuniões com os setores responsáveis da SMS de Curitiba (CE, Coordenação dos Eventos Vitais e Departamento de Atenção Primária, Coordenação da Saúde da Criança) para melhor compreensão do fluxo de informação, organização da rede e dificuldades enfrentadas, de forma que o pré-processamento resultasse em conjunto de dados cuja mineração de dados tivesse maior probabilidade de encontrar conhecimentos novos, válidos e úteis para a gestão. As variáveis mantidas constam no Apêndice C e as excluídas, no Apêndice D, com respectiva justificativa de exclusão.

Quadro 4 – Atividades de pré-processamento realizadas e exemplos.

<b>Atividade</b>	<b>Exemplos</b>
Seleção de dados	Seleção de todos os registros de nascidos vivos entre 1999 e 2017 e respectivos óbitos ocorridos de zero a dois anos de idade.
Limpeza	Exclusão de registros que continham somente número de DN ou número de DO, sem outras variáveis preenchidas. Limpeza de valores código 8, ausente no dicionário de dados do SINASC. Padronização de grafia. Inclusão de dados a partir de outras variáveis, como, por exemplo, inserção de data de nascimento na DO, considerando a informação da idade do óbito em horas.
Codificação	Transformação dos códigos em valores categóricos, conforme dicionário de dados. Exemplo: 1 = sim, 2 = não.
Enriquecimento	A partir do número do CNES do estabelecimento de saúde, consultado no <i>site</i> se o estabelecimento de saúde era SUS ou não.
Construção	Criação de variáveis a partir de outras já existentes: - Idade em dias no momento do óbito, a partir da data de nascimento e data de óbito. - Grupo de evitabilidade do óbito conforme Seade, a partir da causa básica. - Classificação de risco ao nascer a partir da codificação registrada com o código do bairro.
Correção de prevalência	Não necessário, pois o algoritmo utilizado lida com essa prevalência, atribuindo peso maior ao erro associado às classes menores.

Fonte: A autora (2020).

A modelagem para *linkage* dos conjuntos de dados contou com o apoio de colegas do grupo de pesquisa e do estudante Pibic Júnior (2018-2019) e foi

realizada no *software* R<sup>®</sup>. Teve por objetivo criar pares de óbito-nascido vivo, encontrando para cada registro de óbito no SIM o respectivo registro de nascido vivo no SINASC. O R<sup>®</sup> foi escolhido por ser *software* livre que facilita o manuseio dos dados por meio de funcionalidades disponíveis em bibliotecas abertas<sup>8</sup>, das quais foram utilizadas para *linkage*: *openxlsx*, *dplyr*, *lubridate*, *fuzzyjoin*, *stringi*, *stringr* e *readr*. O processo detalhado do *linkage* consta no Apêndice E, com respectiva descrição e *script* realizado no R<sup>®</sup>.

O conjunto de dados resultante do processo de *linkage* foi denominado SIM-SINASC final e serviu de base para os experimentos de mineração de dados. Suas variáveis foram aproximadas das camadas do modelo dos DSS, a fim de direcionar os experimentos KDD e a leitura dos resultados sob a perspectiva de utilidade para a gestão. Reconhecendo a interrelação existente entre camadas, essa aproximação levou em consideração a camada mais contígua à variável. As variáveis desse conjunto de dados final e respectiva descrição, SIS de origem e camada dos DSS aproximada constam no Apêndice C.

**Mineração de dados.** A modelagem desta etapa contou com o apoio de estudante Pibic de Ciência da Computação (2018-2019). Utilizou-se o WEKA por ser *software* livre amplamente utilizado no KDD. Porém, dada a sua limitação, realizaram-se diferentes experimentos em movimento evolutivo, uma vez que os resultados do experimento anterior modelaram o seguinte na busca por padrões que complementassem o conhecimento existente. No total, foram realizados quatro experimentos com diferentes disposições de variáveis, registros e anos no subconjunto de dados, sendo os três primeiros de classificação e o último de descoberta de regras de associação.

Nas tarefas de classificação, utilizou-se o algoritmo C4.5 (QUINLAN, 1993), implementado pelo J48 no *software* WEKA (HALL et al., 2009). Na tarefa de descoberta de regras de associação, aplicou-se o algoritmo Apriori (AGRAWAL; SRIKANT, 1994) com suporte de 0,1 e confiança de 30, no intuito de “captar” as exceções. Os experimentos foram nomeados conforme o conjunto de dados e tarefa utilizada e estão dispostos no Quadro 5, em que o destaque em cinza identifica a alteração no subconjunto de dados em relação ao experimento anterior.

---

<sup>8</sup> Bibliotecas disponíveis em <https://www.r-project.org/>.

Quadro 5 – Experimentos de mineração de dados e características dos subconjuntos utilizadas (variáveis SINASC, variáveis SIM, registros e disposição dos anos).

	<b>Experimento 1 “Classificação SINASC todos”</b>	<b>Experimento 2 “Classificação SIM-SINASC óbitos”</b>	<b>Experimento 3 “Classificação SIM-SINASC óbitos agrupados”</b>	<b>Experimento 4 “Associação* SINASC óbitos agrupados”</b>
Variáveis SINASC	Todas conforme Apêndice C.	Todas conforme Apêndice C, com exclusão daquelas encontradas no experimento 1 mais fortemente relacionadas ao desfecho e consolidadas na literatura.	Todas conforme Apêndice C, com exclusão daquelas encontradas no experimento 1 mais fortemente relacionadas ao desfecho e consolidadas na literatura.	Somente variáveis relacionadas às camadas dos DSS de condições socioambientais gerais e de vida e trabalho.
Variáveis SIM	Evitabilidade do óbito.	Todas conforme Apêndice C.	Todas conforme Apêndice C.	Evitabilidade do óbito.
Registros	Todos os nascidos vivos, independentemente de óbito.	Somente nascidos vivos que foram a óbito.	Somente nascidos vivos que foram a óbito.	Somente nascidos vivos que foram a óbito.
Disposição dos anos por subconjunto	Um ano por subconjunto de dados (total 19 subconjuntos).	Um ano por subconjunto de dados (total 19 subconjuntos).	Todos os anos em um único subconjunto de dados.	Todos os anos em um único subconjunto de dados.

Fonte: A autora (2020).

Nota: Os destaques em cinza identificam a mudança em comparação ao experimento anterior.

\* Para reduzir o nome, o experimento foi denominado “associação”, porém entende-se que se trata de descoberta de regras de associação.

No experimento 1, o subconjunto de dados utilizado continha as variáveis do SINASC conforme Apêndice C, uma única variável do SIM (evitabilidade do óbito) e todos os registros de nascidos vivos, independentemente de óbito. Foram trabalhados com 19 subconjuntos de dados, um para cada ano, obtendo-se como resultado uma árvore de decisão por ano, na qual os valores do desfecho são “não óbito” e os oito grupos de evitabilidade do óbito explicados no referencial teórico. Nas árvores de decisão, identificou-se que as variáveis mais fortemente relacionadas ao desfecho não agregavam ao conhecimento existente por já serem conhecidas e consolidadas na literatura. Assim, para os experimentos 2 e 3, essas variáveis foram excluídas, as variáveis do SIM foram acrescidas e consideraram-se os registros somente de nascidos vivos que foram a óbito, portanto, o desfecho passou a ser exclusivamente o grupo de evitabilidade do óbito, com oito valores.

No experimento 2, continuou-se a trabalhar com cada ano em um subconjunto e as árvores de decisão descobertas não permitiram a análise do todo. Realizou-se, então, o experimento 3, no qual os 19 subconjuntos foram agrupados num único

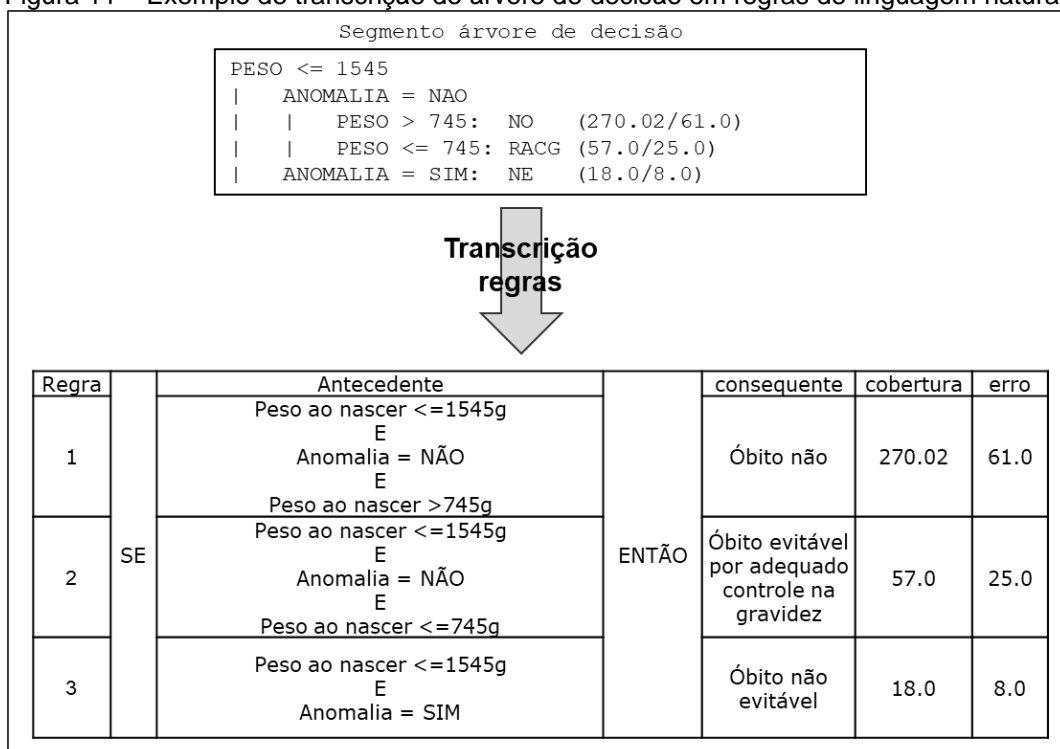
subconjunto de dados. Nesse experimento, identificou-se que as exceções poderiam indicar situações passíveis de atuação direta do gestor da saúde e, para qualificá-las, realizou-se o quarto e último experimento, de descoberta de regras de associação.

Para a qualificação das exceções, no experimento 4, foi criado um subconjunto de dados com variáveis passíveis de atuação direta do gestor (Apêndice C), conforme os seguintes critérios:

- a) origem SINASC por ser antecedente ao óbito;
- b) aproximadas as camadas mais externas do modelo dos DSS: (i) condições socioeconômicas e ambientais gerais; (ii) condições de vida e trabalho, especificamente relacionadas aos serviços de saúde.

**Pós-processamento.** Na tarefa de classificação, os resultados do algoritmo J48 representados por árvores de decisão foram transcritos em linguagem natural para facilitar a compreensão, conforme exemplo da Figura 11. Seu pós-processamento baseou-se na generalização (TEIXEIRA; COLMANETTI; CARVALHO, 2015), identificando-se as variáveis e respectiva posição nas árvores de decisão, sendo a variável do nó-raiz e as mais próximas a ela consideradas mais fortemente relacionadas ao desfecho.

Figura 11 – Exemplo de transcrição de árvore de decisão em regras de linguagem natural.




Fonte: A autora (2020).



Para o pós-processamento das regras de associação, empregaram-se as estratégias de descoberta de regras de exceção e de filtro de regras de associação. Após identificação das regras gerais e de exceção pelo algoritmo, selecionaram-se as regras gerais correspondentes às regras descobertas no experimento 3 para as quais se desejava qualificar as exceções; dessas, consideraram-se interessantes as regras de exceção cujo item de dado que alterava o desfecho se aproximava da camada de condições socioeconômicas, culturais e ambientais gerais, por ser de governabilidade direta do gestor, conforme modelo dos DSS. Para facilitar a compreensão, as regras gerais e de exceção foram transcritas conforme Figura 12.

Figura 12 – Exemplo de transcrição de regras gerais e de exceção em regras de linguagem natural.

<p>Regra geral 1: SE QTDGESTANT_01, ENTÃO obito_NAO (32.2%, 99.3%) Exceção 1.1: SE QTDGESTANT_01 E GESTACAO_2, ENTÃO obito_SIM (0.1%, 62.5%)</p> <p>Regra geral 2: SE QTDGESTANT_02, ENTÃO obito_NAO (14.3%, 99.0%) Exceção 2.1: SE QTDGESTANT_02 E GESTACAO_2, ENTÃO obito_SIM (0.1%, 69.2%)</p> <p>Regra geral 3: SE pelvica_podalica, ENTÃO obito_NAO (3.8%, 96.1%) Exceção 3.1: SE pelvica_podalica E GESTACAO_2, ENTÃO obito_SIM (0.1%, 66.7%)</p>						
<p><b>Transcrição</b> <b>regras</b></p> 						
Regra		antecedente		consequente	suporte	confiança
Geral 1		quantidade de gestação anterior = 01		Óbito NÃO	32,2%	99,3%
Exceção 1.1		quantidade de gestação anterior = 01 E duração da gestação = 22 a 27 semanas		Óbito SIM	0,1%	62,5%
Geral 2		quantidade de gestação anterior = 02		Óbito NÃO	14,3%	99,0%
Exceção 2.1	<b>SE</b>	quantidade de gestação anterior = 02 E duração da gestação = 22 a 27 semanas	<b>ENTÃO</b>	Óbito SIM	0,1%	69,2%
Geral 3		apresentação do recém nascido = pélvica/podálica		Óbito NÃO	3,8%	96,1%
Exceção 2.3		apresentação do recém nascido = pélvica/podálica E duração da gestação = 22 a 27 semanas		Óbito SIM	0,1%	66,7%

Fonte: A autora (2020).

### 4.3.3 Etapa 3 – análise dos resultados

Realizaram-se testes estatísticos para avaliar a relação entre o desfecho “evitabilidade do óbito” e as variáveis encontradas nas árvores de decisão e regras de exceção. Nas árvores de decisão, considerou-se a estratégia de generalização e os testes foram feitos para as variáveis mais frequentes como nó-raiz até o terceiro nó-filho. Nas regras de exceção, os testes estatísticos foram realizados para as variáveis que alteravam o desfecho de RPOS para óbito evitável por qualquer ação da saúde (RAAP, RACG, RAPDTP e RI).

As hipóteses testadas consideraram como variáveis independentes aquelas encontradas pelo KDD e como variável dependente, a evitabilidade do óbito. Para as variáveis qualitativas, foi utilizado o teste de qui-quadrado e, para as quantitativas, devido à distribuição não normal dos grupos, o teste de Kruskal-Wallis com comparação dois a dois com nível de significância corrigido por Bonferroni.

Os testes foram realizados considerando o conjunto de dados do experimento no qual se encontrou a variável, sendo desconsiderados os grupos de evitabilidade do óbito com frequência relativa menor que 0,01% no conjunto de dados contendo todos os nascidos vivos e menor que 1% no conjunto de dados contendo somente os óbitos. Foram desconsiderados também os valores em branco. Utilizou-se o programa SPSS 25 e adotou-se o nível de significância de  $p \leq 0,05$ , com intervalo de confiança de 95%.

Para cumprimento do objetivo específico de “identificar novas variáveis de governabilidade direta da gestão, relacionadas à evitabilidade do óbito entre zero e dois anos de idade, por meio da mineração de dados”, a governabilidade direta da gestão foi inferida pelas camadas do modelo dos DSS às quais a variável foi aproximada. As características desejáveis de validade, novidade e utilidade das regras descobertas pelo KDD foram avaliadas a partir da literatura, incluindo os resultados encontrados na revisão integrativa de identificação de variáveis associadas ao óbito materno-infantil. Para análise também se levou em consideração os critérios de classificação de risco ao nascer sugeridos pelo MS e os critérios adotados por Curitiba.

## 5 RESULTADOS

Os resultados são apresentados conforme as etapas metodológicas: revisão integrativa, pré-processamento e, por último, mineração de dados e pós-processamento.

### 5.1 REVISÃO INTEGRATIVA

A revisão integrativa, objetivando identificar as variáveis associadas à mortalidade materno-infantil, encontra-se no Apêndice A. Esta revisão foi apresentada no formato de pôster nos eventos 15º Congresso Brasileiro de Clínica Médica, realizado de 3 a 5 de outubro de 2019, em Florianópolis (SC), e no XXVII Seminário de Iniciação Científica, realizado de 22 a 24 de outubro de 2019, em Curitiba (PR). Também foi submetida, no formato artigo, para a Revista de Atenção à Saúde.

Nesta revisão identificou-se que as variáveis mais associadas ao óbito foram peso ao nascer e idade materna, ambas variáveis relacionadas à camada individual do modelo dos DSS, sem governabilidade direta do gestor. Destacou-se que para mudanças nessa camada, as intervenções devem ser realizadas a partir das camadas mais externas, das quais evidenciaram-se as variáveis assistência pré-natal, escolaridade materna, renda e saneamento básico. Por fim, reforçou-se a responsabilidade direta dos gestores na redução da mortalidade materna e infantil, por meio de ações que visem reduzir as iniquidades em saúde. As variáveis encontradas nessa revisão ajudaram na análise e discussão dos resultados.

A revisão integrativa de mineração de dados no apoio à gestão em saúde materno-infantil encontra-se no Apêndice B e foi aceita para publicação na Revista Brasileira de Pesquisa em Saúde (*Brazilian Journal of Health Research*). Nesta revisão identificou-se o uso da mineração de dados para diferentes situações ocorridas nos primeiros 1.000 dias de vida – doenças agudas, doenças crônicas não transmissíveis, assistência à saúde, desnutrição, mortalidade, entre outros. Inferiu-se que o processo KDD pode descobrir regras que complementem o que já se conhece e apoiem a gestão, principalmente quando a interdisciplinaridade se faz presente. Porém, identificou-se que todos os experimentos foram pontuais, de desenvolvimento tecnológico ou de execução do KDD, não tendo sido encontrados

relatos de aplicação dos resultados encontrados pelo KDD e avaliação do emprego dessa tecnologia para apoio à gestão em saúde na saúde materno-infantil.

## 5.2 PRÉ-PROCESSAMENTO

O primeiro resultado do pré-processamento foi uma planilha com inconsistências identificadas nos conjuntos de dados SIM e SINASC. São exemplos de inconsistências encontradas: DO sem informação de data de nascimento, mas com idade no momento do óbito; DNV com grafia incorreta da hora de nascimento (29:60); variáveis com valor cujo código inexistente no dicionário de dados disponibilizado; falta de padrão na grafia de variáveis qualitativas; número de DNV duplicada no SINASC; número de DO duplicada no SIM; e mesmo número de DNV em diferentes DOs. Essa planilha será entregue ao CE/Coordenação de Eventos Vitais da SMS para qualificação das bases de dados<sup>9</sup>.

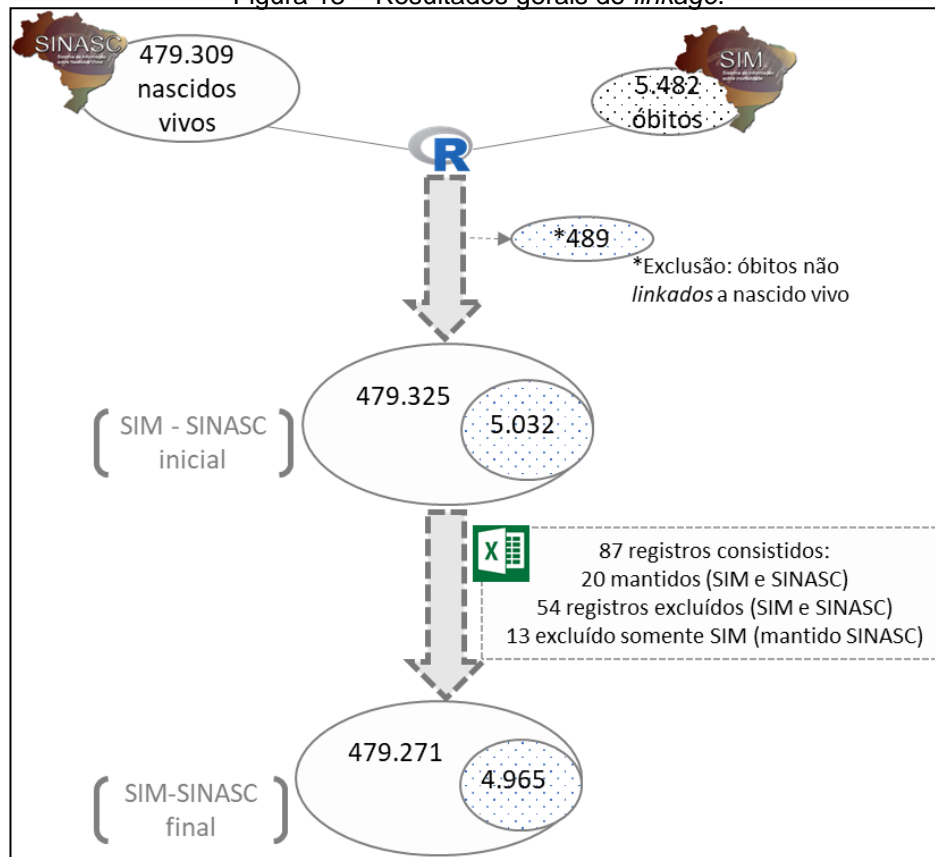
A partir dos conjuntos de dados iniciais do SINASC (479.309 registros de nascidos vivos) e SIM (5.482 registros de óbitos), parearam-se 5.032 óbitos a nascidos vivos (91,8%). Os óbitos não pareados foram descartados (489) e os nascidos vivos não pareados (474.293) foram adicionados aos registros pareados, de maneira que todos os nascidos vivos, independentemente da ocorrência de óbito, estivessem no conjunto de dados. Esse conjunto continha 479.325 registros e foi denominado “SIM-SINASC inicial”. O número maior de registros após *linkage* deveu-se ao pareamento do mesmo óbito a mais de um nascido vivo, como no caso de óbito de gêmeos do mesmo sexo.

Oitenta e sete registros pareados foram verificados quanto à duplicidade e inconsistência, dos quais 20 foram mantidos por estarem corretos, 54, excluídos (46 por não ser possível identificar se o pareamento estava correto e oito por estarem com data de óbito anterior à data de nascimento) e, em 13 registros, foram excluídos somente os dados relacionados ao óbito, por terem sido pareados incorretamente. O conjunto de dados final continha 479.271 nascidos vivos, dos quais 4.965 foram a óbito. A partir da Figura 13, é possível identificar os resultados gerais do *linkage*, cujos detalhes podem ser verificados no Apêndice E.

---

<sup>9</sup> Os resultados dessa pesquisa serão apresentados aos gestores e técnicos dos setores da SMS Curitiba envolvidos: Centro de Epidemiologia - Coordenação de Eventos Vitais e Departamento de Atenção Primária – Rede Mãe Curitibana Vale a Vida e Programa da Criança.

Figura 13 – Resultados gerais do linkage.



Fonte: A autora (2020).

O conjunto de dados SIM-SINASC final continha, portanto, os registros de nascidos vivos entre 1999 e 2017, com informação de ocorrência ou não do óbito, e, nos casos de sua ocorrência, os dados do registro de óbito. Conforme explicitado na seção de materiais e métodos, salienta-se que, devido à limitação do *software* utilizado, foi necessário segmentar esse conjunto de dados em subconjuntos para a realização dos experimentos de KDD. Entretanto, na Tabela 2, é possível identificar o perfil da população estudada constante no conjunto de dados SIM-SINASC final (pós-linkage), com respectiva frequência absoluta e relativa.

Tabela 2 – Perfil da população estudada do conjunto de dados SIM-SINASC final (pós-linkage), com variáveis e respectiva frequência absoluta e relativa.

Variável/classificação	FA (n)	FR (%)
<b>Escolaridade materna – em anos de estudo</b>		
Nenhum	1.486	0,31%
De 1 a 3 anos	11.358	2,37%
De 4 a 7 anos	87.597	18,28%
De 8 a 11 anos	196.018	40,90%
Maior ou igual a 12 anos	158.341	33,04%
Ignorado	1.821	0,38%
(vazio)	22.650	4,73%
<b>Estado civil da mãe</b>		

Solteira	186.994	39,02%
Casada	220.881	46,09%
União estável	40.315	8,41%
Separada	6.505	1,36%
Viúva	1.039	0,22%
Ignorado	779	0,16%
(vazio)	22.758	4,75%
<b>Número de consultas de pré-natal</b>		
Nenhuma	2.729	0,57%
De 1 a 3	10.307	2,15%
De 4 a 6	54.908	11,46%
Maior ou igual a 7	401.784	83,83%
Ignorado	3.039	0,63%
(vazio)	6.504	1,36%
<b>Nascimento SUS/não SUS*</b>		
SUS	155.645	32,48%
Não SUS	131.994	27,54%
Nascimento fora de Curitiba	4.914	1,03%
Ignorado	184.940	38,59%
(vazio)	1.778	0,37%
<b>Tipo de gravidez</b>		
Única	468.141	97,68%
Dupla	10.520	2,20%
Tripla ou mais	421	0,09%
Ignorado	29	0,01%
(vazio)	160	0,03%
<b>Tipo de parto</b>		
Vaginal	210.779	43,98%
Cesáreo	268.287	55,98%
Ignorado	51	0,01%
(vazio)	154	0,03%
<b>Semanas de gestação</b>		
Menor que 22 semanas	103	0,02%
De 22 a 27 semanas	2.196	0,46%
De 28 a 31 semanas	3.599	0,75%
De 32 a 36 semanas	28.134	5,87%
De 37 a 41 semanas	439.775	91,76%
Maior ou igual a 42 semanas	1.959	0,41%
Ignorado	82	0,02%
(vazio)	3.423	0,71%
<b>Raça/cor do nascido vivo</b>		
Branca	327.947	68,43%
Amarela	305	0,06%
Indígena	43	0,01%
Parda	4.738	0,99%
Preta	780	0,16%
(vazio)	145.458	30,35%
<b>Peso ao nascer</b>		
Menor que 1.000 g	2.710	0,57%
Menor que 1.500 g	3.786	0,79%
Menor que 2.500 g	36.115	7,54%
Maior ou igual a 2.500 g	436.457	91,07%
(vazio)	203	0,04%
<b>Sexo</b>		
Masculino	245.153	51,15%
Feminino	234.044	48,83%
Ignorado	74	0,02%
<b>Apgar 1º minuto</b>		
0	520	0,11%
1	2.757	0,58%

2	3.226	0,67%
3	3.786	0,79%
4	5.439	1,13%
5	7.911	1,65%
6	11.068	2,31%
7	24.383	5,09%
8	112.136	23,40%
9	299.788	62,55%
10	6.529	1,36%
99	74	0,02%
(vazio)	1.654	0,35%
<b>Apgar 5º minuto</b>		
0	328	0,07%
1	498	0,10%
2	315	0,07%
3	436	0,09%
4	656	0,14%
5	1.260	0,26%
6	2.360	0,49%
7	5.304	1,11%
8	18.557	3,87%
9	152.361	31,79%
10	295.741	61,71%
99	56	0,01%
(vazio)	1.399	0,29%
<b>Classificação do RN ao nascer considerando curva de crescimento Intergrowth*</b>		
Pequeno para Idade Gestacional	7.178	1,50%
Adequado para Idade Gestacional	95.707	19,97%
Grande para Idade Gestacional	15.658	3,27%
(vazio)	360.728	75,27%
<b>Classificação do RN ao nascer considerando curva de crescimento Fenton*</b>		
Pequeno para Idade Gestacional	12.334	2,57%
Adequado para Idade Gestacional	101.764	21,23%
Grande para Idade Gestacional	4.554	0,95%
(vazio)	360.619	75,24%
<b>Malformação congênita</b>		
Sim	3.013	0,63%
Não	452.071	94,32%
Ignorado	1.400	0,29%
(vazio)	22.787	4,75%
<b>Classificação de risco ao nascer*</b>		
Risco	109.561	22,86%
Não risco	368.565	76,90%
Ignorado	362	0,08%
(vazio)	783	0,16%
<b>Dia da semana do nascimento*</b>		
Segunda-feira	78.426	16,36%
Terça-feira	72.733	15,18%
Quarta-feira	73.327	15,30%
Quinta-feira	73.101	15,25%
Sexta-feira	74.417	15,53%
Sábado	60.258	12,57%
Domingo	47.009	9,81%
<b>Óbito*</b>		
Não óbito	474.306	98,96%
Óbito por causas Mal Definidas	41	0,01%
Óbito por causas Não Classificada / Outras	4	0,00%
Óbitos Não Evitáveis	862	0,18%
Óbitos Redutíveis por Adequada Atenção ao Parto	509	0,11%

Óbitos Redutíveis por Adequado Controle na Gravidez	1.421	0,30%
Óbitos Redutíveis por Ações de Prevenção, Diagnóstico e Tratamento Precoces	996	0,21%
Óbitos Redutíveis por Imunoprevenção	14	0,00%
Óbitos Redutíveis por Parcerias com Outros Setores	1.118	0,23%

Fonte: A autora (2020).

\*Nota: variáveis criadas na etapa de pré-processamento pelas atividades de enriquecimento ou construção

Na Tabela 3, é possível verificar a distribuição por ano dos nascidos vivos conforme ocorrência do desfecho (não óbito e óbito) no conjunto de dados SIM-SINASC final (pós-linkage). À direita da coluna de óbito, está a respectiva estratificação por grupo de evitabilidade conforme classificação Seade. Assim, para o experimento 1, foram considerados todos os registros (479.271) e para os experimentos 2, 3 e 4, considerou-se somente os registros de nascidos vivos que foram a óbito (4.965).

Tabela 3 – Número de registros no conjunto de dados SIM-SINASC final (pós-linkage) por ocorrência do desfecho (óbito e não óbito) por ano, com estratificação dos óbitos por grupo de evitabilidade.

Ano de nascimento	Não óbito	Óbito	Grupo de evitabilidade do óbito							
			Mal Definido	Não classificado / Outro	Não Evitável	Redutível por Adequada Atenção ao Parto	Redutível por Adequado Controle na Gravidez	Ações de Prevenção, Diagnóstico e Tratamento	Redutível por Imunoprevenção	Redutível por Parcerias com Outros Setores
1999	29.174	334	10		85	18	21	128	1	71
2000	28.919	387	8		89	24	25	162	1	78
2001	26.735	352	6		85	33	20	144	1	63
2002	25.996	326	1		55	37	87	76	2	68
2003	24.493	308	2		59	35	96	48		68
2004	24.827	296			44	41	85	52	1	73
2005	24.160	277	4		41	27	92	47		66
2006	24.068	261	1		40	30	82	44		64
2007	23.747	247	3		29	28	88	39	1	59
2008	24.734	260	1	1	39	21	97	42		59
2009	24.343	212	1		30	22	80	23		56
2010	24.793	235			32	19	94	32		58
2011	24.635	235	2		45	23	77	34	2	52
2012	24.478	223	1	1	40	25	84	26	2	44
2013	24.724	210			29	24	81	24	2	50
2014	24.606	209			34	27	70	20		58
2015	24.283	223	1	1	25	31	89	28	1	47
2016	23.019	198		1	29	18	76	19		55
2017	22.572	172			32	26	77	8		29
<b>Total</b>	<b>474.306</b>	<b>4.965</b>	<b>41</b>	<b>4</b>	<b>862</b>	<b>509</b>	<b>1.421</b>	<b>996</b>	<b>14</b>	<b>1.118</b>

Fonte: A autora (2020).



### 5.3 MINERAÇÃO DE DADOS E PÓS-PROCESSAMENTO

O experimento 1 permitiu identificar, por ano, as variáveis que mais influenciaram a ocorrência do óbito. As árvores de decisão resultantes constam no Apêndice F e todas as variáveis encontradas foram esquematizadas na Figura 14 por ano, com respectiva posição na árvore de decisão, representada por números em ordem crescente do nó-raiz para o nó-folha; assim, quanto menor o número, mais forte a relação. A coloração cinza indica a ausência da variável no subconjunto de dados do ano e a grafia do ano em cor vermelha, que o algoritmo não conseguiu encontrar no subconjunto de dados padrões para explicar o desfecho, não gerando árvore de decisão, o que ocorreu no ano de 1999. Do total de 49 variáveis do SINASC utilizadas, 21 (42,85%) apareceram em pelo menos uma árvore de decisão.

Das 21 variáveis, quatro estão relacionadas a outros setores além da saúde, reforçando o modelo dos DSS: “CBO mãe no SINASC”, “bairro de residência”, “escolaridade materna em anos concluídos” e “escolaridade materna ensino concluído”. Dessas, somente “bairro de residência” tem relação mais forte, sendo o primeiro nó filho (posição 2) na árvore de decisão do ano de 2001.

Existem duas variáveis relacionadas à duração da gestação no SINASC, uma categórica ordinal (menos de 22 semanas; 22 a 27 semanas; 28 a 31 semanas; 32 a 36 semanas; 37 a 41 semanas; 42 semanas e mais; ignorado) e outra numérica (número absoluto de semanas). A categórica está presente em toda janela temporal deste estudo, enquanto a numérica foi inserida na DNV em 2012 e passou a ser digitada no SINASC em 2013; a partir de então, apareceu em árvores de decisão em posição mais próxima do nó-raiz que a variável categórica.

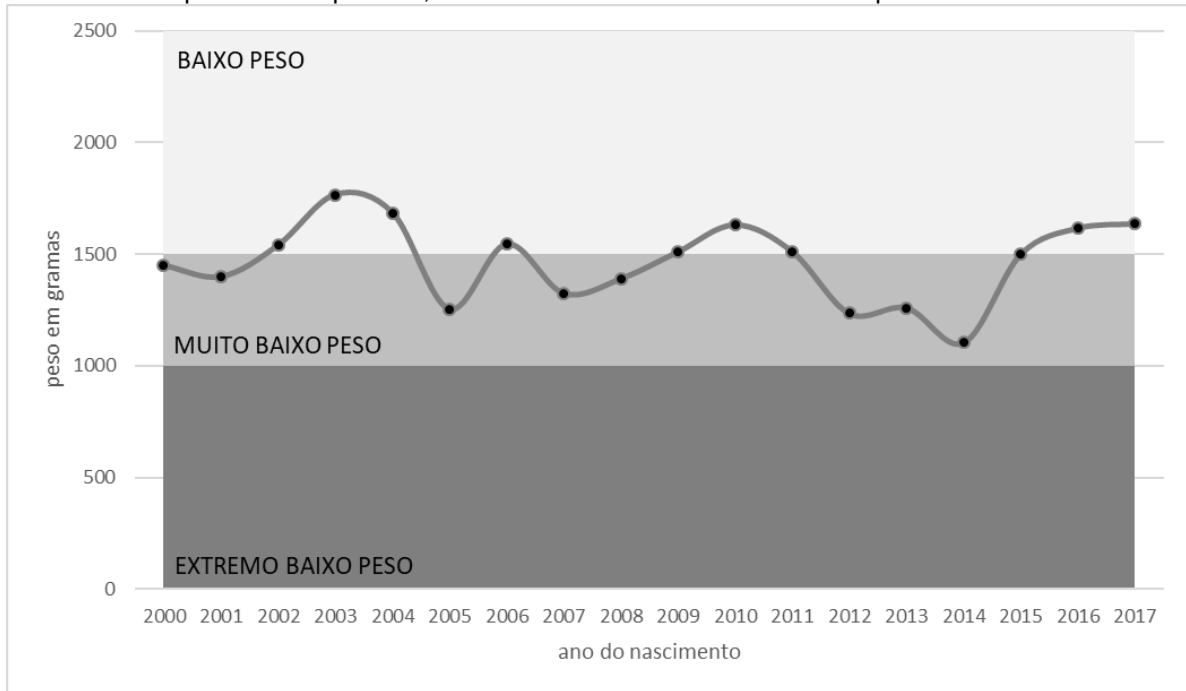
Figura 14 – Variáveis encontradas no experimento 1 por ano, com representação da força da relação ao óbito pela numeração em ordem crescente do nó-raiz ao nó-folha.

variável	ano																		
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
peso ao nascer		1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
malformação			2	2	2	2		2	2	2	2	2	3	2		2	3		2
apgar 5min		2									3		2	3		4	2		3
apgar 1min					3	4						3	4	4		5			
consultas pré-natal categorizada				4	5	3				4									
no. filhos mortos			3												4		5		
CBO mãe no SINASC			3									4							4
duração gestação numérica															2		4	2	
idade materna			3													3			
sexo				4	4														
duração gestação categorizada								3						4					
partos cesáreos anteriores															3		4		
bairro de residência			2																
no. filhos vivos			3																
escolaridade materna em anos concluídos			3																
dia da semana nascimento				3															
turno nascimento														3					
método estimação semanas gestação															3				
recuperação apgar																3			
escolaridade materna ensino concluído																	4		
classificação risco ao nascer																			3

Fonte: A autora (2020).

Identificou-se a forte relação entre óbito e “peso ao nascer”, pela presença dessa variável como nó-raiz em todos os anos, exceto em 1999, quando não houve árvore de decisão gerada. A partir da Figura 15, é possível identificar a variação dos valores de corte de peso ao nascer para cada ano, considerando os conceitos de baixo peso ao nascer. Apesar de não haver padrão de comportamento entre os anos, todos os valores de corte são considerados de baixo peso (< 2.500 g), ficando, em geral, mais próximos da classificação de muito baixo peso (< 1.500 g). Além de peso, as variáveis mais frequentes até o terceiro nó filho foram “malformação”, “Apgar no 5º minuto” e “Apgar no 1º minuto” (Figura 14). Estas e “peso ao nascer” têm sua relação com o óbito consolidada na literatura e foram excluídas do conjunto de dados para o experimento 2, no intuito de encontrar padrões que pudessem complementar o conhecimento existente por meio de padrões válidos, novos e úteis.

Figura 15 – Valores de corte de peso ao nascer encontrados em cada árvore de decisão do experimento 1 por ano, relacionados aos conceitos de baixo peso ao nascer.



Fonte: A autora (2020).

No experimento 2, à exceção do ano de 2015, as árvores de decisão geradas (Apêndice G) continham apenas uma variável. Elas estão esquematizadas na Figura 16, na qual o número em ordem crescente representa a posição do nó-raiz ao nó-folha, a coloração cinza indica a ausência da variável no subconjunto de dados do ano e a grafia do ano em cor vermelha, que o algoritmo não conseguiu encontrar no subconjunto de dados padrões explicação do desfecho, não gerando árvore de decisão, o que ocorreu nos anos de 1999, 2000, 2001, 2007 e 2013.

Figura 16 – Variáveis encontradas no experimento 2 por ano, com representação da força da associação ao óbito pela numeração em ordem crescente do nó-raiz ao nó-folha.

		ano																		
		1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017
variável	classificação de risco ao nascer					1						1	1		1		1	2		
	idade da criança em meses completos						1				1							1		
	duração gestação categorizada								1	1								3		
	idade da criança em dias				1									1						
	duração gestação numérica																		1	1
	local ocorrência óbito																		2	
	grupo de Robson																		3	
	componente do óbito																		4	

Fonte: A autora (2020).

Nota: A coloração cinza indica a ausência da variável nos respectivos anos e a grafia em cor vermelha do ano identifica aqueles em que não foi gerada árvore de decisão.

Do total de 66 variáveis desse experimento, oito (12,12%) apareceram em pelo menos uma árvore de decisão. Das variáveis encontradas, metade é oriunda do SINASC (“classificação de risco ao nascer”, “duração gestação categorizada”, “duração gestação numérica” e “grupo de Robson”) e metade, do SIM (“idade da criança em meses completos”, “idade da criança em dias”, “local ocorrência óbito” e “componente do óbito”). A variável “classificação de risco ao nascer” foi a que mais apareceu, porém sem padrão de comportamento entre os anos. As variáveis do SIM “idade da criança em meses completos”, “idade da criança em dias” e “componente do óbito” referem-se à idade da criança no momento do óbito em dias, meses ou período (neonatal precoce, neonatal tardio, pós-neonatal ou 1-2 anos) e apareceram sem padrão entre os anos.

Devido à redundância das variáveis relacionadas à idade da criança no momento do óbito e variação do nó-raiz entre os anos no experimento 2, realizou-se o experimento 3, com único conjunto de dados contendo os registros de todos os anos. Das quatro regras descobertas, as únicas variáveis presentes no antecedente foram “idade da criança em meses completos” e “classificação de risco ao nascer” e, no conseqüente, todos os óbitos eram evitáveis (Figura 17).

Figura 17 – Regras encontradas no experimento 3.

Regra	SE	Antecedente	ENTÃO	Consequente	Cobertura	Erro
1		idade da criança <=1 mês completo E classificação de risco ao nascer = RISCO		Óbito evitável por adequado controle na gravidez	3313.00	1968.07
2		idade da criança <=1 mês completo E classificação de risco ao nascer = NÃO RISCO		Óbito evitável por parceria com outros setores	505.99	273.85
3		idade da criança <=1 mês completo E classificação de risco ao nascer = IGNORADO		Óbito evitável por adequado controle na gravidez	3.01	2.01
4		idade da criança >1 mês completo		Óbito evitável por parceria com outros setores	1143.00	668.00

Fonte: A autora (2020).

Enquanto, no experimento 2, havia três variáveis relacionadas à idade da criança no momento do óbito, no experimento 3, destacou-se apenas “idade em meses completos”, com valores numéricos de zero a 24 meses. O valor de corte encontrado para essa variável foi de um mês completo, compreendendo, portanto, o período de zero a 60 dias. Na Figura 17, é possível verificar que, enquanto na regra 4 a única variável no antecedente foi “idade em meses completos”, nas regras 1, 2 e 3, “idade em meses completos” combinou-se à “classificação de risco ao nascer” para se relacionar ao grupo de evitabilidade do óbito. Assim, infere-se que a variável “classificação de risco ao nascer” possui relação com o óbito evitável nos primeiros meses de vida e que ações específicas da saúde poderiam evitá-lo, uma vez que na regra com a maior cobertura (Figura 17, regra 1) estima que 3.313 RNs com idade até um mês completo e de risco ao nascer foram a óbito RACG.

Desconsiderou-se a regra 3, devido à baixa cobertura frente às outras regras. As regras 2 e 4 possuem como desfecho RPOS, porém o erro demonstra que existem casos com o mesmo antecedente, mas com outro consequente, ou seja, outro grupo de redutibilidade de óbito. Para qualificar essas exceções, realizou-se o experimento 4, no qual foram descobertas 35 regras de exceção para a regra 2 e 15 para a regra 4.

Recuperando o critério adotado para seleção das regras de associação (considerar interessantes as regras de exceção cujo item de dado que altere o

desfecho se aproxima da camada de condições socioeconômicas, culturais e ambientais gerais, por ser de governabilidade direta do gestor conforme modelo dos DSS), identificaram-se as variáveis “DS de residência” e “vinculação SUS” como aquelas que alteraram o consequente da regra geral. Foram selecionadas as regras de exceção que continham essas variáveis, totalizando quatro regras de exceção para a regra 2 (Figura 18) e cinco para a regra 4 (Figura 19).

Figura 18 – Regra geral 2 e respectivas regras de exceção selecionadas no pós-processamento do experimento 4.

Regra	Antecedente	Consequente	Suporte	Confiança
geral 2	idade da criança <=1 mês completo E classificação de risco ao nascer = NÃO RISCO	óbito evitável por parceria com outros setores	10,1%	46,0%
exceção 2.1	idade da criança <=1 mês completo E classificação de risco ao nascer = NÃO RISCO E <b>DS residência = BN*</b>	óbito evitável por prevenção, diagnóstico e tratamento precoces	0,4%	30,0%
exceção 2.2	idade da criança <=1 mês completo E classificação de risco ao nascer = NÃO RISCO E <b>DS residência = SF*</b>	óbito evitável por prevenção, diagnóstico e tratamento precoces	0,5%	39,1%
exceção 2.3	idade da criança <=1 mês completo E classificação de risco ao nascer = NÃO RISCO E <b>DS residência = MZ*</b>	óbito evitável por prevenção, diagnóstico e tratamento precoces	0,2%	33,3%
exceção 2.4	idade da criança <=1 mês completo E classificação de risco ao nascer = NÃO RISCO E <b>DS residência = TQ*</b>	óbito evitável por prevenção, diagnóstico e tratamento precoces	0,2%	33,3%

Fonte: A autora (2020).

\*Nota: DS na árvore de decisão (Figura 18): BN = Bairro Novo, SF = Santa Felicidade, MZ = Matriz e TQ = Tatuquara.

Figura 19 – Regra geral 4 e respectivas regras de exceção selecionadas no pós-processamento do experimento 4.

Regra		Antecedente		Consequente	Suporte	Confiança
geral 4		idade da criança >1 mês completo		óbito evitável por parceria com outros setores	23,0%	41,6%
exceção 4.1	SE	idade da criança >1 mês completo E <b>DS residência = PR*</b>	ENTÃO	óbito evitável por prevenção, diagnóstico e tratamento precoces	0,7%	36,1%
exceção 4.2		idade da criança >1 mês completo E <b>DS residência = BN*</b>		óbito evitável por prevenção, diagnóstico e tratamento precoces	1,0%	32,7%
exceção 4.3		idade da criança >1 mês completo E <b>DS residência = CJ*</b>		óbito evitável por prevenção, diagnóstico e tratamento precoces	1,1%	36,4%
exceção 4.4		idade da criança >1 mês completo E <b>DS residência = BV*</b>		óbito evitável por prevenção, diagnóstico e tratamento precoces	1,3%	30,8%
exceção 4.5		idade da criança >1 mês completo E <b>vinculação = SUS</b>		óbito evitável por prevenção, diagnóstico e tratamento precoces	7,7%	32,8%

Fonte: A autora (2020).

\*Nota: DS da árvore de decisão (Figura 19): PR = Portão, BN = Bairro Novo, CJ = Cajuru e BV = Boa Vista.

O registro do DS no SINASC iniciou em 2008 e, até 2017, 2.177 crianças nascidas nesse período foram a óbito, das quais 446 (20,5%) não foram consideradas com risco ao nascer, conforme é possível notar na Tabela 4. Ainda, é possível identificar diferença desse percentual entre os DS - enquanto no DS de residência MZ 14,3% dos nascidos vivos que foram a óbito não eram de risco, no DS TQ esse percentual foi de 24,2%.

Tabela 4 – Frequência absoluta e relativa dos nascidos vivos entre 2008 e 2017 que foram a óbito, por DS e classificação de risco ao nascer.

DS de residência	Classificação de risco ao nascer								Total
	Não risco		Risco		Não classificado		Ignorado		
	<i>n</i>	%*	<i>n</i>	%*	<i>n</i>	%*	<i>n</i>	%*	
Bairro Novo	45	17,9	206	82,1					251
Boqueirão	54	22,5	186	77,5					240
Boa Vista	61	22,6	209	77,4					270
CIC	39	21,4	143	78,6					182
Cajuru	51	18,9	219	81,1					270
Matriz	18	14,3	108	85,7					126
Pinheirinho	65	24,1	205	75,9					270
Portão	40	19,4	166	80,6					206
Santa Felicidade	36	18,4	160	81,6					196
Tatuquara	16	24,2	50	75,8					66
em branco	21	21,0	61	61,0	16	16,0	2	2,0	100
<b>Total</b>	<b>446</b>	<b>20,5</b>	<b>1.713</b>	<b>78,7</b>	<b>16</b>	<b>0,7</b>	<b>2</b>	<b>0,1</b>	<b>2.177</b>

Fonte: A autora (2020).

Nota: \* Frequência relativa em relação ao total de óbitos do DS (linha).

Apesar da consulta ao CNES<sup>10</sup>, não foi possível resgatar o estabelecimento de saúde dos registros do período de 1999 a 2005. Do período de 2006 a 2017, o percentual de óbitos das crianças com vinculação ao SUS (1,1%) foi maior que daquelas não vinculadas (0,7%) (Tabela 5).

Tabela 5 – Frequência absoluta e relativa dos nascidos vivos entre 2006 e 2017, por vinculação ao SUS e ocorrência do óbito.

Vinculação SUS	Não óbito		Óbito		Total
	<i>n</i>	%*	<i>n</i>	%*	
SUS	153.956	98,9	1.689	1,1	155.645
Não SUS	131.087	99,3	907	0,7	131.994
Nascimento fora de Curitiba	1.808	97,7	42	2,3	1.850
Estabelecimento de nascimento ignorado	2.151	99,0	21	1,0	2.172
Nascimento ocorrido em domicílio/outro	1.000	97,5	26	2,5	1.026
<b>Total geral</b>	<b>290.002</b>	<b>99,1</b>	<b>2.685</b>	<b>0,9</b>	<b>292.687</b>

Fonte: A autora (2020).

Nota: \* Frequência relativa em relação ao total de cada tipo de vinculação (linha).

Os resultados das regras de exceção obtidas no pós-processamento das regras 2 e 4 evidenciaram que as variáveis “DS de residência” e “vinculação SUS” alteraram o desfecho RPOS para RAPDTP (Figuras 18 e 19), sugerindo que, a depender do DS de residência e da vinculação ao SUS, ações específicas da saúde,

<sup>10</sup> Disponível em: <http://cnes.datasus.gov.br>.



relacionadas à prevenção, diagnóstico e tratamento precoces seriam capazes de evitar o óbito.

Para os testes estatísticos, os grupos de óbito MD, NCO e RI foram desconsiderados por possuírem frequência relativa abaixo dos critérios estabelecidos – 0,01% no conjunto de dados contendo todos os nascidos vivos e 1% no conjunto de dados contendo somente os óbitos (Tabela 6). Os testes foram realizados para avaliar a relação de “peso ao nascer”, “anomalia”, “Apgar no 1º minuto”, “Apgar no 5º minuto”, “idade em meses completos”, “classificação de risco ao nascer”, “DS de residência” e “vinculação SUS” com o grupo de evitabilidade de óbito.

Tabela 6 – Frequência absoluta de todos os grupos de óbito e frequência relativa dos grupos de óbito em relação a todos os nascidos vivos e nascidos vivos que foram a óbito.

<b>Grupo de óbito</b>	<b><i>n</i></b>	<b>% em relação a todos os nascidos vivos (<i>n</i> = 479.271)</b>	<b>% em relação aos nascidos vivos que foram a óbito (<i>n</i> = 4.965)</b>
Não óbito	474.306	99,0%	-
Óbito mal definido	41	0,0%	0,8%
Óbito não classificado/outro	4	0,0%	0,1%
Óbito não evitável	862	0,2%	17,4%
Óbito evitável por adequada atenção ao parto	509	0,1%	10,3%
Óbito evitável por adequado controle na gravidez	1.421	0,3%	28,6%
Óbito evitável por prevenção, diagnóstico e tratamento precoces	996	0,2%	20,1%
Óbito evitável por imunoprevenção	14	0,0%	0,3%
Óbito evitável por parceria com outros setores	1.118	0,2%	22,5%
<b>Total de nascidos vivos</b>	<b>479.271</b>		
<b>Total de nascidos vivos que foram a óbito</b>	<b>4.965</b>		

Fonte: A autora (2020)

Constatou-se relação significativa para todas as variáveis testadas e o grupo de evitabilidade do óbito. Porém vale ressaltar que, para as variáveis quantitativas na comparação dois a dois, alguns grupos não demonstraram diferença significativa entre si, mas eram diferentes dos outros grupos. Informações sobre as variáveis, respectivas camadas do modelo dos DSS, experimento em que foram encontradas e valor de *p* (com especificação da comparação dois a dois das variáveis quantitativas) são apresentadas no Quadro 6.

Quadro 6 – Variáveis encontradas mais fortemente relacionadas ao desfecho, respectivas camadas do modelo dos DSS, experimento em que foram encontradas, teste estatístico e nível de significância.

Variável	Camada do modelo dos DSS					Experimento em que foi encontrada a variável	Teste estatístico e nível de significância
	Idade, sexo e fatores hereditários	Estilo de vida	Redes sociais e comunitárias	Condições de vida, trabalho e acesso a serviços de saúde	Condições socioeconômicas, culturais e ambientais gerais		
Peso ao nascer	✓					1	Kruskal-Wallis p < 0,001 Na comparação dois a dois, não apresentaram diferença significativa entre si, mas diferenciavam-se dos outros grupos: - NE e RAAP (p = 0,121)
Anomalia	✓					1	Qui-quadrado p < 0,001
Apgar no 1º minuto	✓					1	Kruskal-Wallis p < 0,001 Na comparação dois a dois, não apresentaram diferença significativa entre si, mas diferenciavam-se dos outros grupos: - NE e RACG (p = 1,000) - RAPDTP e RPOS (p = 0,051)
Apgar no 5º minuto	✓			✓		1	Kruskal-Wallis p < 0,001 Na comparação dois a dois, não apresentaram diferença significativa entre si, mas diferenciavam-se dos outros grupos: - NE e RACG (p = 1,000) - RAPDTP e RPOS (p = 1,000)
Idade em meses completos	✓					2 3	Kruskal-Wallis p < 0,001 Na comparação dois a dois, não apresentaram diferença significativa entre si, mas diferenciavam-se dos outros grupos: - RAAP e RACG (p = 1,000) - RAPDTP e RPOS (p = 1,000)
Classificação de risco ao nascer	✓			✓		2 3 4	Qui-quadrado p < 0,001
DS				✓	✓	4	Qui-quadrado p = 0,024
Vinculação ao SUS				✓	✓	4	Qui-quadrado p < 0,001

Fonte: A autora (2020).

## 6 DISCUSSÃO

Nesta seção, são discutidos os principais resultados obtidos a partir da pergunta norteadora; depois, são apresentadas as limitações e, por último, as perspectivas de trabalhos futuros.

### 6.1 QUAIS VARIÁVEIS DE GOVERNABILIDADE DIRETA DO GESTOR TÊM POTENCIAL DE SUBSIDIAR O DIRECIONAMENTO DE AÇÕES PARA REDUÇÃO DOS ÓBITOS EVITÁVEIS DE ZERO A DOIS ANOS DE IDADE NO MUNICÍPIO DE CURITIBA?

No experimento 1, o algoritmo corroborou a literatura ao encontrar “malformação congênita”, “Apgar no 1º minuto”, “Apgar no 5º minuto” e “peso ao nascer” como variáveis mais fortemente relacionadas ao óbito.

“Malformação congênita” foi relacionado à mortalidade na Europa (GROEN et al., 2017), América Central (GARCÍA et al., 2014) e Brasil (VELOSO et al., 2019); no entanto, não faz parte dos critérios de risco ao nascer sugeridos pelo MS (BRASIL, 2014a) ou adotados pela SMS de Curitiba (CURITIBA, 2019b).

“Apgar no 1º minuto” representa a conjuntura do parto, enquanto “Apgar no 5º minuto” está relacionado à resposta à assistência ao nascimento e ao prognóstico neonatal (BOUZADA et al., 2018). “Apgar no 5º minuto”, quando menor que 7, é considerado critério de risco ao nascer (CURITIBA, 2019b), o que foi confirmado nas árvores de decisão pela relação dos valores maiores que 7 a não óbitos (anos de 2009, 2014, 2015 e 2017). Encontraram-se também valores de corte de 3 e 5 (anos de 2000 e 2011), suscitando análise mais detalhada. Existe associação entre óbitos neonatais e baixo índice de Apgar no quinto minuto, apesar de esse não ser determinante isolado para desfechos de longo prazo (VARELA et al., 2019), o que pode ser identificado nas árvores de decisão pelo relacionamento do “Apgar no 5º minuto” a outras variáveis como “peso ao nascer” e/ou “malformação congênita” para relacionar-se ao óbito.

O peso ao nascer inferior a 2.500 g é classificado como baixo peso e, independentemente da idade gestacional, considerado um dos principais fatores de risco para a morbimortalidade infantil (UNITED NATIONS CHILDREN’S FUND; WORLD HEALTH ORGANIZATION, 2004; VELOSO et al., 2019). A variável “peso

ao nascer” foi o nó-raiz nas árvores de decisão geradas no experimento 1, destacando sua forte relação com o óbito. Apesar de ser considerado risco ao nascer o peso menor que 2.500 g (BRASIL, 2014a; CURITIBA, 2019b), os valores de corte encontrados variaram entre os anos, porém se aproximaram do muito baixo peso (inferior a 1.500 g), aproximando-se ao encontrado por Lansky et al. (2014).

Ainda no experimento 1, a duração da gestação apareceu como segundo nó quando passou a ser registrada também de forma numérica, evidenciando que a coleta do dado bruto permite maior flexibilidade de análise (PINHEIRO; TORRES, 2009). Sua categorização em anos anteriores pode subestimar a prematuridade (CURITIBA, 2019b), considerada fator de risco à mortalidade (TEIXEIRA et al., 2016; DEMITTO et al., 2017). Os valores de corte encontrados foram a partir de 28 a 31 semanas para a variável categórica e de 30 e 31 semanas para a variável numérica, ambas abaixo do valor considerado critério de risco – menor que 37 semanas (BRASIL, 2014a; CURITIBA, 2019b).

Assim como a duração da gestação, o registro do número de consultas acontece tanto de forma numérica quanto categorizada no SINASC e somente a categorizada apareceu nas árvores de decisão. Não é sugerido como critério pelo MS (BRASIL, 2014a), porém é adotado por Curitiba quando menor que quatro (CURITIBA, 2019b). Contudo, houve relação de mais de sete consultas com o óbito, bem como nenhuma consulta ao não óbito, possibilitando criar hipóteses que convergem ao encontrado por Koshida et al. (2016) de que, mais que o número de consultas, a assistência obstétrica de qualidade contribui para redução dos óbitos neonatais.

A escolaridade e idade materna, quantidade de filhos vivos e mortos são critérios considerados pela SMS de Curitiba para classificação de risco ao nascer e apareceram como terceiro, quarto e quinto nós nas árvores de decisão, demonstrando que outras variáveis antecedentes devem ser consideradas para a sua relação com o óbito. De 2008 a 2017 houve redução de 63,15% de mães com ensino fundamental incompleto e de 44,7% na proporção de gravidez na adolescência, quando Curitiba alcançou a proporção de 9,0%, a menor entre todas as capitais brasileiras (CURITIBA, 2019b). Essas reduções são reflexos de melhores condições sociais (CURITIBA, 2018a) e de ações integradas de cuidado continuado e coordenado à saúde materno-infantil no município, norteadas pela Rede Mãe Curitibana Vale a Vida, implantada em 1999 (CURITIBA, 2019c).

As variáveis com relação mais forte ao óbito evidenciadas no experimento 1 estão relacionadas à camada mais interna do modelo dos DSS, referentes às características biológicas e individuais. Estas, são influenciadas pelas camadas mais externas (ORGANIZAÇÃO MUNDIAL DA SAÚDE; COMISSÃO PARA OS DETERMINANTES SOCIAIS DA SAÚDE, 2010; MOORE et al., 2017), das quais identificaram-se as variáveis relacionadas a serviços de saúde e bairro de residência, que apareceu com forte relação somente no ano de 2001. O local de residência pode ser fator de risco para a mortalidade, pois está relacionado a outros fatores, como saneamento básico, renda e esgotamento sanitário (ALMEIDA; GOMES; NASCIMENTO, 2014; BÜHLER et al., 2014).

Para o cumprimento da premissa do KDD de complementar o conhecimento existente a partir de padrões novos, válidos e úteis (FAYYAD et al., 1996), a exclusão das variáveis consolidadas na literatura como fatores de risco (peso, malformação, Apgar no primeiro e quinto minuto) permitiu pressupor que “classificação de risco ao nascer” possuía relação com o óbito visto que apareceu como nó-raiz em árvores de decisão do experimento 2. Tal suposição ganhou força no experimento 3, quando tal variável foi uma das duas únicas presentes nos antecedentes das regras da árvore de decisão.

A árvore de decisão do experimento 3 reforçou o potencial da relação da classificação de risco com o óbito evitável, junto da idade da criança em meses completos. Segundo a United Nations Children’s Fund (2018), o primeiro mês de vida é o período mais crucial para a sobrevivência infantil. O óbito ocorrido até 28 dias faz parte do componente neonatal, que corresponde a aproximadamente 70% dos óbitos de menores de um ano no Brasil (FRANÇA; LANSKY, 2016). Há concentração dos óbitos de menores de um ano em Curitiba no período neonatal (CURITIBA, 2018a); contudo, o valor de corte encontrado foi menor ou igual a um mês completo, correspondendo ao período de zero a 60 dias, além dos 28 do período neonatal.

Na qualificação das exceções (experimento 4), evidenciou-se que as variáveis “DS de residência” e “vinculação SUS” alteram o desfecho RPOS para RAPDTP, sendo óbito evitável por ação da saúde. Essas variáveis, aproximadas da camada de condições socioeconômicas e ambientais gerais e, portanto, de governabilidade direta do gestor, reforçam a influência dos DSS na mortalidade nos primeiros 1.000 dias de vida. Piores condições socioeconômicas, de moradia, saneamento básico e

renda estão associados ao óbito infantil (BARBOSA et al., 2014; BÜHLER et al., 2014; SADOVSKY et al., 2018). Silva e Paes (2019) identificaram que o Programa Bolsa Família teve impacto positivo na redução dos níveis de mortalidade infantil ao tentar superar a desigualdade pela distribuição de renda condicional para famílias pobres. Corrobora-se assim que a redução do óbito evitável pode se dar por mudanças estruturais, como políticas públicas, com vistas à redução de iniquidades em saúde (DAHLGREN; WHITEHEAD, 1991).

O baixo nível socioeconômico é sugerido como critério de risco pelo MS (BRASIL, 2014a), mas não adotado por Curitiba (2019b). Apesar de Curitiba possuir IDHM muito alto – 0,823 (CURITIBA, 2018b), nota-se heterogeneidade no índice por Unidade de Desenvolvimento Humano (UDH)<sup>11</sup>, com variação de 0,50 (baixo) a 0,96 (muito alto) (PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO, 2020). Assim, suscita-se a necessidade de reflexão sobre a classificação de risco ao nascer utilizado atualmente por Curitiba, com ampliação de seus critérios, considerando principalmente as vulnerabilidades específicas de certos locais. A reavaliação periódica dos critérios de risco possibilita a captação e vigilância das crianças conforme perfil epidemiológico atual (SOARES et al., 2019). Existe limitação na DNV no que tange à coleta de dados socioeconômicos e ambientais gerais (SOARES et al., 2019), mas em Curitiba tais características poderiam ser inferidas pelo DS de residência e vinculação ao SUS.

Leal (2020) identificou desigualdades regionais, barreiras no acesso e inadequação da atenção pré-natal no serviço público. A relação entre a qualidade do pré-natal e as desigualdades sociais pode ser demonstrada com o aumento do percentual de pré-natal adequado em gestantes de maior renda, em municípios maiores e com melhores IDHs (TOMASI et al., 2017). Em Curitiba, estima-se que 50% da população é beneficiária de plano de saúde<sup>12</sup>, no entanto, Silva, Luiz e Bahia (2019) sinalizaram a relação entre vinculação a um plano de saúde de menor

---

<sup>11</sup> “As Unidades de Desenvolvimento Humano (UDH) são áreas dentro das regiões metropolitanas que podem ser uma parte de um bairro, um bairro completo ou, em alguns casos, até um município pequeno. A homogeneidade socioeconômica é o que define os limites das UDHs, que são formadas a partir da agregação dos setores censitários do IBGE.” (PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO, 2020. Disponível em: <https://www.br.undp.org/content/brazil/pt/home/idh0/atlas-do-desenvolvimento-humano/atlas-do-desenvolvimento-humano-nas-regioes-metropolitanas-brasi.html>.

<sup>12</sup> Cálculo realizado pela autora, resultante do quociente entre número de beneficiários de plano de saúde em dezembro 2019 (<http://www.ans.gov.br/perfil-do-setor/dados-e-indicadores-do-setor>) e população estimada para 2019 (<https://www.ibge.gov.br/cidades-e-estados/pr/curitiba.html>).

abrangência (sem cobertura de internações) e utilização da rede de serviços do SUS. Portanto, há de se considerar a possibilidade de o SUS absorver casos de maior complexidade e conseqüentemente com maior risco de ocorrência de desfechos indesejáveis, logo, seria leviana a conclusão simplista de baixa qualidade assistencial do SUS como causa do óbito. De qualquer maneira, as ações relacionadas à variável vinculação SUS foram consideradas como de governabilidade direta do gestor, o que justifica o seu destaque dentre as variáveis identificadas.

A influência do DS de residência pode ser demonstrada pela diferença das TMI's dos DSs de Curitiba, com taxas que variaram de 5,1 a 15,1/1.000 nascidos vivos no ano de 2016 (preliminar), conforme pode ser identificado na Figura 20 (CURITIBA, 2018a). Os determinantes sociais da saúde quando desiguais podem causar iniquidades em saúde sistemáticas, relevantes, evitáveis, injustas e desnecessárias (DAHLGREN; WHITEHEAD, 1991; WORLD HEALTH ORGANIZATION, 2011).

Figura 20 – TMI's por DS e ano de 2007 a 2016 (preliminar), no município de Curitiba.

<b>Distritos Sanitários</b>	<b>2007</b>	<b>2008</b>	<b>2009</b>	<b>2010</b>	<b>2011</b>	<b>2012</b>	<b>2013</b>	<b>2014</b>	<b>2015</b>	<b>2016*</b>
<b>Santa Felicidade</b>	8,4	7,4	10,7	10,3	8,1	11,9	8,7	9,1	6,8	8,4
<b>Boa Vista</b>	10,7	8,3	8,6	7,2	8,1	5,8	8,8	6,2	7,9	6,0
<b>Boqueirão</b>	10,3	12,0	8,5	8,9	8,8	7,8	7,1	6,6	6,6	11,9
<b>Portão</b>	7,5	7,8	7,6	8,6	9,8	6,9	9,7	8,4	7,9	6,0
<b>Pinheirinho</b>	11,9	12,9	14,8	14,0	10,7	10,7	9,2	9,0	12,4	7,5
<b>Cajuru</b>	15,4	11,3	8,06	9,3	9,7	10,7	6,8	8,4	10,2	10,3
<b>Matriz</b>	7,8	8,2	6,5	5,3	5,9	8,0	9,4	4,9	8,1	5,1
<b>Bairro Novo</b>	11,2	9,3	6,3	8,9	11,2	9,7	8,6	9,0	9,9	8,3
<b>CIC</b>	10,5	11,1	9,4	8,9	5,8	15,0	11,4	7,2	10,5	8,3
<b>Tatuquara**</b>	0	0	0	0	0	0	0	0	0	15,1
<b>Curitiba</b>	10,5	9,9	9,0	9,1	8,7	9,5	8,8	7,7	8,9	8,6

Fonte: SMS/CE/CEV/SIM/SINASC \* 2016- dado preliminar. \*\*O DS Tatuquara foi criado em 2016

Fonte: Curitiba (2018a, p. 72).

Como estratégia para promoção da equidade no SUS de Curitiba, instituiu-se em 2017 o Índice de Vulnerabilidade das Áreas de Abrangência das Unidades Municipais de Saúde (IVAB), calculado a partir do Índice de Vulnerabilidade das Famílias do Paraná (IVF-PR) do Instituto Paranaense de Desenvolvimento

Econômico e Social (Ipardes) e da população do Censo do Instituto Brasileiro de Geografia e Estatística (IBGE) por área de abrangência de UMS. Calcula-se o quociente entre o número de pessoas de alta vulnerabilidade de acordo com o IVF-PR e a população geral da área de abrangência da UMS, multiplica-se por 100 e ordenam-se as UMSs em ordem crescente do índice obtido, dividindo-se em tercís: baixo IVAB, médio IVAB e alto IVAB (CURITIBA, 2018c). O IVAB é indicativo de condições socioeconômicas e ambientais já utilizado pela saúde e poderia ser utilizado para ampliar os critérios de classificação de risco ao nascer.

A classificação de risco destacada é de risco ao nascer, porém vale lembrar que situações de risco acontecem também no decorrer do crescimento e evolução da criança e devem ser consideradas na sua atenção e vigilância (BRASIL, 2014a). O risco deve ser constantemente reavaliado no acompanhamento pela UMS, pois uma criança que não apresentava risco ao nascer pode passar a tê-lo de acordo com fatores de risco evolutivos, como internações ou infecções repetidas, desnutrição instalada, desenvolvimento psicomotor insatisfatório para a idade, casos na família de tuberculose, HIV/AIDS, drogadição, entre outros (CURITIBA, 2012). No entanto, a classificação de risco ao nascer realizada a partir de dados da DNV é importante, pois orienta as ações das equipes de saúde e desencadeia a busca ativa dessas crianças (CURITIBA, 2012, 2019b).

Os resultados reforçam a abrangência e interrelação das variáveis relacionadas ao desfecho, alicerçado pelo modelo dos DSS. Ainda, permitiram evidenciar a relação da classificação de risco ao nascer e o óbito evitável por ações da saúde, demonstrando maior potencial quando levados em consideração o DS de residência e a vinculação ao SUS, variáveis que representam condições socioeconômicas e ambientais e podem levar a iniquidades em saúde. De acordo com Marmot (2005), não existe razão biológica para diferenças de expectativas de vida e reduzir as iniquidades sociais é questão de justiça social.

## 6.2 LIMITAÇÕES

Apesar de SIM e SINASC possuírem dados complementares com grande potencial de geração de informação, quando analisados em conjunto, eles não são integrados, sendo necessário o *linkage*. Em estudo que avaliou o *linkage* dos conjuntos de dados desses SISs em todas as capitais brasileiras, Curitiba, Porto



Alegre, Campo Grande, Vitória e Recife tiveram a melhor taxa de pareamento (acima de 97%) (MAIA et al., 2017). Apesar de o CE/Coordenação de Eventos Vitais da SMS de Curitiba realizar a qualificação dos dados preenchidos nos instrumentos de registro (CURITIBA, 2019b), foram identificados registros incompletos com número de DNV e data de nascimento em branco, com erros de digitação como nome da mãe escritos de diferentes maneiras e com inconsistências como mesmo número de DNV ou DO em diferentes registros. Essas falhas podem ter influenciado no *linkage* e, conseqüentemente, nos resultados.

A qualidade do preenchimento dos dados pode limitar o potencial dos SISs na geração de informação (SILVA et al., 2014; FERNANDES; CHIAVEGATTO FILHO, 2019) e deve ser considerado no KDD, em que a influência da qualidade dos dados pode ser resumida pelo acrônimo GIGO (*Garbage In, Garbage Out*<sup>13</sup>) (GOLDSCHMIDT; PASSOS; BEZERRA, 2015). Um dos maiores desafios a serem vencidos no SINASC é a ausência ou incompletude dos dados (BONILHA et al., 2018). Falhas no preenchimento podem ter deixado de parear ou pareado registros equivocadamente. Essas falhas podem ocorrer por diversos fatores – falta de clareza nos manuais, diversidade de profissionais responsáveis pelo preenchimento, preenchimento de algumas variáveis em detrimento de outras, falha na digitação etc. (AGRANONIK; JUNG, 2019). Podem ser minimizadas por sensibilização e capacitação dos profissionais envolvidos e redução de etapas, uma vez que a coleta e digitação desses dados acontecem em tempos e locais diferentes. Estratégias informatizadas do MS, como o e-SUS (BRASIL, 2014b), são exemplos de redução de etapas, com coleta e digitação ao mesmo tempo, encurtando caminhos no repasse dos dados.

A limitação do *software* utilizado também pode ter influenciado nos resultados. O *WEKA* foi escolhido por ser um *software* livre, com interface gráfica amigável e intuitiva (WITTEN; FRANK, 2005), além da familiaridade da autora com a ferramenta. No entanto, seu uso para grandes conjuntos de dados pode ser desafiador pois a sua linguagem de programação em Java necessita de quantidade de memória no computador geralmente maior que o disponibilizado<sup>14</sup>. Apesar de ter sido aumentada a memória disponibilizada para o ambiente Java, ainda assim foi

---

<sup>13</sup> “Entra lixo, sai lixo” (tradução livre da autora)

<sup>14</sup> <https://www.cs.waikato.ac.nz/ml/weka/bigdata.html>

necessário realizar a segmentação do conjunto de dados e diversos experimentos. Salienta-se que existem alternativas sugeridas pelos criadores<sup>9,15</sup> para análise de grandes conjuntos de dados, mas que não foram possíveis de serem concretizadas em virtude do tempo disponível.

Embora o uso de dados de SISs consolidados em território nacional (BRASIL, 2011a, 2011b) vise a garantir a replicação do estudo para outros municípios brasileiros, há de se considerar o possível esgotamento de análise de suas variáveis (SARTORELLI et al., 2017). As atividades de enriquecimento e construção na etapa de pré-processamento (GOLDSCHMIDT; PASSOS; BEZERRA, 2015) aumentaram as possibilidades de descoberta de padrões válidos, novos e potencialmente úteis, que foram confirmadas pelo encontro da relação do óbito com as variáveis “classificação de risco ao nascer”, “idade da criança em meses completos”, “DS de residência” e “vinculação SUS”. A limitação na DNV no que tange à coleta de dados socioeconômicos e ambientais gerais (SOARES et al., 2019) poderia ser suprimida pela integração com bases de dados de outros sistemas de informação, mas que não foi possível de ser executada pela dificuldade de acesso e tempo.

Segundo Goldschmidt, Passos e Bezerra (2015), as atividades realizadas considerando o processo KDD podem ser organizadas em três grandes grupos – desenvolvimento tecnológico, que abrange a concepção, desenvolvimento, aprimoramento e otimização de algoritmos; execução do KDD, que envolve a utilização dos algoritmos já desenvolvidos na busca por padrões que possam complementar o conhecimento existente e, por último, aplicação dos resultados, que envolve a utilização dos resultados no contexto em que foi realizado o processo KDD. Na revisão integrativa identificou-se que todos os estudos foram pontuais, de desenvolvimento tecnológico ou de execução do KDD. Este estudo limitou-se à execução do KDD, reforçando a necessidade de aplicação na prática dos resultados obtidos e, principalmente, avaliação do emprego dessa tecnologia em saúde para apoio à gestão no enfrentamento da mortalidade nos primeiros 1.000 dias de vida. A Avaliação de Tecnologias em Saúde (ATS) contribui para melhorar o processo decisório em todos os pontos de atenção em saúde, desde a assistência prestada pelos profissionais ao estabelecimento de políticas públicas (ELIAS, 2017).

---

<sup>15</sup> [https://waikato.github.io/weka-wiki/classifying\\_large\\_datasets/](https://waikato.github.io/weka-wiki/classifying_large_datasets/)

### 6.3 PERSPECTIVAS DE TRABALHOS FUTUROS

Considerando os resultados obtidos e as limitações identificadas, é possível vislumbrar algumas sugestões de trabalhos futuros:

- a) avaliar a qualidade dos dados dos SISs pela verificação de características como cobertura, completude e consistência. Seus resultados poderiam indicar com mais precisão pontos a ser melhorados, possibilitando a proposição de capacitações mais bem direcionadas no que tange à coleta e registro do dado e avaliações sistemáticas e automatizadas para detecção de inconsistências;
- b) criar sistema que permita integração automática dos SISs;
- c) avaliar variáveis coletadas nos SISs e sua contribuição efetiva para a tomada de decisão, com proposição de inclusão de mais variáveis que representem condições socioeconômicas e ambientais gerais ou a inserção de chave de identificação única presente em sistemas de informação para além da saúde, como, por exemplo, o Cadastro de Pessoa Física, o que possibilitaria *linkage* de bases de dados oriundas de outras áreas, como da assistência social;
- d) analisar as regras não exploradas;
- e) realizar experimentos agregando dados de outras fontes de informação, como prontuário eletrônico do paciente e outros SISs, como o Sistema de Informação de Agravos de Notificação, o Sistema de Informações Hospitalares e sistemas de informação de outros setores, abrangendo variáveis de todas as camadas do modelo dos DSS;
- f) ampliar o escopo para os óbitos fetais e aborto, abrangendo todo o período dos primeiros 1.000 dias de vida;
- g) realizar o processo KDD no WEKA com as alternativas sugeridas pelos criadores para grandes conjuntos de dados ou utilizar outras linguagens de programação ou ferramentas, como por exemplo Python e Rapid Minner, que permitam trabalhar o conjunto de dados de maior magnitude que o próprio WEKA;
- h) propor a automatização, no SINASC, da classificação de risco ao nascer no SINASC;

- i) reavaliar os critérios de classificação de risco ao nascer, considerando a redução no valor de corte de peso ao nascer e ampliação dos critérios com inclusão de condições socioeconômicas e ambientais, que poderiam ser representadas pelo IVAB;
- j) avaliar o emprego do processo KDD para apoio à gestão na organização da atenção em saúde.

## 7 CONSIDERAÇÕES FINAIS

O modelo dos DSS representa as influências no desenvolvimento da criança nos primeiros 1.000 dias de vida. Dada a complexidade dessa interação, apesar das ações que vem sendo realizadas, óbitos nesse período continuam ocorrendo, com destaque para os evitáveis. A reconstrução de coorte histórica de janela temporal de quase 20 anos e mais de 400 mil registros, aliada à abordagem mista de análise de dados, que compreendeu algoritmos, estatística descritiva e testes estatísticos convencionais, permitiu identificar dificuldades e fragilidades no percurso de transformação do dado em informação, corroborar fatores relacionados à mortalidade já estabelecidos na literatura e descobrir novas variáveis de governabilidade direta do gestor com potencial de contribuição no enfrentamento do problema, atingindo o objetivo proposto.

As limitações foram consideradas no desenvolvimento do trabalho e apontadas na discussão, destacando-se a qualidade dos conjuntos de dados, a falta de integração entre os SISs e possível esgotamento de análise de suas variáveis, uma vez que representam relações com a mortalidade já consolidadas na literatura.

Apesar de o estudo retratar os achados de um único município, o desfecho estudado não é problema exclusivo dele. Para garantir sua replicação, utilizaram-se SISs de abrangência nacional presentes em todos os municípios brasileiros e *softwares* livres, rodados em computadores domésticos. O *script* obtido permite a replicação do *linkage*, superando a falta de integração entre os SISs.

A constatação da limitada quantidade de variáveis representativas da camada mais externa do modelo dos DSS nos SISs utilizados sugere a necessidade de reavaliação dos dados coletados, com inclusão de novas variáveis que representem condições socioeconômicas e ambientais gerais ou inserção de chave de identificação única presente em sistemas de informação para além da saúde, como, por exemplo, o Cadastro de Pessoa Física, o que possibilitaria *linkage* de bases de dados oriundas de outras áreas, como da assistência social.

Recuperando a contribuição científica de identificar valores diferentes dos estabelecidos na literatura para fatores de risco já conhecidos e descobrir novas variáveis relacionadas ao óbito, considerando as especificidades locais e a governabilidade do gestor sobre elas, consideram-se como principais achados o valor de corte de peso ao nascer abaixo de 2.500 g, a necessidade de atenção mais

vigilante para além de 30 dias e o potencial da classificação de risco ao nascer relacionado com o óbito evitável, especialmente quando combinado a variáveis ligadas às camadas mais externas do modelo dos DSS – DS de residência e vinculação ao SUS.

A relação do óbito evitável por ações da saúde à classificação de risco ao nascer, DS de residência e vinculação ao SUS reforça a responsabilidade direta dos gestores na redução da mortalidade nos primeiros 1.000 dias de vida por meio de ações que visem à melhoria da atenção à saúde materno-infantil e redução de iniquidades em saúde. Ainda, suscita a necessidade de reavaliação dos critérios de risco ao nascer atualmente utilizados, com sugestão de inclusão de critérios que representem condições socioeconômicas e ambientais, que poderiam ser representadas pelo IVAB. Dessa maneira, contempla-se a contribuição social de poder subsidiar os gestores na tomada de decisão, com vistas à redução de iniquidades em saúde, além da melhora da qualidade da atenção prestada à saúde materno-infantil.

## REFERÊNCIAS

- AGRANONIK, M.; JUNG, R. O. Qualidade dos sistemas de informações sobre nascidos vivos e sobre mortalidade no Rio Grande do Sul, Brasil, 2000 a 2014. **Ciência e Saúde Coletiva**, Rio de Janeiro, v. 24, n. 5, p. 1945-1958, maio 2019. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-81232019000501945&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232019000501945&lng=en&nrm=iso). Acesso em: 20 dez. 2019.
- AGRAWAL, R.; IMIELINSKI, T.; SWAMI, A. Mining association rules between sets of itens in large databases. *In*: ACM SIGMOD CONFERENCE MANAGEMENT OF DATA, 1993, New York. **Proceedings [...]**. [S.l.: sn], 1993.
- AGRAWAL, R.; SRIKANT, R. Fast algorithms for mining association rules. *In*: INTERNATIONAL CONFERENCE ON VERY LARGE DATA BASES, 20., 1994, Santiago de Chile. **Proceedings [...]**. [S.l.: s.n.], 1994. p. 487-499.
- ALMEIDA, M. C. S.; GOMES, C. M. S.; NASCIMENTO, L. F. C. Análise espacial da mortalidade neonatal no estado de São Paulo, 2006-2010. **Revista Paulista de Pediatria**, São Paulo, v. 32, n. 4, p. 374-380, dez. 2014. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0103-05822014000400374&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-05822014000400374&lng=en&nrm=iso). Acesso em: 5 maio 2018.
- ALMEIDA FILHO, N.; BARRETO, M. L. Desenho de pesquisa em epidemiologia. *In*: ALMEIDA FILHO, N.; BARRETO, M. L. **Epidemiologia & saúde: fundamentos, métodos, aplicações**. Rio de Janeiro: Guanabara Koogan, 2014. p. 165-174.
- APGAR, V. A proposal for a new method of evaluation of the newborn infant. **Current Researches in Anesthesia and Analgesia**, v. 32, n. 4, p. 260-267, 1953. Disponível em: <https://profiles.nlm.nih.gov/ps/access/cpbkkg.pdf>. Acesso em: 31 ago. 2019.
- APGAR, V. **Evaluation of the newborn infant-second report**. 1958. Disponível em: <https://jamanetwork.com/journals/jama/article-abstract/325068>. Acesso em: 31 ago. 2019.
- AQUINO, E. M. L.; BARRETO, S.; SZKLO, M. Estudos de coorte. *In*: ALMEIDA FILHO, N.; BARRETO, M. L. **Epidemiologia & saúde: fundamentos, métodos, aplicações**. Rio de Janeiro: Guanabara Koogan, 2014. p. 203-214.
- BARBOSA, T. A. G. S. et al. Determinantes da mortalidade infantil em municípios do Vale do Jequitinhonha, Minas Gerais, Brasil. **Revista Mineira de Enfermagem**, Belo Horizonte, v. 18, n. 4, p. 907-914, 2014.
- BLACK, R. E. et al. Maternal and child undernutrition: global and regional exposures and health consequences. **The Lancet**, v. 371, n. 9608, p. 243-260, 2008. Disponível em: [https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736\(07\)61690-0.pdf](https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(07)61690-0.pdf). Acesso em: 26 nov. 2018.
- BLACK, R. E. et al. Maternal and child undernutrition and overweight in low-income and middle-income countries. **The Lancet**, v. 382, n. 9890, p. 427-451, 2013.

Disponível em: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(13\)60937-X/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(13)60937-X/fulltext). Acesso em: 15 nov. 2018.

BONILHA, E. A. et al. Cobertura, completude e confiabilidade das informações do Sistema de Informações sobre Nascidos Vivos de maternidades da rede pública no município de São Paulo, 2011. **Epidemiologia e Serviços de Saúde**, Brasília, DF, v. 27, n. 1, e201712811, 2018. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S2237-96222018000100310&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S2237-96222018000100310&lng=en&nrm=iso). Acesso em: 3 jan. 2020.

BORGELT, C. **Apriori**. 2010. Disponível em: <http://www.borgelt.net/doc/apriori/apriori.html>. Acesso em: 8 nov. 2019.

BOUZADA, M. C. F. et al. Resposta aos procedimentos de reanimação neonatal no quinto minuto de vida em recém-nascidos Apgar  $\leq 3$  no primeiro minuto. **Revista Médica de Minas Gerais**, Belo Horizonte, v. 28, supl. 6, S280608, 2018. Disponível em: <http://www.rmmg.org/artigo/detalhes/2416>. Acesso em: 3 jun. 2019.

BRASIL. Ministério da Saúde. **Biblioteca Virtual em Saúde**, Vigilância em Saúde, Sistemas de Informação, 2008. Disponível em: [http://bvsmms.saude.gov.br/bvs/svs/inf\\_sist\\_informacao.php](http://bvsmms.saude.gov.br/bvs/svs/inf_sist_informacao.php). Acesso em: 10 nov. 2018.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância Epidemiológica. **Guia de vigilância epidemiológica**. 7. ed. Brasília, DF: MS, 2009a. Disponível em: <http://www.epi.uff.br/wp-content/uploads/2013/10/Guia-de-Vigil%C3%A2ncia-Epidemiol%C3%B3gica-%E2%80%93-7%C2%AAedi%C3%A7%C3%A3o-2010.pdf>. Acesso em: 10 nov. 2018.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Secretaria de Atenção à Saúde. **Manual de vigilância do óbito infantil e fetal e do Comitê de Prevenção do Óbito Infantil e Fetal**. 2. ed. Brasília, DF: MS, 2009b. Disponível em: [http://bvsmms.saude.gov.br/bvs/publicacoes/manual\\_obito\\_infantil\\_fetal\\_2ed.pdf](http://bvsmms.saude.gov.br/bvs/publicacoes/manual_obito_infantil_fetal_2ed.pdf). Acesso em: 10 nov. 2019.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde. **Manual de instruções para o preenchimento da declaração de nascido vivo**. Brasília, DF: MS, 2011a. Disponível em: [http://www.saude.ms.gov.br/wp-content/uploads/sites/88/2015/11/inst\\_dn.pdf](http://www.saude.ms.gov.br/wp-content/uploads/sites/88/2015/11/inst_dn.pdf). Acesso em: 9 abr. 2018.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise de Situação de Saúde. **Manual de instruções para o preenchimento da declaração de óbito**. Brasília, DF: MS, 2011b. Disponível em: [http://svs.aids.gov.br/download/manuais/Manual\\_Instr\\_Preench\\_DO\\_2011\\_jan.pdf](http://svs.aids.gov.br/download/manuais/Manual_Instr_Preench_DO_2011_jan.pdf). Acesso em: 9 abr. 2018.

BRASIL. Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. **Síntese de evidências para**



**políticas de saúde:** mortalidade perinatal. Brasília, DF: MS, 2012a. (Série B. Textos Básicos de Saúde).

BRASIL. Ministério da Saúde. **Indicadores e Dados Básicos – IDB.** Brasília, DF: MS, 2012b. Disponível em: <http://tabnet.datasus.gov.br/cgi/idb2012/matriz.htm#mort>. Acesso em: 24 nov. 2018.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Ações Programáticas Estratégicas. **Atenção à saúde do recém-nascido:** guia para os profissionais de saúde. 2. ed. atual. Brasília, DF: MS, 2014a.

BRASIL. Ministério da Saúde. Secretaria de Atenção à Saúde. Departamento de Atenção Básica. **Diretrizes nacionais de implantação da estratégia e-SUS AB.** Brasília, DF: MS, 2014b.

BRASIL. Ministério da Saúde. Departamento de Informação e Análise Epidemiológica. **Indicadores de mortalidade que utilizam a metodologia do busca ativa,** Taxa de mortalidade infantil (menores de 1 ano), 2019. Disponível em: <http://svs.aids.gov.br/dantps/aceso-a-informacao/acoes-e-programas/busca-ativa/indicadores-de-saude/mortalidade/>. Acesso em: 1 jun. 2019.

BÜHLER, H. F. et al. Análise espacial de indicadores integrados determinantes da mortalidade por diarreia aguda em crianças menores de 1 ano em regiões geográficas. **Ciência e Saúde Coletiva**, Rio de Janeiro, v. 19, n. 10, p. 4131-4140, 2014.

COELI, C. M. et al. Sistemas de informação em saúde. *In*: MEDRONHO, R. A. (Ed.). **Epidemiologia**. 2. ed. São Paulo: Atheneu, 2009. p. 525-534.

CUNHA, A. J. L. A.; LEITE, A. J. M.; ALMEIDA, I. S. The pediatrician's role in the first thousand days of the child: the pursuit of healthy nutrition and development. **Jornal de Pediatria**, Porto Alegre, v. 91, n. 6, p. S44-S51, 2015.

CURITIBA. Secretaria Municipal da Saúde. **Diretriz da atenção à saúde da criança curitibana.** Curitiba: SMS, 2012.

CURITIBA. Secretaria Municipal da Saúde. **Plano Municipal de Saúde 2018-2021.** Curitiba: SMS, 2018a. Disponível em: <http://www.saude.curitiba.pr.gov.br/a-secretaria/plano-municipal-de-saude-e-programacao-anual-de-saude.html>. Acesso em: 13 nov. 2019.

CURITIBA. **Perfil da cidade de Curitiba.** 2018b. Disponível em: <http://www.curitiba.pr.gov.br/conteudo/perfil-da-cidade-de-curitiba/174>. Acesso em: 13 nov. 2018.

CURITIBA. Decreto n. 638, de 21 de junho de 2018. Institui o Índice de Vulnerabilidade das Áreas de Abrangência das Unidades Municipais de Saúde – IVAB na Secretaria Municipal da Saúde de Curitiba. **Diário Oficial do Município**, Curitiba, 21 de junho de 2018c.

CURITIBA. **Secretaria Municipal da Saúde**. 2019a. Disponível em: <http://www.saude.curitiba.pr.gov.br/>. Acesso em: 30 nov. 2019.

CURITIBA. Secretaria Municipal da Saúde. **Perfil dos nascidos vivos em Curitiba: série histórica de dez anos (2008 a 2017)**. Curitiba: SMS, 2019b.

CURITIBA. Secretaria Municipal da Saúde. **Rede Mãe Curitibana Vale a Vida: assistência ao pré-natal, parto e puerpério**. Curitiba: SMS, 2018c. Disponível em: [http://www.saude.curitiba.pr.gov.br/images/REDE%20M%C3%83E%20CURITIBANA%20-%20ASSIST%C3%84NCIA%20AO%20PR%C3%89\\_NATAL-PARTO-PERP%C3%89RIO%2019.pdf](http://www.saude.curitiba.pr.gov.br/images/REDE%20M%C3%83E%20CURITIBANA%20-%20ASSIST%C3%84NCIA%20AO%20PR%C3%89_NATAL-PARTO-PERP%C3%89RIO%2019.pdf). Acesso em: 13 nov. 2018.

DAHLGREN, G.; WHITEHEAD, M. **Polices and strategies to promote social equity in health**. Stockholm: Institute for Future Studies, 1991.

DEMITTO, M. O. et al. Gestaç o de alto risco e fatores associados ao  bito neonatal. **Revista da Escola de Enfermagem da USP**, S o Paulo, n. 51, p. e03208, 2017. Disponível em: [http://www.scielo.br/pdf/reeusp/v51/pt\\_1980-220X-reeusp-51-e03208.pdf](http://www.scielo.br/pdf/reeusp/v51/pt_1980-220X-reeusp-51-e03208.pdf). Acesso em: 9 set. 2019.

DIAS, B. A. S.; SANTOS NETO, E. T.; ANDRADE, M. A. C. Classifica es de evitabilidade dos  bitos infantis: diferentes m todos, diferentes repercuss es? **Cadernos de Sa de P blica**, Rio de Janeiro, n. 33, e00125916, 2017. Disponível em: <http://www.scielo.br/pdf/csp/v33n5/1678-4464-csp-33-05-e00125916.pdf>. Acesso em: 10 ago. 2019.

ELIAS, F. T. S. Avalia o de tecnologias em sa de: prop sitos e desenvolvimento no mundo e no pa s. *In*: TOMA, T. S. et al. (Org.). **Avalia o de tecnologias de sa de & pol ticas informadas por evid ncias**. S o Paulo: Instituto de Sa de, 2017. p. 15-27.

FAYYAD, U. M. et al. **Advances in knowledge discovery and data mining**. California: MIT Press, 1996.

FENTON, T. R.; KIM, J. H. A systematic review and meta-analysis to revise the Fenton growth chart for preterm infants. **BMC Pediatrics**, v. 13, n. 59, p. 1-13, 2013. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3637477/pdf/1471-2431-13-59.pdf>. Acesso em: 20 set. 2019.

FERNANDES, Fernando Timoteo; CHIAVEGATTO FILHO, Alexandre Dias Porto. Perspectivas do uso de minera o de dados e aprendizado de m quina em sa de e seguran a no trabalho. **Rev. bras. sa de ocup.**, v. 44, e13, 2019. Disponível em: <http://dx.doi.org/10.1590/2317-6369000019418>. Acesso em: 25 jan. 2020.

FRAN A, E.; LANSKY, S. Mortalidade infantil neonatal no Brasil: situa o, tend ncias e perspectivas. *In*: ENCONTRO NACIONAL DE ESTUDOS POPULACIONAIS, 16., 2016. **Anais [...]**. [S.l.: s.n.], 2016. p. 1-29.

GARC A, M. A. et al. Detecci n ecogr fica de anomal as cong nitas en 76.155 nacimientos en Bogot  y Cali, 2011-2012. **Biom dica**, v.34, p. 379-386, 2014.

Disponível em: <http://dx.doi.org/10.7705/biomedica.v34i3.2259>. Acesso em: 3 dez. 2019.

GOLDSCHMIDT, R.; PASSOS, E.; BEZERRA, E. **Data mining**: conceitos, técnicas, algoritmos, orientações e aplicações. 2. ed. Rio de Janeiro: Elsevier, 2015.

GROEN, H. et al. Stillbirth and neonatal mortality in pregnancies complicated by major congenital anomalies: findings from a large European cohort. **Prenatal Diagnosis**, n. 37, p. 1100-1111, 2017. Disponível em: <https://obgyn.onlinelibrary.wiley.com/action/showCitFormats?doi=10.1002%2Fpd.5148>. Acesso em: 3 dez. 2019.

HALL, M. et al. **Weka**: the Weka data mining software. 2009. Disponível em: <http://www.cs.waikato.ac.nz/ml/weka>. Acesso em: 28 ago. 2018.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Cidades e estados do Brasil**. 2019. Disponível em: <https://www.ibge.gov.br/cidades-e-estados/pr/curitiba.html>. Acesso em: 30 nov. 2019.

JEGANATHAN, R. et al. Factors associated with recovery from 1 minute Apgar score <4 in live, singleton, term births: an analysis of Malaysian National Obstetrics Registry data 2010-2012. **BMC Pregnancy and Childbirth**, v. 17, n. 110, 2017. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5385027/>. Acesso em: 30 ago. 2019.

KOHAVI, R.; QUINLAN, R. **C5. 1.3 Decision Tree Discovery**. 1999. Disponível em: <http://ai.stanford.edu/~ronnyk/treesHB.pdf>. Acesso em: 31 jan. 2019.

KOSHIDA, S. et al. Possible prevention of neonatal death: a regional population-based study in Japan. **Yonsei Medical Journal**, v. 57, n. 2, mar. 2016. Disponível em: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4740536/>. Acesso em: 28 dez. 2019.

KROPIWIEC, M. V.; FRANCO, S. C.; AMARAL, A. R. Fatores associados à mortalidade infantil em município com índice de desenvolvimento humano elevado. **Revista Paulista de Pediatria**, São Paulo, v. 35, n. 4, p. 391-398, dez. 2017. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0103-05822017000400391&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-05822017000400391&lng=en&nrm=iso). Acesso em: 11 dez. 2019.

LANSKY, S. et al. Pesquisa Nascer no Brasil: perfil da mortalidade neonatal e avaliação da assistência à gestante e ao recém-nascido. **Caderno de Saúde Pública**, Rio de Janeiro, v. 30, supl. 1, p. S192-S207, 2014. Disponível em: 10.1590/0102-311X00133213. Acesso em: 6 dez. 2019.

LEAL, M. C. et al. Prenatal care in the Brazilian public health services. **Rev Saude Publica**, v. 54, n. 8, p. 1-12, 2020. Disponível em: <https://doi.org/10.11606/s1518-8787.2020054001458>. Acesso em 25 jan. 2020

MAIA, L. T. S. et al. Uso do linkage para a melhoria da completude do SIM e do SINASC nas capitais brasileiras. **Revista de Saúde Pública**, São Paulo, v. 51, n.

112, 2017. Disponível em:  
[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0034-89102017000100300&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0034-89102017000100300&lng=en&nrm=iso). Acesso em: 20 dez. 2019.

MAMELI, C.; MAZZANTINI, S.; ZUCCOTTI, G. Nutrition in the first 1000 days: the origin of childhood obesity. **International Journal of Environmental Research and Public Health**, v. 13, n. 9, p. 838, 2016.

MARIN, H. F. Sistemas de informação em saúde: considerações gerais. **Journal of Health Informatics**, v. 2, n. 1, p. 20-24, 2010. Disponível em: <http://www.jhi-sbis.saude.ws/ojs-jhi/index.php/jhi-sbis/article/view/4/52>. Acesso em: 10 nov. 2018.

MARMOT, M. Social determinants of health inequalities. **The Lancet**, n. 365, p. 1099-1104, 2005. Disponível em:  
[https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(05\)71146-6/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(05)71146-6/fulltext). Acesso em: 20 dez. 2019.

MILANI, C. S.; CARVALHO, D. R. Pós-processamento em KDD. **Revista de Engenharia e Tecnologia**, Ponta Grossa, v. 5, n. 1, p. 151-162, abr. 2013. Disponível em:  
<http://revistas2.uepg.br/index.php/ret/article/view/11473/209209209447>. Acesso em: 15 nov. 2018.

MOORE, T. et al. **The first thousand days: an evidence paper**. Melbourne: Murdoch Children's Research Institute, 2017. Disponível em:  
<https://www.suicideinfo.ca/resource/the-first-thousand-days-an-evidence-paper/>. Acesso em: 31 ago. 2019.

ORGANIZAÇÃO MUNDIAL DA SAÚDE (OMS); COMISSÃO PARA OS DETERMINANTES SOCIAIS DA SAÚDE (CDSS). **Redução das desigualdades no período de uma geração**. Igualdade na saúde através da ação sobre os seus determinantes sociais. Relatório Final da Comissão para os Determinantes Sociais da Saúde. Lisboa: OMS, 2010.

ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE (OPAS). Rede Interagencial de Informação para a Saúde. **Indicadores básicos para a saúde no Brasil: conceitos e aplicações**. 2. ed. Brasília, DF, 2008.

PAPAGEORGHIU, A. T. et al. International standards for fetal growth based on serial ultrasound measurements: the fetal growth longitudinal study of the Intergrowth-21st project. **The Lancet**, v. 384, n. 9946, p. 869-879, 2014. Disponível em: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(14\)61490-2/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(14)61490-2/fulltext). Acesso em: 20 set. 2019.

PINHEIRO, R. S.; TORRES, T. Z. G. Análise exploratória de dados. *In*: MEDRONHO, R. A. (Ed.). **Epidemiologia**. 2. ed. São Paulo: Atheneu, 2009. p. 323-341.

PROGRAMA DAS NAÇÕES UNIDAS PARA O DESENVOLVIMENTO (PNUD). **Atlas do desenvolvimento humano no Brasil**. 2020. Disponível em: <http://www.atlasbrasil.org.br/2013/pt/download/>. Acesso em: 10 mar. 2020.

QUINLAN, J. R. **C4.5: programs for machine learning**. San Mateo: Morgan Kaufmann, 1993.

REZENDE, S. O. et al. Mineração de dados. *In*: REZENDE, S. O. (Org.). **Sistemas inteligentes: fundamentos e aplicações**. Barueri: Manole, 2003. p. 307-335.

SADOVSKY, A. D. I. et al. Iniquidades socioeconômicas em nascimentos prematuros em quatro estudos brasileiros de coortes de nascimento. **Jornal de Pediatria**, Porto Alegre, v. 94, n. 1, p. 15-22, 2018.

SARTORELLI, A. P. et al. Fatores que contribuem para mortalidade infantil utilizando a mineração de dados. **Saúde e Pesquisa**, Maringá, v. 10, n. 1, p. 33-41, jan./abr. 2017.

SAVE THE CHILDREN. **Nutrition in the first 1,000 days: state of the world's mothers 2012**. 2012. Disponível em: <https://www.savethechildren.org/content/dam/usa/reports/advocacy/sowm/sowm-2012.pdf>. Acesso em: 15 nov. 2018.

SILVA, E. S. A.; PAES, N. A. Programa Bolsa Família e a redução da mortalidade infantil nos municípios do Semiárido brasileiro. **Ciência e Saúde Coletiva**, Rio de Janeiro, v. 24, n. 2, p. 623-630, fev. 2019. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1413-81232019000200623&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1413-81232019000200623&lng=en&nrm=iso). Acesso em: 20 dez. 2019.

SILVA, C. G.; LUIZ, R. R.; BAHIA, L. Vinculação aos planos de saúde de menor abrangência e seu impacto na utilização da rede de serviços do SUS: uma análise a partir da Pesquisa Nacional de Saúde de 2013. **JBES: Brazilian Journal of Health Economics/Jornal Brasileiro de Economia da Saúde**, v. 11, n. 2, 2019. Disponível em: [http://docs.bvsalud.org/biblioref/2019/09/1021110/05\\_ao\\_artigo-vinculacao-aos-planos.pdf](http://docs.bvsalud.org/biblioref/2019/09/1021110/05_ao_artigo-vinculacao-aos-planos.pdf). Acesso em: 01 abr. 2020.

SOARES, D. C. et al. Programas de vigilancia de recién nacidos de riesgo: experiencias municipales brasileñas. **Cultura de los Cuidados**, v. 23, n. 53, p. 316-329, 2019. Disponível em: <http://hdl.handle.net/10045/91849>. Acesso em: 2 jan. 2020.

TEIXEIRA, G. A. et al. Risk factors for neonatal mortality in the life of first week. **Revista de Pesquisa: Cuidado é Fundamental Online**, Rio de Janeiro, v. 8, n. 1, p. 4036-4046, jan. 2016. Disponível em: <http://www.seer.unirio.br/index.php/cuidadofundamental/article/view/3943>. Acesso em: 11 dez. 2019.

TEIXEIRA, R. S.; COLMANETTI, J. B. D.; CARVALHO, D. R. Post-processing of classifiers – KDD. **Iberoamerican Journal of Applied Computing**, v. 5, n. 1, p. 1-6, abr. 2015.

TOMASI, E. et al. Qualidade da atenção pré-natal na rede básica de saúde do Brasil: indicadores e desigualdades sociais. **Cad. Saúde Pública**, Rio de Janeiro, v. 33, n. 3, e00195815, 2017. Disponível em:

[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0102-311X2017000305001&lng=pt&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-311X2017000305001&lng=pt&nrm=iso). Acesso em: 25 jan.2020.

UNITED NATIONS CHILDREN´S FUND (UNICEF). **Levels & trends in child mortality**. New York, 2017. Disponível em:

[https://www.unicef.org/publications/files/Child\\_Mortality\\_Report\\_2017.pdf](https://www.unicef.org/publications/files/Child_Mortality_Report_2017.pdf). Acesso em: 18 jul. 2018.

UNITED NATIONS CHILDREN´S FUND (UNICEF). **Every child alive: the urgent need to end newborn deaths**. 2018. Disponível em:

<https://weshare.unicef.org/Package/2AMZIF2P99KK>. Acesso em: 17 jul. 2018.

UNITED NATIONS CHILDREN´S FUND (UNICEF); WORLD HEALTH ORGANIZATION (WHO). **Low birthweight: country, regional and global estimates**. New York, 2004. Disponível em:

<http://apps.who.int/iris/bitstream/10665/43184/1/9280638327.pdf>. Acesso em: 6 jun. 2019.

VARELA, A. R. et al. Fetal, neonatal, and post-neonatal mortality in the 2015 Pelotas (Brazil) birth cohort and associated factors. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 35, n. 7, p. e00072918, 2019. Disponível em:

<https://www.scielosp.org/pdf/csp/2019.v35n7/e00072918/en>. Acesso em: 31 ago. 2019.

VELOSO, F. C. S. et al. Análise dos fatores de risco na mortalidade neonatal no Brasil: uma revisão sistemática e metanálise de estudos observacionais. **Jornal de Pediatria**, Porto Alegre, v. 95, n. 5, p. 519-530, set. 2019. Disponível em:

[http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0021-75572019000600519&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0021-75572019000600519&lng=en&nrm=iso). Acesso em: 6 dez. 2019.

VICTORA, C. 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**, v. 3, n. 4, p. e199-e205, 2015.

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. **The Lancet**, v. 384, n. 9946, p. 857-868, 2014. Disponível em: [https://www.thelancet.com/journals/lancet/article/PIIS0140-6736\(14\)60932-6/fulltext](https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(14)60932-6/fulltext). Acesso em: 19 set. 2019.

WITTEN, I. H.; FRANK, E. **Data mining practical machine learning tools and techniques**. 2. ed. San Francisco: Elsevier, 2005.

WORLD HEALTH ORGANIZATION (WHO). **Declaração política do Rio sobre determinantes sociais da saúde**. Rio de Janeiro, 2011. Disponível em:

[http://www.who.int/sdhconference/declaration/Rio\\_political\\_declaration\\_portuguese.pdf](http://www.who.int/sdhconference/declaration/Rio_political_declaration_portuguese.pdf). Acesso em: 2 dez. 2019.

WORLD HEALTH ORGANIZATION (WHO). **Birth defects surveillance training: facilitator's guide**. Geneva, 2015. Disponível em: [https://www.who.int/nutrition/publications/birthdefects\\_training\\_facilitators\\_guide/en/](https://www.who.int/nutrition/publications/birthdefects_training_facilitators_guide/en/). Acesso em: 5 jun. 2019.

## APÊNDICE A – ARTIGO DE REVISÃO INTEGRATIVA SOBRE VARIÁVEIS ASSOCIADAS À MORTALIDADE MATERNO-INFANTIL

VARIÁVEIS ASSOCIADAS À MORTALIDADE MATERNO E INFANTIL: UMA  
REVISÃO INTEGRATIVA

VARIABLES ASSOCIATED WITH MATERNAL AND CHILD MORTALITY: AN  
INTEGRATIVE REVIEW

Determinantes Sociais da Saúde e Mortalidade materno-infantil  
*Social Determinants of Health and maternal-child mortality*

Descritores: Mortalidade Infantil. Mortalidade Materna. Determinantes Sociais da Saúde.  
*Keywords: Infant Mortality. Maternal Mortality. Social Determinants of Health.*

### Resumo

**Introdução:** A mortalidade materna e infantil reflete a situação sanitária e a qualidade da assistência. As taxas de mortalidade no Brasil apresentam diferenças internas e são superiores à de outros países em desenvolvimento. **Objetivo:** Identificar variáveis associadas à mortalidade materna e infantil. **Materiais e Métodos:** Revisão integrativa, consultadas as bases Literatura Latino-Americana e do Caribe em Ciências da Saúde, Medical Literature Analysis and Retrieval System Online e Banco de Dados em Enfermagem. Considerados artigos publicados entre 1999-2018, cenário Brasil e idioma português, com seleção por dois pesquisadores calibrados com nível de concordância Kappa substancial. As variáveis associadas significativamente às mortalidades foram aproximadas às camadas do modelo dos Determinantes Sociais de Saúde. **Resultados:** Identificou-se 522 artigos e selecionados 74 com concordância Kappa = 0,793. Na camada de características individuais destacaram-se peso ao nascer e idade materna; na camada do estilo de vida, tabagismo na gestação; em relação às influências sociais predominou estado civil materno; nas condições de vida e trabalho, a assistência pré-natal e escolaridade materna; e nas condições socioeconômicas e ambientais gerais evidenciou-se saneamento básico e renda. **Conclusões:** As variáveis mais encontradas foram peso ao nascer e idade materna, relacionadas à camada individual, sem governabilidade direta do gestor. Para mudanças nessa camada, as intervenções devem ser realizadas a partir das camadas mais externas, nas quais destacaram-se assistência pré-natal, escolaridade materna, renda e saneamento básico. Reforça-se a responsabilidade direta dos gestores na redução da mortalidade materna e infantil, por meio de ações que visem reduzir as iniquidades em saúde.

### Abstract

**Introduction:** Maternal and child mortality reflect the local health situation and the quality health care. Brazilian mortality rates differ internally and are higher than others developing countries. **Objectives:** To identify variables associated with maternal and child mortality.



**Materials and Methods:** Integrative review, researching Latin American and Caribbean Health Science Literature, Medical Analysis and Retrieval System Online and Nursing Database, considering only articles published between 1999 and 2018, Brazil scenario and Portuguese language. Two researchers selected the articles after performing calibrations until they reached a substantial agreement level, calculated by Kappa. Variables significantly associated with mortality were extracted from articles and approximated to the layers of the Social Determinants of Health. **Results:** Five hundred and twenty-two articles were identified after Kappa = 0,793. Seventy-four articles were selected. In individual characteristics layer, birth weight and maternal age were highlighted; in lifestyle layer, smoking in pregnancy; about social influences, maternal marital status predominated; in living and working conditions, prenatal care and maternal education; and in general socioeconomic and environmental conditions, sanitation and rent were evidenced. **Conclusions:** The most noticed variables were birth weight and maternal age; both related to individual layer, therefore, with lack of directly governability by the managers. In order to change this tier, interventions shall be performed in outer layers, highlighting: prenatal care, maternal education, income, and basic sanitation. Results strength managers direct responsibility to reduce maternal and child mortality throughout actions aimed at reducing health inequities.

## INTRODUÇÃO

A Taxa de Mortalidade Infantil (TMI) no Brasil em 2017 foi de 13,4/1.000 nascidos vivos, seis vezes mais alta que a do Japão (2/1.000 nascidos vivos), uma das menores do mundo e mais alta que a de outros países em desenvolvimento, como a Argentina (9/1.000 nascidos vivos)<sup>1</sup>. A diferença entre os países também pode ser identificada dentro do Brasil, onde as regiões Norte e Nordeste obtiveram as maiores TMI - 17,3 e 15,8/1.000 nascidos vivos respectivamente, enquanto as regiões Sudeste e Sul apresentaram as menores, respectivamente 11,7/1.000 e 10,1/1.000 nascidos vivos.

A meta estabelecida frente às Nações Unidas era a redução de 66% na mortalidade materna no mundo entre 1990 e 2015, porém o resultado foi de 43%. Cerca de 99% dos óbitos maternos ocorrem em países em desenvolvimento, cuja RMM é de 239/100.000 nascidos vivos, elevada quando comparada a de países desenvolvidos (12/100.000 nascidos vivos)<sup>2</sup>. A Razão de Mortalidade Materna (RMM) no Brasil aumentou de 62 para 64/100.000 nascidos vivos entre 2015 e 2017, principalmente nas regiões Norte e Nordeste.

A TMI é definida pelo número de crianças que morrem antes de completar um ano de idade para cada 1.000 nascidos vivos<sup>3</sup>. Esta taxa apresenta três componentes baseados na idade quando da ocorrência do óbito. O primeiro, mortalidade neonatal precoce diz respeito aos óbitos ocorridos de 0 até 6 dias completos; o segundo, mortalidade neonatal tardia, abrange os óbitos ocorridos de 7 a 27 dias completos e o terceiro componente é a mortalidade pós-neonatal, que contempla os óbitos de 28 a 364 dias de vida completos<sup>4</sup>.

A morte materna é definida pela morte de mulheres devido a qualquer causa relacionada com ou agravada pela gravidez, excluindo-se causas acidentais, durante a gestação ou dentro de um período de 42 dias após a gravidez<sup>5</sup>. Pode ser direta quando ocorre por complicações obstétricas durante a gravidez, parto ou puerpério por omissões, intervenções ou tratamento incorreto, ou indireta quando por doenças prévias à gestação ou agravadas por esta ou desenvolvidas neste período, porém sem causa gestacional<sup>6</sup>. A RMM se dá pelo número de óbitos maternos por 100 mil nascidos vivos<sup>4</sup>.

A TMI e a RMM são consideradas indicadores de saúde utilizados para analisar variações populacionais, temporais e geográficas, contribuir na avaliação dos níveis de saúde e desenvolvimento socioeconômico e servir de base para planejamento de políticas voltadas

para a saúde, como atenção ao pré-natal, parto e puerpério<sup>4</sup>. Por meio delas é possível identificar situações de desigualdades, salientando a lacuna entre ricos e pobres e reforçando a existência de situações de iniquidade em saúde<sup>2,4</sup>.

O modelo dos Determinantes Sociais da Saúde (DSS) proposto por Dahlgren e Whitehead<sup>7</sup> (1991) é composto por cinco camadas concêntricas que demonstram as principais influências na saúde e a relação entre elas, onde as mais externas influenciam e modulam as mais internas. Na mais externa estão as condições socioeconômicas, culturais e ambientais gerais que ao sofrerem interferências estruturais podem reduzir iniquidades em saúde, demonstrando sua influência sobre as outras. A próxima camada retrata o modo e as condições nas quais os indivíduos vivem e trabalham, demonstrando que pessoas em desvantagem social são mais vulneráveis aos riscos em saúde. A terceira camada destaca a influência do apoio social de familiares, amigos ou pessoas próximas na saúde do indivíduo. A seguir, está a camada que retrata o comportamento e estilo de vida individual, na qual as escolhas pessoais são importantes, mas que claramente sofrem influências das mais externas, visto que essas escolhas dependem também de acesso à informação, pressão de amigos e familiares, acesso a alimentos e ambientes saudáveis, entre outros. Na camada mais interna estão características mais individuais como idade, sexo e condições genéticas, que possuem pouca possibilidade de mudança e são influenciadas pelas camadas mais externas.

Considerando a importância das TMI e RMM, os resultados obtidos nacionalmente e a interação entre todas as camadas dos DSS, esta revisão integrativa, tem por objetivo identificar variáveis associadas à mortalidade materna e infantil que possam direcionar os gestores no estabelecimento de ações no enfrentamento dessa mortalidade.

## MÉTODOS

Esta revisão, adaptada do método Preferred Reporting Items for Systematic reviews and Meta-Analysis (PRISMA)<sup>8</sup>, envolve as etapas de identificação, seleção, elegibilidade e inclusão. Para a identificação dos artigos, realizou-se a busca em agosto de 2018 nas bases de dados Literatura Latino-Americana e do Caribe em Ciência da Saúde (LILACS), Medical Literature Analysis and Retrieval System Online (MEDLINE) e Banco de Dados em Enfermagem (BDENF). Foram utilizados os descritores: “mortalidade infantil” AND “fatores de risco” OR “fatores de proteção” OR “fatores socioeconômicos” e “mortalidade materna” AND “fatores de risco” OR “fatores de proteção” OR “fatores socioeconômicos”. Os critérios utilizados para a busca foram: publicações somente do tipo artigo, publicados entre 1999 e 2018, com disponibilidade de texto completo, estudos feitos no Brasil e idioma português.

Após a exclusão dos artigos duplicados, a etapa de seleção foi realizada por duas pesquisadoras, de forma independente, pela leitura do título e resumo, aplicando-se o seguinte critério de seleção: qualquer situação relacionada à mortalidade materna e/ou mortalidade infantil e/ou mortalidade até dois anos de idade. Foram realizadas calibrações até atingir-se um nível de concordância substancial<sup>9</sup>. No caso de discordância, o artigo era analisado por uma terceira pesquisadora.

Os artigos selecionados foram avaliados pela leitura na íntegra. O critério de inclusão foi o uso de teste estatístico para verificação da significância da associação encontrada entre variável e desfecho estudado. Dos artigos incluídos, foram extraídas as seguintes características: título e autor para identificação do artigo, problema de pesquisa para apontar a mortalidade estudada, local do estudo para entendimento do contexto, base de dados para reconhecimento de possíveis vieses e identificação das bases mais utilizadas, variáveis com associação significativa à mortalidade, sendo aproximadas às camadas dos DSS e síntese dos resultados.

O local do estudo foi agrupado conforme unidades federativas (UF) ou Brasil, quando estudo realizado na Federação como um todo. Agrupamentos também foram realizados em relação às fontes de dados<sup>10</sup>: secretarias municipais e estaduais de saúde foram agrupadas como secretarias de saúde; cartórios, hospitais e unidades básicas de saúde foram agrupadas como registros oficiais; cemitérios, igrejas, delegacias, funerárias, farmácias, líderes comunitários e curandeiras como registros não oficiais.

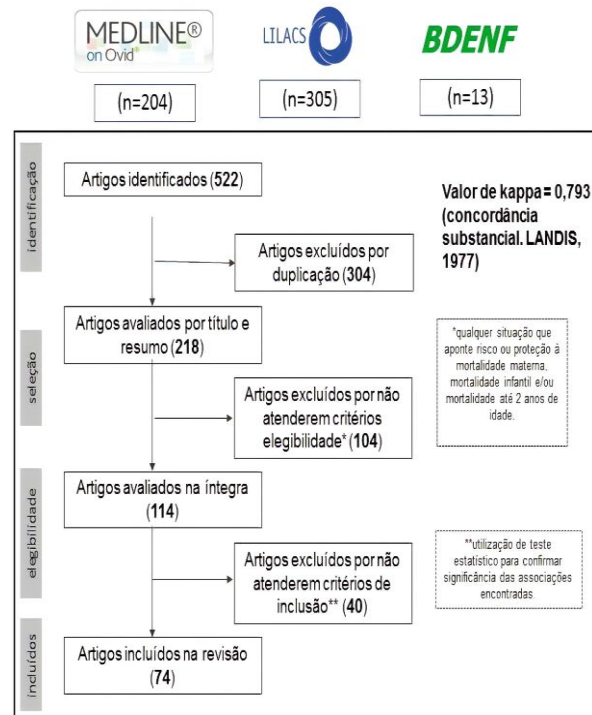
As variáveis encontradas foram aproximadas das camadas dos DSS7, nominados da mais interna para a mais externa: determinantes individuais, determinantes comportamentais, determinantes de influência social, determinantes de condições de vida e trabalho e determinantes socioeconômico e ambiental geral.

Algumas variáveis foram agrupadas a fim de facilitar a análise dos dados: grupo afecções do recém-nascido - oligodrâmnio, hiperglicemia no recém-nascido, hipoxemia no recém-nascido, sepse neonatal e síndrome do desconforto respiratório; grupo local do nascimento - tipo de hospital, nascimento em pronto-socorro e complexidade do hospital de nascimento; grupo morbidade materna - distúrbio psiquiátrico materno, endometrite, hemorragia vaginal, agressão física em gestantes, pré-eclâmpsia e hipertensão arterial materna; grupo intercorrências da gestação - descolamento prematuro de placenta, sofrimento fetal e infecções urinárias e grupo renda - PIB per capita, razão entre a renda média dos 10% mais ricos e a dos 40% mais pobres, concentração de renda, renda per capita e número de indigentes.

## **RESULTADOS E DISCUSSÃO**

Foram encontrados nas bases de dados MEDLINE, LILACS e BDENF, 204, 305 e 13 artigos respectivamente, totalizando 522. Desses, 304 foram excluídos por estarem duplicados e 218 avaliados por título e resumo. O valor de Kappa obtido foi de 0,793, classificado como concordância substancial<sup>9</sup>. Após seleção discordante, 23 artigos foram avaliados por uma terceira pesquisadora. Dos 114 artigos eleitos para avaliação por leitura na íntegra, 74 foram incluídos na revisão (Figura 1).

Figura 1 - Resultados de cada etapa da revisão adaptado do PRISMA



Fonte: Adaptado de Moher et al., 2009

Em relação ao problema de pesquisa, 69 artigos abordaram a mortalidade infantil e desses, quase metade versavam sobre mortalidade neonatal (31). Devido a instituição de políticas globais voltadas para a diminuição da mortalidade infantil, houve um decréscimo do componente pós-neonatal<sup>11</sup>, que poderia explicar o menor número, nesta revisão, de estudos com foco nesse período. A redução desse componente colaborou para o aumento do componente neonatal no Brasil a partir da década de 90<sup>11</sup>, o que reforça a importância de estudos desse período uma vez que é o principal componente da mortalidade infantil<sup>12</sup>.

A mortalidade materna foi discutida em seis artigos, um deles associado à mortalidade infantil. A baixa quantidade de estudos prejudica o aprofundamento na investigação para fatores de risco do referido desfecho<sup>13</sup>. A maioria das mortes maternas são evitáveis e do tipo obstétricas, resultantes de déficits na atenção à saúde da gestante desde a gestação até o puerpério<sup>14</sup>. Uma das metas preconizadas pelos Objetivos de Desenvolvimento do Milênio (ODM) era a redução da mortalidade infantil em 75% até 2015, a qual não foi alcançada no Brasil. A nova meta brasileira diante dos Objetivos de Desenvolvimento Sustentável (ODS) é de 30 mortes maternas a cada 100.000 nascidos vivos até 2030, o que contrasta com uma RMM de 64 em 2017<sup>15</sup>. A mortalidade até dois anos foi contemplada em dois artigos, ambos também tratavam de mortalidade infantil.

A camada dos determinantes individuais, com 25 variáveis aproximadas, foi a segunda camada mais associada à mortalidade (50 artigos) (Tabela 1). Apesar de serem consideradas com pouca ou nenhuma possibilidade de mudança, as características individuais são úteis para a identificação de grupos mais vulneráveis, propiciando o direcionamento das ações<sup>7</sup>. Dessa camada, a variável idade materna foi uma das que mais apareceram associada à mortalidade. Gravidez na adolescência foi encontrada como uma variável de risco<sup>16-20</sup> ao passo que idade materna maior ou igual a 35 anos foi considerada risco em outros estudos<sup>21,22</sup> e sem significância estatística em outro<sup>23</sup>. A influência entre as camadas do modelo dos DSS é

exemplificada pela correlação da idade jovem da mãe às condições de vida, principalmente baixa inserção social, com impactos no comportamento reprodutivo e morbimortalidade das crianças<sup>24</sup>.

Tabela 1 – Variáveis aproximadas à camada dos determinantes individuais e respectiva quantidade de artigos\*

variáveis	<i>n</i>
Peso ao nascer	28
Prematuridade	28
Idade materna	23
Tipo de parto	13
Morbidade materna	12
Anormalidades congênitas	11
Tipo de gestação	7
Índice de Apgar no 1º e 5º minutos	7
Cor materna	6
Afecções do recém-nascido	5
Intercorrências na gestação	5
Natimortalidade	4
Infecção pelo vírus da dengue na gestação	2
História de óbito fetal	1
Número de gestações anteriores	1
Número de filhos vivos	1
Idade dos recém-nascidos internados em UTI	1
Estatura para idade	1
Peso para estatura	1
História de aborto provocado	1
Doenças infecciosas e parasitárias associadas à criança	1
Apresentação pélvica	1
Intercorrências no parto	1
Nascimento pós-termo	1
Índice de Apgar no 1º minuto	1

\*nota: um mesmo artigo pode ter apresentado mais de uma variável

Ainda associado às características individuais da mãe, a infecção materna pelo vírus da dengue durante a gestação foi considerada fator de risco tanto para óbitos maternos, neonatais e infantil<sup>25</sup> como apenas para óbito materno<sup>26</sup>. A infecção durante a gestação triplica o risco de óbito materno, e em casos de dengue grave o risco aumenta cerca de 27 vezes<sup>27</sup>. Desse modo, verifica-se a influência da camada de condições de vida e trabalho do modelo dos DSS na individual à medida que os gestores propõem intervenções em relação à condição que os indivíduos vivem<sup>7</sup>, como a ampliação de ações coletivas de prevenção no combate à dengue e outras arboviroses, especialmente na população gestante a qual é considerada grupo vulnerável para risco de agravamento e morte por dengue<sup>28</sup>.

O baixo peso ao nascer foi encontrado como fator de risco<sup>21,29-36</sup> enquanto outro estudo encontrou associação com a mortalidade somente o muito baixo peso<sup>37</sup> ou o extremo baixo peso ao nascer<sup>16,38</sup>. O baixo peso<sup>16,39-40</sup> e o muito baixo peso ao nascer<sup>41</sup> são considerados uns dos principais fatores associados à mortalidade neonatal. Em contrapartida, alguns estudos não encontraram significância estatística na associação do peso ao nascer à mortalidade<sup>10,42-43</sup>. A assistência ao pré-natal de baixo risco monitora na gestante possíveis causas de diminuição

do peso ao nascer, como violência, anemias, Doenças Sexualmente Transmissíveis (DST's), Infecções do Trato Urinário (ITU's) e comorbidades prévias como asma<sup>44</sup>. Dessa maneira, percebe-se, novamente, a influência das camadas mais externas do modelo dos DSS nas camadas mais internas<sup>7</sup>, onde a assistência ao pré-natal (camada dos determinantes de condições de vida e trabalho) reflete no peso ao nascer (camada dos determinantes individuais).

Outra influência de camada mais externa na camada individual pode ser exemplificada pela baixa qualidade do atendimento ao neonato na sala de parto pode resultar em um baixo valor de Apgar<sup>19,45</sup>. O índice de Apgar se refere à avaliação clínica do recém-nascido nos 1º e 5º minutos de vida<sup>46</sup>, e quando menor que 7, no 1º e 5º minutos, pode indicar asfixia e por isso ser predictor de risco de morte<sup>39</sup>. Encontrou-se como fator de risco quando Apgar menor do que 7 no 1º minuto<sup>24,47</sup> e 5º minuto<sup>24,29-30,40,47-49</sup>, porém um dos estudos considerou risco quando menor ou igual a 3 no 5º minuto<sup>50</sup>.

À camada dos determinantes comportamentais foi aproximada somente a variável tabagismo na gestação (Tabela 2), a qual apresentou-se como variável de risco associada à mortalidade infantil<sup>51</sup>. Intervenções nessa camada requerem, dentre outras, ações de educação em saúde no intuito de estimular a mudança de comportamento<sup>7</sup>. As campanhas antitabagismo são amplamente divulgadas na mídia e a televisão parece ser a mídia mais adequada para o fornecimento de materiais educativos a gestantes. Apesar das campanhas, permanece baixo o nível de conhecimento das gestantes sobre as consequências do uso do tabaco<sup>52</sup>.

Tabela 2 – Variáveis aproximadas à camada dos determinantes comportamentais e respectiva quantidade de artigos\*

variáveis	<i>n</i>
Tabagismo na gestação	1

\*nota: um mesmo artigo pode ter apresentado mais de uma variável

A camada dos determinantes de influência social totalizou três variáveis (Tabela 3). O suporte social no período de gestação reduz comportamentos de risco e a chance de as crianças não terem cuidados adequados no futuro<sup>53-54</sup>. Desse modo, é importante que os gestores considerem ações que favoreçam a criação e fortalecimento de redes de apoio e autocuidado<sup>7</sup>. O fato de a mulher não ter companheiro durante a gestação é um fator de risco para mortalidade infantil<sup>55-58</sup>, assim como o apoio paterno é fator de proteção<sup>59</sup>.

Tabela 3 – Variáveis aproximadas à camada dos determinantes de influência social e respectiva quantidade de artigos\*

variáveis	<i>n</i>
Estado civil materno	6
Apoio paterno na gestação	2
Coabitação dos pais	1

\*nota: um mesmo artigo pode ter apresentado mais de uma variável

A camada dos determinantes de condições de vida e trabalho foi a que mais apareceu associada à mortalidade (55 artigos), bem como foi a que apresentou a maior diversidade de variáveis (34) (Tabela 4). Dentre elas destaca-se a escolaridade materna, a qual associou-se à mortalidade infantil e materna. Um tempo de escolaridade menor que 4 anos é considerado

risco para mortalidade infantil<sup>21,60</sup>, assim como ausência de instrução materna<sup>61</sup>, enquanto a partir dos oito anos de escolaridade as RMM se aproximaram do preconizado pela OMS<sup>62</sup>.

Tabela 4 – Variáveis aproximadas à camada dos determinantes de condições de vida e trabalho e respectiva quantidade de artigos\*

variáveis	<i>n</i>
Assistência pré-natal	20
Escolaridade materna	18
Local do nascimento	11
Residência	7
Escolaridade	6
Uso de corticoide antenatal	4
Tipo de trabalho materno	3
Condição socioeconômica materna	2
Número de cômodos no domicílio	2
Ventilação mecânica no recém-nascido	2
Utilização de partograma	2
Leitos de UTI neonatal	2
Médicos por mil habitantes	2
Cobertura do PACS/PSF	2
Leitos obstétricos	2
Número de eletrodomésticos no domicílio	1
Tipo de moradia	1
Meio de transporte utilizado para assistência ao parto	1
Vínculo à maternidade durante o pré-natal	1
Trabalho familiar na agricultura	1
Consultas de puericultura	1
Período de internação do recém-nascido em UTI	1
Tempo de internação da gestante para o parto	1
Nutrição parenteral no recém-nascido	1
Uso de surfactante	1
Avaliações fetais durante o trabalho de parto	1
Escore de CRIB	1
Critérios laboratoriais de <i>near miss</i> materno	1
Presença ou ausência de salas de parto	1
Cobertura vacinal	1
Local do óbito infantil	1
Número de centros e postos de saúde	1
Uso de tubo endotraqueal no recém-nascido	1
Internação em berçário	1

\*nota: um mesmo artigo pode ter apresentado mais de uma variável

O acesso a serviços de saúde e assistência ao pré-natal pode ser inferida pelo número de consultas de pré-natal - relacionadas à mortalidade infantil quando o número de consultas inferior a seis<sup>63-64</sup> ou quatro para crianças que nasceram em domicílio<sup>65</sup>, e na sua ausência<sup>12,16,66-68</sup>. O acompanhamento pré-natal de qualidade permite identificar e intervir precocemente em potenciais situações de risco que possam causar danos à saúde materno-infantil<sup>69</sup>, assim como vem se mostrando importante para a prevenção de morbimortalidade infantil<sup>45,70-71</sup> e diminuição de mortalidade de recém-nascidos prematuros<sup>72</sup>.

Ainda relacionada à atenção à saúde materno-infantil, a criação da Rede Cegonha ocorreu em um contexto onde, apesar da ampliação na cobertura do acompanhamento pré-natal, os indicadores de mortalidade não estavam adequados e importantes causas de morbimortalidade perinatal e materna como sífilis congênita e hipertensão arterial mantinham suas taxas de incidência elevadas<sup>44</sup>. De 2015 para 2016 o número de casos de sífilis em gestantes foi de 33.365 para 37.436, e os casos de sífilis congênita, de 19.288 para 20.474<sup>73-74</sup>. A sífilis congênita é resultado do manejo inadequado de sífilis em gestantes e acarreta alta morbimortalidade para os conceitos, sendo considerada um problema de saúde pública<sup>75</sup>. Somado a isso, o não atendimento de padrões estabelecidos em protocolos nacionais e o início tardio de pré-natal são outros indícios de que apesar da alta cobertura da Atenção Pré-Natal no âmbito da Atenção Primária à Saúde, a mesma não tem sido de qualidade<sup>76-77</sup>.

Também na camada de condições de vida e trabalho, o tempo prolongado para se chegar ao local do parto se mostrou como fator de risco para mortalidade infantil<sup>33,40,78-80</sup>. Meios móveis e acessos rodoviários adequados são fatores que diminuem o tempo de acesso ao local do parto e assim contribuem para redução de óbitos infantis por causas evitáveis<sup>81</sup>.

Por fim, à camada dos determinantes gerais foram aproximadas 17 variáveis, presentes em 22 artigos (Tabela 5). A camada mais externa representa grande influência sobre todas as outras e as intervenções nela realizadas são de governabilidade direta dos gestores, pois envolve o estabelecimento de políticas públicas além da saúde, como aquelas voltadas para redução da pobreza<sup>7</sup>. Em regiões mais afastadas de capitais e em áreas consideradas periféricas, principalmente naquelas com condições socioeconômicas desfavoráveis, há aumento do risco de mortalidade neonatal<sup>82-84</sup>. O nascimento em regiões com altas taxas de pobreza aumenta a chance de morte neonatal<sup>85</sup>. Em uma cidade da região Nordeste, onde a taxa de mortalidade infantil é considerada alta (15,8/1.000 nascidos vivos)<sup>86</sup>, a mortalidade aumenta, entre outros fatores, em populações menos favorecidas situadas em zonas mais periféricas<sup>87</sup>.

O acesso à água encanada promove proteção ao desfecho de morte infantil<sup>85,87-90</sup>. Em estratos de maior carência social onde os domicílios não possuem abastecimento de água, coleta de lixo e esgotamento sanitário ligado à rede central, há maior risco de doenças infecciosas e parasitárias e mortalidade infantil<sup>91-95</sup>. Em uma análise da evolução da mortalidade infantil na região metropolitana de São Paulo as políticas públicas voltadas para melhoria de saneamento básico mostraram eficácia na diminuição da mortalidade pós-neonatal<sup>96</sup>.

Tabela 5 – Variáveis aproximadas à camada dos determinantes gerais e respectiva quantidade de artigos\*

variáveis	<i>n</i>
Saneamento básico	11
Renda	6
Região geográfica	4
Taxa de fecundidade	3
Taxa de pobreza	3
Desnutrição infantil	3
Esperança de vida ao nascer	1
Fornecimento de auxílio governamental	1
Turbidez da água	1
Crianças de 10 a 14 anos que trabalham	1
Índice de envelhecimento	1
Despesa pessoal em relação à despesa total em saúde	1
Proporção do valor da produção vegetal em relação ao	1



total do estado	
Taxa de natalidade	1
Índice de Gini	1
Índice de desenvolvimento humano municipal	1
Áreas de exclusão social	1

\*nota: um mesmo artigo pode ter apresentado mais de uma variável

A renda pode influenciar a saúde materno-infantil por meio do acesso a boa qualidade de alimentos, pré-natal adequado e identificação precoce de morbidades<sup>97</sup>. Iniquidades econômicas foram associadas ao nascimento de prematuros<sup>98</sup>. Maior risco de morte infantil está associado a desnutrição<sup>99</sup>, maior concentração de renda<sup>100</sup> e baixo índice per capita<sup>101</sup>. Associado a proteção está um maior crescimento do PIB per capita<sup>95,102</sup>. Com essas associações, evidencia-se a influência e a interação entre as camadas do modelo dos DSS e a importância de políticas públicas intersetoriais no intuito de reduzir as iniquidades em saúde, uma vez que não existe razão biológica para diferenças de expectativas de vida<sup>103</sup>.

## CONCLUSÃO

A despeito de políticas e programas nacionais, a RMM e TMI brasileiras são maiores que de outros países em desenvolvimento e apresentam diferenças internas, com maiores valores nas regiões Norte e Nordeste, evidenciando iniquidades em saúde.

Corroborando com o encontrado em outros estudos, a assistência pré-natal, escolaridade materna, peso ao nascer e idade materna foram as variáveis que mais se associaram à mortalidade materna e infantil assim como saneamento básico e renda. Dado que estes fatores de risco estão bem estabelecidos na literatura e os valores de TMI e RMM continuam aquém do esperado, as estratégias voltadas para essa mortalidade não estão se mostrando suficientes.

A aproximação das variáveis às camadas do modelo dos DSS permitiu evidenciar a camada individual, mais interna, da qual fazem parte peso ao nascer e idade materna. Tal camada é de governabilidade limitada pelos gestores, porém mudanças podem ocorrer por meio de influência das camadas mais externas, como a dos determinantes de condições de vida e trabalho e gerais, evidenciadas por assistência pré-natal e renda respectivamente. Essas camadas são passíveis de intervenção por meio da ação exclusiva dos gestores.

Dessa maneira, é de responsabilidade dos gestores o estabelecimento de intervenções transversais à todas as camadas e o direcionamento de ações nos locais onde há maior necessidade, com o objetivo de reduzir iniquidades em saúde.

## AGRADECIMENTOS

Agradecemos, pelo financiamento, ao Conselho Nacional de Desenvolvimento Científico e Tecnológico – CNPq, modalidade Iniciação Científica (IC) e ao Programa de Suporte à Pós-graduação de Instituições Comunitárias de Educação Superior – PROSUC/CAPES, modalidade taxa, para sua realização.

## REFERÊNCIAS

1. UNICEF. Levels and Trends in Child Mortality Report 2018. [acesso em 27 agosto de 2019]. Disponível em: <http://portaldeboaspraticas.iff.fiocruz.br/wp-content/uploads/2019/06/UN-IGME-Child-Mortality-Report-2018.pdf>
2. OPAS Brasil, Folha informativa – Mortalidade materna, 2018. [acesso em 24 junho de 2019]. Disponível em: [https://www.paho.org/bra.../index.php?option=com\\_content&view=article&id=5741:folha-informativa-mortalidade-materna&Itemid=820](https://www.paho.org/bra.../index.php?option=com_content&view=article&id=5741:folha-informativa-mortalidade-materna&Itemid=820)
3. Brasil. Ministério da Saúde. Mortalidade Perinatal Síntese de Evidências para Políticas de Saúde. Brasília, 2012. [acesso em 27 agosto de 2019]. Disponível em: [http://bvsmms.saude.gov.br/bvs/publicacoes/sintese\\_evidencias\\_mortalidade\\_perinatal.pdf](http://bvsmms.saude.gov.br/bvs/publicacoes/sintese_evidencias_mortalidade_perinatal.pdf)
4. OPAS. Indicadores básicos para a saúde no Brasil: conceitos e aplicações. Brasília, 2008. [acesso em 24 junho de 2019]. Disponível em: <http://tabnet.datasus.gov.br/tabdata/livroidb/2ed/indicadores.pdf>
5. Brasil. Ministério da Saúde. Estudo da Mortalidade de Mulheres de 10 a 49 anos, com Ênfase na Mortalidade Materna. Relatório Final. Brasília, 2006. [acesso em 24 de junho de 2019]. Disponível em: [http://bvsmms.saude.gov.br/bvs/publicacoes/cd11\\_03estudo\\_mortalidade\\_mulher.pdf](http://bvsmms.saude.gov.br/bvs/publicacoes/cd11_03estudo_mortalidade_mulher.pdf)
6. Brasil. Ministério da Saúde. Manual dos Comitês de Mortalidade Materna. Brasília, 2007. [acesso em 24 de junho de 2019]. Disponível em: [http://bvsmms.saude.gov.br/bvs/publicacoes/comites\\_mortalidade\\_materna\\_3ed.pdf](http://bvsmms.saude.gov.br/bvs/publicacoes/comites_mortalidade_materna_3ed.pdf)
7. Dahlgren G, Whitehead M. Policies and strategies to promote social equity in health. Stockholm: Institute for Future Studies; 1991.
8. Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(7): e1000097.
9. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics, 1977; 33:59-174.
10. Barbosa TAGS, Coelho KR, Andrade GN, Bittencourt SDA, Leal MC, Gazzinelli A. Determinantes da mortalidade infantil em municípios do Vale do Jequitinhonha, Minas Gerais, Brasil. Rev. Mineira de Enfermagem. 2014;18(4):907-914.
11. Lansky S, França E, Leal MC. Mortalidade perinatal e evitabilidade: revisão de literatura. Rev Saúde Pública. 2002;36(6):759-772.
12. Lima EFA, Sousa AI, Griep RH, Primo CC. Fatores de risco para mortalidade neonatal no município de Serra, Espírito Santo. Rev. Bras. Enferm. 2012;65(4):578-85.

13. Morse ML, Fonseca SC, Barbosa MD, Calil MB, Eyer FPC. Mortalidade materna no Brasil: o que mostra a produção científica nos últimos 30 anos?. *Cad. Saúde Pública*. 2011;27(4):623-638.
14. Faria DR, Sousa RC, Costa TJNM, Leite ICG. Mortalidade materna em cidade-polo de assistência na região Sudeste: tendência temporal e determinantes sociais. *Rev. Med. Minas Gerais*. 2012;22(1):1-128.
15. PortalODS. Mortalidade materna cresce no Brasil. [acesso em 27 de agosto de 2019]. Disponível em: <http://portalods.com.br/noticias/mortalidade-materna-cresce-no-brasil/>
16. Almeida MF, Alencar GP, Schoeps D, Novaes HMD, Campbell O, Rodrigues LC. Sobrevida e fatores de risco para mortalidade neonatal em uma coorte de nascidos vivos de muito baixo peso ao nascer, na Região Sul do Município de São Paulo, Brasil. *Cad. Saúde Pública*. 2011;27(6):1088-1098.
17. Andrade CLT, Szwarcwald CL. Análise espacial da mortalidade neonatal precoce no Município do Rio de Janeiro, 1995-1996. *Cad. Saúde Pública*. 2001;17(5):1199-1210.
18. Simões VMF, Silva AAM, Bettiol H, Filho FL, Tonial SR, Mochel EG. Características da gravidez na adolescência em São Luís, Maranhão. *Rev Saúde Pública*. 2003;37(5):559-565.
19. Ribeiro AM, Guimarães MJ, Lima MC, Sarinho SW, Coutinho SB. Fatores de risco para mortalidade neonatal em crianças com baixo peso ao nascer. *Rev Saúde Pública*. 2009;43(2):246-255.
20. Mombelli MA, Sass A, Molena CAF, Téston EF, Marcon SS. Fatores de risco para mortalidade infantil em municípios do Estado do Paraná, de 1997 a 2008. *Rev Paul Pediatr*. 2012;30(2):187-194.
21. Aquino TA, Guimarães MJB, Sarinho SW, Ferreira LOC. Fatores de risco para a mortalidade perinatal no Recife, Pernambuco, Brasil, 2003. *Cad. Saúde Pública*. 2007;23(12):2853-2861.
22. Gonçalves AC et al. Tendência da mortalidade neonatal na cidade de Salvador (Bahia-Brasil), 1996-2012. *Rev. Bras. Saude Mat. Infant*. 2015;15(3):337-347.
23. Oliveira LC, Costa AAR. Óbitos fetais e neonatais entre casos de near miss materno. *Rev. Assoc. Med. Bras*. 2013;59(5):487-494.
24. Kropiwiec MV, Franco SC, Amaral AR. Fatores associados à mortalidade infantil em municípios com índice de desenvolvimento humano elevado. *Rev Paul Pediatr*. 2017;35(4):391-398.
25. Feitoza HAC, Koifman S, Koifman RJ, Saraceni V. Os efeitos maternos, fetais e infantis decorrentes da infecção por dengue durante a gestação em Rio Branco, Acre, Brasil, 2007-2012. *Cad. Saúde Pública*. 2017;33(5):e00178915.

26. Mota AKM, Filho ALM, Saraceni V, Koifman S. Mortalidade materna e incidência de dengue na Região Sudeste do Brasil: estudo ecológico no período 2001-2005. *Cad. Saúde Pública*. 2012;28(6):1057-1066.
27. Paixão ES et al. Dengue in pregnancy and maternal mortality: a cohort analysis using routine data. *Scientific Reports*. 2018;8(1):1-6.
28. Nascimento LB, Siqueira CM, Coelho GE, Junior JBS. Dengue em gestantes: caracterização dos casos no Brasil, 2007-2015\*. *Epidemiol. Serv. Saúde*. 2017;26(3):433-442.
29. Doldan RV, Costa JSD, Nunes MF. Fatores associados à mortalidade infantil no Município de Foz do Iguaçu, Paraná, Brasil: estudo de caso-controle. *Epidemiol. Serv. Saúde*. 2011;20(4):491-498.
30. Fréu CM, Mottin LM, Migott MD, Martinelli IB, Nunes ML, Geib LTC. Determinantes da mortalidade neonatal de uma coorte de nascidos vivos em Passo Fundo – RS, 2003-2004. *Revista da AMRIGS*. 2008; 52(2):97-102.
31. Lansky S, França E, César CC, Neto LCM, Leal MC. Mortes perinatais e avaliação da assistência ao parto em maternidades do Sistema Único de Saúde em Belo Horizonte, Minas Gerais, Brasil, 1999. *Cad. Saúde Pública*. 2006;22(1):117-130.
32. Shikamura SE, Carvalho MS, Aerts DRGC, Flores R. Distribuição espacial do risco: modelagem da mortalidade infantil em Porto Alegre, Rio grande do Sul, Brasil. *Cad. Saúde Pública*. 2001;17(5):1251-1261.
33. Nascimento RM, Leite AJM, Almeida NMGS, Almeida PC, Silva CF. Determinantes da mortalidade neonatal: estudo caso-controle em Fortaleza, Ceará, Brasil. *Cad. Saúde Pública*. 2012;28(3):559-572.
34. Horta BL, Gigante DP, Candiota JS, Barros FC, Victora CG. Monitorização da mortalidade na coorte de nascimentos de 1982 a 2006, Pelotas, RS. *Rev Saúde Pública*. 2008;42(Supl. 2):108-114.
35. Hernandez AR, Silva CH, Agranonik M, Quadros FM, Goldani MZ. Análise de tendências das taxas de mortalidade e de seus fatores de risco na cidade de Porto Alegre, Rio Grande do Sul, Brasil, no período de 1996 a 2008. *Cad. Saúde Pública*. 2011;27(11):2188-2196.
36. Nascimento EMR, Costa MCN, Mota ELA, Paim JS. Estudo de fatores de risco para óbitos de menores de um ano mediante compartilhamento de bancos de dados. *Cad. Saúde Pública*. 2008;24(11):2593-2602.
37. Pinheiro CEA, Peres MA, d’Orsi E. Aumento na sobrevivência de crianças de grupos de peso baixo ao nascer em Santa Catarina. *Rev Saúde Pública*. 2010;44(5):776-784.
38. Oliveira TG et al. Escore de Apgar e mortalidade neonatal em um hospital localizado na zona sul do município de São Paulo. *Einstein (São Paulo)*. 2012;10(1):22-28.

39. Ferrari RAP, Bertolozzi MR, Dalmas JC, Giroto E. Fatores determinantes da mortalidade neonatal em um município da Região Sul do Brasil. *Rev Esc Enferm USP*. 2013; 47(3):531-538.
40. Lansky S *et al*. Pesquisa *Nascer no Brasil*: perfil da mortalidade neonatal e avaliação da assistência à gestante e ao recém-nascido. *Cad. Saúde Pública*. 2014;30Sup:S192-S207.
41. Duarte JLMB, Mendonça GAS. Fatores associados à morte neonatal em recém-nascidos de muito baixo peso em quatro maternidades no Município do Rio de Janeiro, Brasil. *Cad. Saúde Pública*. 2005;21(1):181-191.
42. Oliveira GS, Lima MCBM, Lyra CO, Oliveira AGRC, Ferreira MAF. Desigualdade espacial da mortalidade neonatal no Brasil: 2006 a 2010. *Ciênc. Saúde Coletiva*. 2013;18(8):2431-2441.
43. Sanders LSC, Pinto FJM, Medeiros CRB, Sampaio RMM, Viana RAA, Lima KJ. Mortalidade infantil: análise de fatores associados em uma capital do Nordeste Brasileiro. *Cad. Saúde Colet*. 2017;25(1):83-89.
44. Brasil. Ministério da saúde. Atenção ao pré-natal de baixo risco. *Cadernos de Atenção Básica*. Brasília, 2012. [acesso em 24 de junho de 2019]. Disponível em: [http://bvsmms.saude.gov.br/bvs/publicacoes/cadernos\\_atencao\\_basica\\_32\\_prenatal.pdf](http://bvsmms.saude.gov.br/bvs/publicacoes/cadernos_atencao_basica_32_prenatal.pdf)
45. Paulucci RS, Nascimento LFC. Mortalidade neonatal em Taubaté: um estudo caso-controle. *Rev Paul de Pediatr*. 2007;25(4):358-363.
46. Apgar V. A proposal for a new method of evaluation of the newborn infant. *Curr Res Anesth Analg*. 1953;32(4):260-267.
47. Maran E, Uchimura TT. Mortalidade neonatal: fatores de risco em um município no sul do Brasil. *Rev. Eletr. de Enfermagem*. 2008;10(1):29-38.
48. Silva CF, Leite AJM, Almeida NMGS, Leon ACMP, Olofin I, Perinatal RNNS. Fatores associados ao óbito neonatal de recém-nascidos de alto risco: estudo multicêntrico em Unidades Neonatais de Alto Risco no Nordeste brasileiro. *Cad. Saúde Pública*. 2014;30(2):355-368.
49. Almeida MFB *et al*. Fatores perinatais associados ao óbito precoce em prematuros nascidos nos centros da Rede Brasileira de Pesquisas Neonatais. *J Pediatr*. 2008;84(4):300-307.
50. Prigenzi MLH, Trindade CEP, Rugolo LMSS, Silveira LVA. Fatores de risco associados à mortalidade de recém-nascidos de muito baixo peso na cidade de Botucatu, São Paulo, no período 1995-2000. *Rev. Bras. Saúde Mater. Infant*. 2008;8(1):93-101.

51. Geib LTC, Fréu CM, Brandão M, Nunes ML. Determinantes sociais e biológicos da mortalidade infantil em coorte de base populacional em Passo Fundo, Rio Grande do Sul. *Ciênc. Saúde Coletiva*. 2010;15(2):363-370.
52. Bertani AL, Garcia T, Tanni SE, Godoy I. Prevenção do tabagismo na gravidez: importância do conhecimento materno sobre os malefícios para a saúde e opções de tratamento disponíveis. *J. Bras. Pneumologia*. 2015;41(2):175-181.
53. Cunha AJLA, Leite AJM, Almeida IS. The pediatrician's role in the first thousand days of the child: the pursuit of healthy nutrition and development. *J Pediatr*. 2015;91(6): S44-S51.
54. Moore TG, Arefabid N, Derry A, West S. The First Thousand Days: An Evidence Paper. Murdoch Children's Research Institute: Melbourne, Centre for Community Child Health, 2017. [acesso em 03 de junho de 2019]. Disponível em: <https://www.suicideinfo.ca/resource/the-first-thousand-days-an-evidence-paper/>
55. Ribeiro FD, Ferrari RAP, Sant'Anna FL, Dalmas JC, Giroto E. Extremos de idade materna e mortalidade infantil: análise entre 2000 e 2009. *Rev Paul Pediatr*. 2014;32(4):381-388.
56. Schoeps D et al. Fatores de risco para mortalidade neonatal precoce. *Rev Saúde Pública*. 2007;41(6):1013-1022.
57. Jobim R, Aerts D. Mortalidade infantil evitável e fatores associados em Porto Alegre, Rio Grande do Sul, Brasil, 2000-2003. *Cad. Saúde Pública*. 2008;21(1):179-187.
58. França E, Souza JM, Guimarães MDC, Goulart EMA, Colosimo E, Antunes CMF. Associação entre fatores sócio-econômicos e mortalidade infantil por diarreia, pneumonia e desnutrição em região metropolitana do Sudeste do Brasil: um estudo caso-controle. *Cad. Saúde Pública*. 2001;17(6): 1437-1447.
59. Oliveira EFV, Gama SGN, Silva CMFP. Gravidez na adolescência e outros fatores de risco para mortalidade fetal e infantil no Município do Rio de Janeiro, Brasil. *Cad. Saúde Pública*. 2010;26(3):567-578.
60. Leite RMB, Araújo TVB, Albuquerque RM, Andrade ARS, Neto PJD. Fatores de risco para mortalidade materna em área urbana do Nordeste do Brasil. *Cad. Saúde Pública*. 2011;27(10):1977-1985.
61. Neto OLM, Barros MBA. Fatores de risco para mortalidade neonatal e pós-neonatal da Região Centro-Oeste do Brasil: linkage entre bancos de dados de nascidos vivos e óbitos infantis. *Cad. Saúde Pública*. 2000;16(2):477-485.
62. Carreno I, Bonilha ALL, Costa JSD. Perfil epidemiológico das mortes maternas ocorridas no Rio Grande do Sul, Brasil: 2004-2007. *Rev. Bras. Epidemiologia*. 2012;15(2):396-406.
63. Pereira PK, Vieira CL, Santos JFC, Lima LA, Legay LF, Lovisi GM. Avaliação de desfechos perinatais/infantis em partos de pacientes com transtornos mentais maiores

- de um hospital psiquiátrico do Rio de Janeiro, Brasil. *Cad. Saúde Pública*. 2014;30(8):1654-1666.
64. Silva CF, Leite AJM, Almeida NMGS, Gondim RC. Fatores de risco para mortalidade infantil em município do Nordeste do Brasil: linkage entre bancos de dados de nascidos vivos e óbitos infantis – 2000 a 2002. *Rev. Bras. Epidemiologia*. 2006;9(1):69-80.
  65. Silva ZP et al. Características dos nascidos vivos, das mães e mortalidade neonatal precoce na Região Metropolitana de São Paulo, Brasil. *Cad. Saúde Pública*. 2009;25(9):1981-1989.
  66. Giglio MRP, Lamounier JA, Neto OLM. Via de parto e risco para mortalidade neonatal em Goiânia no ano de 2000. *Rev. Saúde Publica*. 2005;39(3):350-357.
  67. Leal MC, Bittencourt SDA, Torres RMC, Niquini RP, Junior PRBS. Determinantes do óbito infantil no Vale do Jequitinhonha e nas regiões Norte e Nordeste do Brasil. *Rev Saúde Pública*. 2017;51:12.
  68. Almeida MF et al. Partos Domiciliares acidentais na região sul do Município de São Paulo. *Rev Saúde Pública*. 2005;39(3):366-375.
  69. Maia LTS, Souza WV, Mendes ACG. Diferenciais nos fatores de risco para a mortalidade infantil em cinco cidades brasileiras: um estudo de caso-controle com base no SIM e no SINASC. *Cad. Saúde Pública*. 2012;28(11):2163-2176.
  70. Araújo BF, Bozzetti MC, Tanaka AC. A. Mortalidade neonatal precoce no município de Caxias do Sul: um estudo de coorte. *J Pediatr*. 2000;76(3):200-206.
  71. Viellas EF, Gama SGN, Carvalho ML, Pinto LW. Fatores associados à agressão física em gestantes e os desfechos negativos no recém-nascido. *J Pediatr*. 2013;89(1):83-90.
  72. Araújo BF *et al.* Influência do local de nascimento e do transporte sobre a morbimortalidade de recém-nascidos prematuros. *J Pediatr*. 2011;87(3):257-262.
  73. Brasil. Ministério da saúde. Boletim Epidemiológico - Sífilis 2016. Secretária de Vigilância em Saúde. [acesso em 24 de junho de 2019]. Disponível em: [http://portalarquivos2.saude.gov.br/images/pdf/2016/outubro/31/2016\\_030\\_Sifilis-publicacao2.pdf](http://portalarquivos2.saude.gov.br/images/pdf/2016/outubro/31/2016_030_Sifilis-publicacao2.pdf)
  74. Brasil. Ministério da saúde. Boletim Epidemiológico – Sífilis 2017. Secretária de Viigilância em Saúde. [acesso em 24 de junho de 2019]. Disponível em: <http://portalarquivos.saude.gov.br/images/pdf/2017/novembro/13/BE-2017-038-Boletim-Sifilis-11-2017-publicacao-.pdf>
  75. Cardoso ARP, Araújo MAL, Cavalcante MS, Frota MA, Melo SP. Análise dos casos de sífilis gestacional e congênita nos anos de 2008 a 2010 em Fortaleza, Ceará, Brasil. *Ciênc. saúde colet*. 2018;23(2):563-574.

76. Luz LA, Aquino R, Medina MG. Avaliação da qualidade da Atenção Pré-Natal no Brasil. *Saúde Debate*. 2018;42(2):111-126.
77. Mendes RB et al. Avaliação da qualidade do pré-natal a partir das recomendações do programa de humanização no pré-natal e nascimento. *Ciê e saúde colet*. [acesso em 29 de agosto de 2019]. Disponível: <http://www.cienciaesaudecoletiva.com.br/artigos/avaliacao-da-qualidade-do-prenatal-a-partir-das-recomendacoes-do-programa-de-humanizacao-no-prenatal-e-nascimento/16895?id=16895>
78. Almeida WS, Szwarcwald CL. Mortalidade infantil e acesso geográfico ao parto nos municípios brasileiros. *Rev Saúde Pública*. 2012;46(1):68-76.
79. Pereira APE, Gama SGN, Leal MC. Mortalidade infantil em uma amostra de nascimentos do município do Rio de Janeiro, 1999-2001: “linkage” com o Sistema de Informação de Mortalidade. *Rev. Bras. Saúde Mater. Infant*. 2007;7(1):83-88.
80. Troncon JK, Netto DLQ, Rehder PM, Cecatti JG, Surita FG. Mortalidade materna em um centro de referência do Sudeste Brasileiro. *Rev Bras Ginecol Obstet*. 2013;35(9):388-393.
81. Zuanazzi PT, Cabral PHV, Stella MA, Moraes GI. O tempo de percurso a cidades polos regionais influencia na mortalidade em pequenos municípios gaúchos. *Cad. Saúde Pública*. 2017;33(12):e00187515.
82. Almeida MCS, Gomes CMS, Nascimento LFC. Análise espacial da mortalidade neonatal no estado de São Paulo, 2006-2010. *Rev Paul Pediatr*. 2014;32(4):374-380.
83. Nascimento LFC, Bastista GT, Dias NW, Catelani CS, Becker D, Rodrigues L. Análise espacial da mortalidade neonatal no Vale do Paraíba, 1999 a 2001. *Rev Saúde Pública*. 2007;41(1):94-100.
84. Neto OLM, Barros MBA, Martelli CMT, Silva SA, Cavenaghi SM, Jr JBS. Diferenças no padrão de ocorrência da mortalidade neonatal e pós-neonatal no Município de Goiânia, Brasil, 1992-1996: análise espacial para identificação das áreas de risco. *Cad. Saúde Pública*. 2001;17(5):1241-1250.
85. Zanini RR, Moraes AB, Giugliani ERJ, Riboldi J. Determinantes contextuais da mortalidade neonatal no Rio Grande do Sul por dois modelos de análise. *Rev Saúde Pública*. 2011;45(1):79-89.
86. Brasil. Ministério da Saúde. Informação e análise epidemiológica Indicadores que utilizam a metodologia do Busca Ativa. Brasília, 2019. [acesso em 26 de junho de 2019]. Disponível em: <http://svs.aids.gov.br/dantps/acesso-a-informacao/acoes-e-programas/busca-ativa/indicadores-de-saude/mortalidade/>
87. Costa MCN, Azi PA, Paim JS, Silva LMV. Mortalidade infantil e condições de vida: a reprodução das desigualdades sociais em saúde na década de 90. *Cad. Saúde Pública*. 2001;17(3):555-567.



88. Sousa TRV, Filho PAML. Análise por dados em painel do status de saúde no Nordeste Brasileiro. *Rev Saúde Pública*. 2008;42(5):796-804.
89. Teixeira JC, Guilhermino RL. Análise da associação entre saneamento e saúde nos estados brasileiros, empregando dados secundários do banco de dados indicadores e dados básicos para a saúde 2003 – IDB 2003. *Eng. Sanit. Ambient*. 2006;11(3):277-282.
90. Filho JGB, Pontes LRFSK, Miná DL, Barreto ML. Mortalidade infantil e condições sociodemográficas no Ceará, em 1991 a 2000. *Rev Saúde Pública*. 2007;41(6):1023-1031.
91. Vilela MBR, Bonfim C, Medeiros Z. Mortalidade infantil por doenças infecciosas e parasitárias: reflexo das desigualdades sociais em um município do Nordeste do Brasil. *Rev. Bras. Saúde Mater. Infant*. 2008;8(4):455-461.
92. Filho JGB, Kerr LRFS, Miná DL, Barreto ML. Distribuição espacial da taxa de mortalidade infantil e principais determinantes no Ceará, Brasil, no período 2000-2002. *Cad. Saúde Pública*. 2007;23(5):1173-1185.
93. Bühler HF, Ignotti E, Neves SMAS, Hacon SS. Análise espacial de indicadores integrados determinantes da mortalidade por diarreia aguda em crianças menores de 1 ano em regiões geográficas. *Ciênc. Saúde Coletiva*. 2014;19(10):4131-4140.
94. Costa SS, Heller L, Brandão CCS, Colosimo EA. Indicadores sanitários sentinelas da mortalidade infantil: um estudo ecológico de base municipal em quatro estados brasileiros. *Cad. Saúde Coletiva*. 2005;13(1):131-150.
95. Costa MCN, Mota ELA, Paim JS, Silva LMV, Teixeira MG, Mendes CMC. Mortalidade infantil no Brasil em períodos recentes de crise econômica. *Rev Saúde Pública*. 2003;37(6):699-706
96. Holcman MM, Latorre MRDO, Santos JLF. Evolução da mortalidade infantil na região metropolitana de São Paulo, 1980-2000. *Rev Saúde Pública*. 2004;38(2):180-186.
97. Blencowe H et al. Born Too Soon: The global epidemiology of 15 million preterm births. *Reproductive Health*. 2013;10(1):S2.
98. Sadovsky ADI, Matijasevich A, Santos IS, Barros FC, Miranda AE, Silveira MF. Iniquidades socioeconômicas em nascimentos prematuros em quatro estudos brasileiros de coortes de nascimento. *J Pediatr*. 2018;94(1):15-22.
99. Einloft PR, Garcia PC, Piva JP, Bruno F, Kipper DJ, Fiori RM. Perfil epidemiológico de dezesseis anos de uma unidade de terapia intensiva pediátrica. *Rev Saúde Pública*. 2002;36(6):728-733.

100. Fischer TK, Lima D, Rosa R, Osório D, Boing AF. A mortalidade infantil no Brasil: série histórica entre 1994-2004 e associação com indicadores socioeconômicos em municípios de médio e grande porte. *Rev. Medicina Ribeirão Preto.* 2007;40(4):559-566.
101. Boing AF, Boing AC. Mortalidade infantil por causas evitáveis no Brasil: um estudo ecológico no período 2000-2002. *Cad. Saúde Pública.* 2008;24(2):447-455.
102. Lourenço EC et al. Variáveis de impacto na queda da mortalidade infantil no Estado de São Paulo, Brasil, no período de 1998 a 2008. *Ciênc. Saúde Coletiva.* 2014;19(7):2055-2062.
103. Marmot M. Social determinants of health inequalities. *Lancet.* 2005;365:1099-1104.

## APÊNDICE B – ARTIGO DE REVISÃO INTEGRATIVA SOBRE MINERAÇÃO DE DADOS NO APOIO À GESTÃO EM SAÚDE MATERNO-INFANTIL

### MINERAÇÃO DE DADOS NO APOIO À GESTÃO EM SAÚDE MATERNO-INFANTIL: UMA REVISÃO INTEGRATIVA

#### DATA MINING TO MATERNAL AND CHILD HEALTH MANAGEMENT SUPORT: INTEGRATIVE REVIEW

##### RESUMO

Introdução: a despeito do conhecimento existente e ações já implementadas, desfechos indesejáveis perduram nos primeiros 1.000 dias de vida, com impactos na vida adulta e terceira idade. A Mineração de Dados (MD) é uma etapa do *Knowledge Discovery in Database* (KDD), processo computacional que visa a descobrir conhecimentos novos, válidos e úteis em base de dados, fornecendo subsídios para a tomada de decisão. Objetivo: verificar o potencial do uso da MD para apoio à gestão em saúde materno-infantil. Métodos: busca adaptada do método PRISMA, realizada nos repositórios BVS, PubMed, Scopus (Elsevier) e IEEE Xplore, utilizando-se os descritores “data mining” AND health AND child\* OR maternal OR pregnan\*. Resultados: foram incluídos 29 documentos, dos quais nove corroboraram a literatura, dez focaram a acurácia do modelo e dez encontraram resultados compreensíveis, válidos, novos e úteis, como associação de novos elementos químicos ao ozônio e asma, grupo étnico australiano e natimortalidade, raça indígena americana e prematuridade, vacinação e prematuridade, gestantes idosas e baixa adesão ao pré-natal. Conclusão: os resultados demonstram o potencial do uso da MD para apoio à gestão, especialmente quando a interdisciplinaridade se faz

presente. Identificou-se uma lacuna entre o uso da MD e a aplicação efetiva dos seus resultados.

**PALAVRAS-CHAVE:** Saúde materno-infantil. Determinantes sociais da saúde. Mineração de dados.

## **ABSTRACT**

Despite existing knowledge and actions deployed so far, negative outcomes persist over the first 1,000 days of life, arising impacts on adults and elders. Data mining (DM) is part of KDD (Knowledge Discovery in Database), a computational process focused on new, valid and useful knowledge discovery in database in order to endorse decision making. This integrative review aims to verify use potential of MD to maternal and child health management support. Search process was adapted from PRISMA method, applied on BVS, PubMed, Scopus (Elsevier) and IEEE Xplore repositories through the MeSH® terms “data mining” AND health AND child \* OR maternal OR pregnan\*. Twenty-nine documents were included: nine confirmed previous findings, ten focused on models accuracy (lack of discussion with previous studies), and ten discovered understandable, valid, new and useful results – associations between new chemical elements and O3 on asthma, specific Australian ethnicity and stillbirth, Indian American race and prematurity, vaccination and prematurity, geriatric pregnancy and low prenatal adherence, etc. Results indicated the DM’s potential to maternal and child health management support, notably in interdisciplinary domains. A gap was identified between the use of DM and its results’ effective application.

**KEYWORDS:** Maternal and Child Health. Social Determinants of Health. Data mining.

## INTRODUÇÃO

Os primeiros 1.000 dias de vida são considerados críticos, pois situações ocorridas nesse período afetam a vida e a saúde da gestante e da criança, com repercussões na fase adulta e terceira idade<sup>1</sup>. Essas repercussões podem ser positivas, como aumento da estatura, melhoria do desenvolvimento cognitivo, motor e socioemocional, da performance escolar, da capacidade de aprendizagem, de trabalho e produtividade. Entretanto, implicações negativas suscitam elevação da criminalidade, analfabetismo, obesidade, doenças cardíacas, problemas de saúde mental, morbidade e mortalidade na infância<sup>2-7</sup>.

Esse processo saúde-doença é influenciado por condições individuais (idade, sexo e estilo de vida) e coletivas (ambientais e socioeconômicas), representadas pelos Determinantes Sociais da Saúde (DSS)<sup>8</sup>. O modelo compreende um conjunto de cinco camadas concêntricas que demonstram as principais influências na saúde e respectivo nível de intervenção, por meio do qual é possível inferir o grau de governabilidade do Estado.

As camadas mais internas remetem às condições individuais, de menor governabilidade do Estado, e as mais externas, às condições gerais, de maior governabilidade do Estado e que influenciam e modulam as camadas mais individuais. Mudanças focadas nas camadas mais internas requerem ações pontuais, de prevenção ou educação em saúde, enquanto as intervenções voltadas às camadas mais externas são mais efetivas e demandam ações de promoção, com mudanças estruturais por meio de estratégias econômicas e/ou acordos envolvendo outros setores além da saúde<sup>8</sup>.

A despeito do conhecimento existente e ações já implementadas, dada a complexidade dos primeiros 1.000 dias de vida, desfechos indesejáveis, e por vezes evitáveis, perduram, ocasionando problemas de saúde pública<sup>7,9</sup>. Nesse sentido, a Mineração de Dados (MD) pode auxiliar na descoberta de novos conhecimentos que possam vir a subsidiar a tomada de decisão<sup>10</sup>.

A MD é a segunda etapa do processo de *Knowledge Discovery in Database* (KDD) ou Descoberta de Conhecimento em Bases de Dados. A primeira etapa é o pré-processamento e a última, o pós-processamento. O KDD é um processo computacional que visa extrair informação e conhecimento implícito e escondido em uma base de dados, que podem ser úteis e até então desconhecidos<sup>10</sup>. O algoritmo identifica padrões, relacionando variáveis e valores com a ocorrência ou não do desfecho estudado. Assim, a partir desses resultados (padrões), é possível identificar associações entre fatores de exposição e desfecho. Para gestão em saúde, é importante que essas associações estejam representadas de forma compreensível, de maneira que seres humanos possam identificar as causas (fatores de exposição) e consequências do problema de saúde (desfecho)<sup>11-12</sup>.

A associação entre fatores de exposição e desfecho pode ser testada por métodos estatísticos convencionais. Entretanto, os resultados podem não ser de fácil interpretação, especialmente quando mais de duas variáveis são consideradas simultaneamente<sup>13</sup>. Testes estatísticos apresentam inegável contribuição para delinear o comportamento dos dados<sup>14</sup>, porém, quando se dispõe de grande volume de dados, processos oriundos da computação, como a MD, podem vir a contribuir<sup>10,15</sup>. Os testes estatísticos convencionais requerem um conhecimento prévio de dependência entre as variáveis, pressuposto não exigido para a realização da MD, aumentando a possibilidade de descoberta de novos conhecimentos<sup>10-11,16</sup>,

cuja validade e utilidade precisam ser verificadas por outros meios, como a avaliação de especialistas, a literatura existente e a própria estatística convencional<sup>14</sup>.

Considerando a importância dos primeiros 1.000 dias de vida, a permanência de desfechos indesejáveis nesse período e a finalidade da MD de descobrir conhecimentos novos, válidos e úteis, o objetivo deste estudo é identificar, por meio de uma revisão integrativa, os DSS dos primeiros 1.000 dias descobertos pela MD para apoio à gestão em saúde materno-infantil.

## **MÉTODOS**

As etapas foram adaptadas do método *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA)<sup>17</sup>. Para a busca dos documentos, foram adotados os descritores “data mining” AND health AND child\* OR maternal OR pregnan\*, nos repositórios BVS, PubMed, Scopus (Elsevier) e IEEE Xplore, em maio de 2018. O filtro utilizado foi de documentos do tipo artigo, exceto para o repositório IEEE Xplore, em que também foram consideradas conferências, dada a natureza de documentos mais voltados à computação. Não foram aplicados critérios de temporalidade nem idioma, no intuito de ampliar o escopo dos documentos.

Após a exclusão dos duplicados, foram analisados o título e/ou resumo. Para a leitura integral, foram selecionados documentos que adotassem a MD (KDD, tarefas de mineração, algoritmos, etc.) para o enfrentamento de qualquer situação relacionada à saúde materno-infantil, incluindo gestação, parto, puerpério, nascimento até os dois anos de idade da criança, compreendendo, dessa maneira, todo o período dos primeiros 1.000 dias de vida. Essa etapa de seleção foi realizada por dois pesquisadores independentes e avaliada a partir do teste de concordância

Kappa. As calibrações foram realizadas até atingir um valor de Kappa maior que 0,80<sup>18</sup>. Os documentos com indicativo de seleção divergentes foram avaliados por um terceiro pesquisador independente.

A elegibilidade dos documentos foi baseada na leitura integral a partir dos seguintes critérios: idioma em alfabeto romano e representação compreensível dos resultados obtidos pela MD, permitindo identificar os fatores de exposição e desfecho. Foram excluídos documentos disponíveis somente por compra ou comutação bibliográfica internacional, que contemplassem população infantil cuja faixa etária não estivesse especificada ou fora dos primeiros mil dias de vida. As características extraídas, respectiva descrição e finalidade estão relacionadas no Quadro 1.

Para melhor compreensão, as características de temporalidade, quantidade de registros e variáveis das bases de dados, área de publicação, aderência à literatura e DSS serão melhor detalhadas.

Para a temporalidade, quantidade de registros e de variáveis das bases de dados, após a extração de cada característica, foi calculada a mediana e verificada a posição dos estudos em relação a ela (inferior ou superior).

Com relação à área de publicação, para as revistas, foi consultado o portal de revistas da BVS (<http://portal.revistas.bvs.br/index.php?lang=pt>) e NLM Catalog/PubMed (<https://www.ncbi.nlm.nih.gov/nlmcatalog?cmd=historysearch&querykey=3>). Consideraram-se como área de publicação em saúde aqueles documentos cuja área temática e assuntos (BVS) ou *current subset* e *broad subject term(s)* (PubMed) mencionavam somente a saúde. Quando havia menção à área computacional (por exemplo, informática em saúde), entendeu-se como área de publicação interdisciplinar. A área de publicação em informática foi determinada



quando não havia nenhuma menção à saúde (por exemplo, anais de congressos de MD, de modelagem e sistemas inteligentes, de engenharia de *software* e inteligência artificial etc.).

A aderência das associações à literatura foi considerada conforme o relatado pelos autores de cada estudo: corroboram a literatura, divergem da literatura ou interessantes frente à literatura. Tanto os resultados divergentes quanto os interessantes diferem da literatura, porém os primeiros foram considerados erros, justificados por possíveis vieses e limitações do estudo, enquanto os segundos, vistos como surpreendentes, por haver plausibilidade para o achado, sugerindo outros estudos. Um mesmo estudo pode ter encontrado diferentes associações.

A aproximação dos fatores de exposição às camadas do modelo dos DSS visou a inferir o potencial de utilidade do padrão descoberto para a gestão em saúde, considerando o pressuposto de que, quanto mais externa for a camada do modelo dos DSS, maiores serão a governabilidade do Estado e a probabilidade de resultados efetivos<sup>8</sup>.

Assim, os estudos com potencial de contribuição à gestão em saúde materno-infantil incluíram aqueles que estavam alinhados ao objetivo do KDD<sup>10</sup> de identificar padrões compreensíveis (critério para elegibilidade), válidos, novos (aderência à literatura) e potencialmente úteis (aproximação às camadas mais externas do modelo dos DSS).

## **RESULTADOS**

Foram identificados 834 documentos nos repositórios: PubMed (268), BVS (212), IEEE Xplore (111) e Scopus Elsevier (243). Trezentos e sessenta e seis foram excluídos por duplicidade e 468, avaliados por dois pesquisadores independentes na

etapa de seleção. O valor de Kappa obtido foi de 0,884, sendo considerado uma concordância quase perfeita<sup>18</sup>. Vinte e cinco documentos foram avaliados por um terceiro pesquisador, pela divergência de indicação de seleção pelos dois primeiros.

A partir da Figura 1, é possível identificar o processo de seleção, inclusão e exclusão até a obtenção dos 29 documentos incluídos.

Os desfechos estudados variaram de uma doença específica (oito estudos), um fluxo ou processo de trabalho na assistência (oito estudos) a situações de saúde mais abrangentes (13 estudos). A prematuridade foi o desfecho mais estudado (10,0%).

Dez estudos encontraram, dentre seus achados, associações interessantes, sendo considerados com potencial de contribuição à gestão em saúde materno-infantil (34,5%). Dos 19 estudos restantes, dez apresentaram alguns resultados, porém não discutiam as associações encontradas à luz da literatura, focando a acurácia do modelo computacional utilizado (34,5%); oito encontraram somente resultados que corroboravam a literatura, não agregando novas descobertas (27,6%); e um estudo encontrou associações que corroboravam e diferiam da literatura (3,4%).

Dos dez estudos com potencial de contribuição à gestão, cinco foram publicados na área interdisciplinar (50,0%), quatro, na saúde (40,0%) e um, na informática (10,0%) – único documento com esse potencial, dentre os 11 publicados na área.

Com relação às bases de dados, a menor janela temporal foi de quatro meses e a maior, de 20 anos. Houve variação também com relação à quantidade de registros (menos de cem e acima de um milhão) e variáveis (menos de dez e acima de 1.500). Alguns estudos não informaram tais características – quatro não

informaram a janela temporal, três, a quantidade de registros e 11, a quantidade de variáveis. Foram obtidas as seguintes medianas para cada característica da base de dados: cinco anos, 5.903 registros e 401 variáveis. Considerando apenas os dez estudos com potencial de contribuição à gestão, o número inferior e superior à mediana de cada característica foi próximo (Tabela 1).

A aproximação dos fatores às camadas do modelo dos DSS<sup>8</sup> demonstrou que dos dez estudos com potencial de contribuição à gestão, dois aproximavam-se da camada individual e oito, das mais externas do modelo, de condições socioambientais gerais ou intermediárias, abrangendo situações de vida e trabalho, incluindo o acesso a serviços de saúde.

Todos os 29 estudos foram experimentos transversais e pontuais, não tendo sido evidenciada em nenhum a aplicação dos resultados obtidos pela MD. Os estudos incluídos nesta revisão e respectivas características extraídas são demonstrados no Quadro 2.

## **DISCUSSÃO**

A possibilidade de aplicação da MD em diferentes contextos nos primeiros 1.000 dias de vida foi identificada pela diversidade de desfechos estudados: doenças bucais<sup>34</sup>, com forte componente ambiental<sup>23,28</sup>; doenças crônicas transmissíveis<sup>39</sup>, agudas<sup>14</sup> ou raras<sup>41</sup>; falhas na assistência, tanto na atenção primária à saúde<sup>32-33,38</sup> quanto na secundária ou terciária<sup>13,29</sup>; situações de saúde como desnutrição infantil<sup>15,40</sup>, mortalidade materna<sup>35-36</sup>, mortalidade de neonatos até crianças de cinco anos de idade<sup>22,24,42-43</sup>, entre outros. A prematuridade foi o desfecho mais estudado<sup>19,21,25</sup>, corroborando sua importância em relação aos óbitos

neonatais. Estima-se que, em 2017, ocorreram 2,5 milhões de óbitos neonatais, dos quais dois terços eram prematuros<sup>45</sup>.

O produto final de uma base de dados é o conhecimento<sup>46</sup>. Ressalta-se que sua qualidade interfere no processo KDD e seus resultados<sup>10</sup>. Em alguns estudos, para minimizar essas interferências, pesquisadores recorreram a algumas estratégias como a imputação<sup>26-27,35</sup>, exclusão de registros<sup>13,22,32,41</sup> e de variáveis consideradas redundantes<sup>19,40</sup>. Outros indícios da falta de qualidade nos bancos foram associações encontradas, porém relatadas como tendo possibilidade de viés devido à não completude dos registros<sup>24</sup>.

Alguns estudos não relataram a quantidade de registros<sup>27-28,36</sup>, de variáveis utilizadas<sup>13-14,27-29,31-33,38-39,44</sup> ou a janela temporal das bases de dados<sup>30,34</sup>. A falta dessas informações dificulta o dimensionamento do banco de dados, a reprodutibilidade do experimento e a possibilidade de agregação de novas variáveis em estudos futuros.

A conformação da base de dados e a demonstração dos resultados encontrados estão diretamente relacionadas à finalidade das atividades do KDD: desenvolvimento tecnológico (concepção, desenvolvimento, aprimoramento e otimização de algoritmos), execução (uso de algoritmos desenvolvidos na busca de conhecimentos) e aplicação de resultados (aplicação efetiva dos resultados obtidos)<sup>11</sup>. Uma das dificuldades desta revisão foi comparar diferentes finalidades de atividades do KDD sob o ponto de vista de contribuição à gestão.

Para a atividade de desenvolvimento tecnológico, por exemplo, uma menor quantidade de variáveis diminui os parâmetros dos algoritmos e aumenta a eficiência computacional<sup>47</sup>. Por exemplo, na previsão de prematuridade, a inclusão de centenas de variáveis às sete mais fortemente associadas aumentou a acurácia em

apenas 0,03 na curva ROC<sup>19</sup>. Ainda, o estudo com a menor quantidade de registros foi considerado com potencial de contribuição à gestão, por ter encontrado um ponto de inflexão no tempo decorrido de cirurgia, a partir do qual há uma redução significativa na prevenção de leucomalácia periventricular<sup>30</sup>.

Ao avaliar a quantidade de estudos com potencial de contribuição à gestão em relação às medianas de temporalidade da base de dados, quantidade de registros e variáveis, o número de estudos com valores inferiores ou superiores a elas ficou próximo, suscitando a hipótese de que o tamanho da base de dados não é condição indispensável, pois, quando a MD descobre conhecimentos potencialmente úteis para o apoio à tomada de decisão, a base de dados atendeu ao esperado, independentemente de seu tamanho<sup>48</sup>.

Nesse sentido, dez documentos trouxeram resultados compreensíveis, válidos e novos (considerados, nesta revisão, como interessantes)<sup>13-14,19,21,24,30-31,33-34,38</sup> e sua utilidade ficou evidenciada ao identificar que as camadas mais externas do modelo dos DSS foram as que mais tiveram aproximações. São essas camadas que influenciam e modulam as mais internas e onde se encontra a maior possibilidade de resultados efetivos, pois as ações nelas demandam promoção de saúde, políticas públicas e outras estratégias intersetoriais, de governabilidade única e exclusiva do Estado<sup>8</sup>.

Nove estudos encontraram associações que corroboraram e/ou divergiram da literatura, não agregando nenhum novo conhecimento. Considerando que os resultados da MD dependem dos dados<sup>19</sup>, a não descoberta de algo novo pode ser devido ao uso de bases de dados com variáveis cuja associação ao desfecho já está bem estabelecida<sup>49</sup>. Para maior probabilidade de descoberta de conhecimento que possa auxiliar no planejamento e ação mais assertiva, é desejável, além das

variáveis já conhecidas, outras novas, contemplando todas as camadas possíveis, aumentando a possibilidade de encontrar conhecimentos novos e úteis para o estabelecimento de hipóteses distintas<sup>24</sup>.

Ressalta-se que, para a aderência à literatura, característica por meio da qual foram inferidas a validade e a novidade dos resultados, foi acatado o relatado pelos autores de cada estudo, o que pode ser considerado um viés importante, visto que os estudos são de áreas diferentes e possuem finalidades distintas, tendo em comum apenas o uso da MD. Contudo, a descoberta de resultados interessantes demonstra o cumprimento da finalidade da MD, de descoberta de conhecimento, fomentando o estabelecimento de novas hipóteses, que, ao serem confirmadas, poderão auxiliar na tomada de decisão<sup>10-11</sup>.

Para um resultado mais efetivo, a MD exige a interdisciplinaridade, pois os problemas que se propõe a resolver não são da área computacional, sendo ela um meio para contribuir com a solução<sup>50</sup>. Espera-se que os resultados encontrados sejam interpretados por especialistas na área para avaliação de sua utilidade<sup>14,20</sup>. Detectou-se que os estudos publicados na área de informática foram os que mais pecaram nesse sentido, pois, em sua grande maioria, não foram considerados com potencial de contribuição à gestão<sup>15,20,25-28,32,35-36,40</sup>, sugerindo que a falta de interdisciplinaridade interfere na utilidade da MD para a gestão.

## **CONCLUSÃO**

Variáveis redundantes e registros ausentes ou incompletos, reforçam a necessidade da sensibilização contínua dos profissionais sobre a importância dos registros corretos, bem como do trabalho interdisciplinar no desenvolvimento dos sistemas, a fim de evitar redundância de variáveis.

Pela identificação da diversidade dos problemas de saúde abordados, obtenção de resultados considerados compreensíveis, novos, interessantes e úteis, com a aproximação das variáveis às camadas mais externas do modelo dos DSS, em que reside a maior governabilidade do Estado, conclui-se que a MD possui potencial de contribuição à gestão em saúde materno-infantil, principalmente quando a interdisciplinaridade faz-se presente.

No entanto, notou-se que, dos resultados interessantes, a maioria dos desfechos estudados foi de doenças específicas (asma, cárie precoce na infância, leucomalácia periventricular etc.) ou de processos/serviços de saúde (consulta pré-natal, consulta puerperal, retorno à emergência pediátrica etc.), com poucos desfechos diretamente relacionados à garantia da vida. A ausência de mais resultados interessantes relacionados à mortalidade suscita a necessidade de discussão da inclusão de variáveis nos sistemas de informação em saúde além das já estabelecidas, o que possibilitaria novas descobertas de conhecimento.

Foi possível perceber a ausência de interdisciplinaridade principalmente por parte dos estudos publicados na área de informática, refletindo na obtenção de resultados sem potencial de contribuição à gestão. Ainda, identificou-se uma importante lacuna entre o desenvolvimento e uso da MD e a aplicação dos resultados obtidos, visto que todos os estudos foram ou de desenvolvimento tecnológico ou de execução, com experiências transversais e pontuais. A aproximação dessas áreas distintas gera benefícios tanto para a área-meio (computação) quanto para a área-fim (saúde), uma vez que os resultados produzidos pela computação serão úteis e válidos ao agregar ao que já se conhece e ser efetivamente aplicáveis, de forma a melhorar a saúde da população.

## AGRADECIMENTOS

Agradecemos ao Programa de Suporte à Pós-Graduação de Instituições Comunitárias de Educação Superior (PROSUC/CAPES), pelo financiamento, modalidade taxa, para a realização.

## REFERÊNCIAS

1. Black RE, Allen LH, Bhutta ZA, Caulfield LE, Onis M, Ezzati M, Mathers C, Rivera J. Maternal and child undernutrition: global and regional exposures and health consequences. *The Lancet*, 2008 Jan [cited 2018 Nov 26]; 371(9608):243-260. Available from: [https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736\(07\)61690-0.pdf](https://www.thelancet.com/pdfs/journals/lancet/PIIS0140-6736(07)61690-0.pdf).
2. Comissão para os Determinantes Sociais da Saúde [CDSS]. *Redução das desigualdades no período de uma geração: igualdade na saúde através da ação sobre os seus determinantes sociais: relatório final* [Internet]. 2010 [acessado 2018 Jun 01]. Disponível em: [http://www.who.int/eportuguese/publications/Reducao\\_desigualdades\\_relatorio2010.pdf](http://www.who.int/eportuguese/publications/Reducao_desigualdades_relatorio2010.pdf)
3. Black RE, Victora CG, Walker SP, Bhutta ZA, Christian P, de Onis M, Ezzati M, Grantham-McGregor S, Katz J, Martorell R, Uauy R. Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, 2013 Jun; 382(9890):427-451.
4. Cunha AJAD, Leite AJM, Almeida IS. The pediatrician's role in the first thousand days of the child: the pursuit of healthy nutrition and development. *J Pediatr*, 2015 Nov/Dec; 91(6 S1):S44-S51.



5. Victora GC, Horta BL, Mola CL, Quevedo L, Pinheiro RT, Gigante D, Gonçalves H, Barros FC. Association between breastfeeding and intelligence, educational attainment, and income at 30 years of age: a prospective birth cohort study from Brazil. *The Lancet Glob Health*; 2015 Apr; 3(4):e199-e205.
6. Mameli C, Mazzantini S, Zuccotti G. Nutrition in the first 1000 days: the origin of childhood obesity. *Int J Environ Res Public Health*; 2016 Sep; 13(9):838.
7. Moore, T et al. *The first thousand days: an evidence paper* [Internet]. 2017 Sep [cited 2018 Set 07]. Available from: <https://www.suicideinfo.ca/resource/the-first-thousand-days-an-evidence-paper/>.
8. Dahlgren G, Whitehead M. *Policies and strategies to promote social equity in health*. Stockholm: Institute for future studies; 1991.
9. Brasil, Ministério da Saúde, Secretaria de Atenção à Saúde, Departamento de Ações Programáticas Estratégicas. *Política Nacional de Atenção Integral à Saúde da Criança: orientações para implementação* [Internet]. Brasília: Ministério da Saúde; 2018 [acessado 2019 mar. 22]. Disponível em: [http://www.saude.pr.gov.br/arquivos/File/Politica\\_Nacional\\_de\\_Atencao\\_Integral\\_a\\_Saude\\_da\\_Crianca\\_PNAISC.pdf](http://www.saude.pr.gov.br/arquivos/File/Politica_Nacional_de_Atencao_Integral_a_Saude_da_Crianca_PNAISC.pdf).
10. Fayyad U, Shapiro G, Smyth P. From data mining to knowledge discovery in data base. *AI Magazines*; 1996; 17(3):37-54.
11. Goldschmidt R, Passos E, Bezerra E. *Data mining: conceitos, técnicas, algoritmos, orientações e aplicações*. 2. ed. Rio de Janeiro: Elsevier; 2015.
12. Carvalho DR, Escobar LFA, Almeida DT. Pontos de atenção para o uso da mineração de dados da saúde. *Informação & Informação*; 2014 Jan/Apr; 19(1):249-272.

13. Hu YH, Tai CT, Chen SCC, Lee HW, Sung SF. Predicting return visits to the emergency department for pediatric patients: applying supervised learning techniques to the Taiwan National Health Insurance Research Database. *Comput Methods Programs Biomed*; 2017 Jun; 144:105-112.
14. Toti G, Vilalta R, Lindner P, Lefer B, Macias C, Price D. Analysis of correlation between pediatric asthma exacerbation and exposure to pollutant mixtures with association rule mining. *Artif Intell Med*; 2016 Nov;74:44-52.
15. Ariyadasa SN, Munasinghe KL, Senanayake HDS, Fernando MGNAS. Knowledge extraction to mitigate child malnutrition in developing countries (Sri Lankan context). In: *Proceedings of the 4th International Conference on Intelligent Systems, Modelling and Simulation*; 2013 Jan 29-31; Bangkok, Thailand. New Jersey: IEEE; 2013.
16. Witten IH, Frank E. *Data mining practical machine learning tools and techniques*. 2nd. ed. San Francisco: Elsevier; 2005.
17. Moher D, Liberati A, Tetzlaff J, Altman DG. The Prisma Group (2009) preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*; 2009 Jul; 6(7):e1000097.
18. Landis J, Koch GG. The measurements of agreement for categorical data. *Biometrics*; 1977 Mar; 33(1):159-174.
19. Goodwin LK, Iannacchione MA, Hammond WE, Crockett P, Maher S, Schlitz K. Data mining methods find demographic predictors of preterm birth. *Nurs Res*; 2001 Nov/Dec; 50(6):340-345.
20. Silva D, Alahakoon D, Dharmage S. Cluster analysis using the GSOM: patterns in epidemiology. In: *Proceedings of the 3rd International Conference on Information and*

*Automation for Sustainability*; 2007 Dec 4-6; Melbourne, Australia. New Jersey: IEEE; 2007.

21. Orozova-Bekkevold I, Jensen H, Stensballe L, Olsen J. Maternal vaccination and preterm birth: using data mining as a screening tool. *Pharm World Sci*; 2007 Jun; 29(3): 205-212.

22. Vianna RC, Moro CM, Moysés SJ, Carvalho D, Nievola JC. Data mining and characteristics of infant mortality. *Cad Saude Publica*; 2010 Mar; 26(3):535-542.

23. Loucoubar C, Paul R, Bar-Hen A, Huret A, Tall A, Sokhna C, et al. An exhaustive, non-euclidean, non-parametric data mining tool for unraveling the complexity of biological systems--novel insights into malaria. *PLoS One*; 2011 Sep; 6(9):e24085.

24. Matharage S, Alahakoon O, Alahakoon D, Kapurubandara S, Nayyar R, Mukherji M et al. Analysing stillbirth data using dynamic self organizing maps. In: *Proceedings of the 22nd International Workshop on Database and Expert Systems Applications*; 2011 29 Aug.-2 Sep; Toulouse, France. Berlin: Springer; 2011.

25. Chen HY, Chuang CH, Yang YJ, Wu TP. Exploring the risk factors of preterm birth using data mining. *Expert Syst Appl*; 2011 May; 38(5):5384-5387.

26. Belle V, Lisboa P. Automated selection of interaction effects in sparse kernel methods to predict pregnancy viability. In: *Proceedings of the 2013 IEEE Symposium on Computational Intelligence and Data Mining*; 2013 Apr 16-19; Singapore, Singapore. New Jersey: IEEE; 2013.

27. Kolyshkina I, Brownlow M, Taylor J. Improving every child's chance in life. In: *Proceedings of the 13th International Conference on Data Mining Workshops*; 2013 Dec 7-10; Dallas, USA. New Jersey: IEEE; 2013.

28. Zacarias O, Bostrom H. Strengthening the health information system in Mozambique through malaria incidence prediction. In: *Proceedings of the 2013 IST-*

*Africa Conference & Exhibition*; 2013 May 29-31; Nairobi, Kenya. New Jersey: IEEE; 2013.

29. Windal F, Jeribi K, Ficheur G, Degoul S, Martinot A, Beuscart R, Renard JM. Pediatric emergency department crowding: survival tree clustering for length of patient stay. *Stud Health Technol Inform*; 2014; 205:1095-1099.

30. Jalali A, Buckley EM, Lynch JM, Schwab PJ, Licht DJ, Nataraj C. Prediction of periventricular leukomalacia occurrence in neonates after heart surgery. *IEEE J Biomed Health Inform*; 2014 Jul; 18(4):1453-1460.

31. Huff MD, Weisman D, Adams J, Li S, Green J, Malone LL, Clemmons S. The frequency of tetracycline resistance genes co-detected with respiratory pathogens: a database mining study uncovering descriptive trends throughout the United States. *BMC Infect Dis*; 2014 Aug 25; 14:460.

32. Jawad F, Choudhury TUR, Najeeb A, Faisal M, Nusrat F, Shamita RC et al. Data mining techniques to analyze the reason for home birth in Bangladesh. In: *Proceedings of the 16th International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing*; 2016 June 1-3; Takamatsu, Japan. New Jersey: IEEE; 2016.

33. Le Meur N, Gao F, Bayat S. Mining care trajectories using health administrative information systems: the use of state sequence analysis to assess disparities in prenatal care consumption. *BMC Health Serv Res*; 2015 May 15; 15:200.

34. Ivancevic V, Tusek I, Tusek J, Knezevic M, Elheshk S, Lukovic I. Using association rule mining to identify risk factors for early childhood caries. *Comput Methods Programs Biomed*; 2015 Nov; 122(2):175-181.

35. Sundararaman A, Ramanathan SV. A novel approach to estimate maternal mortality rate and its determinants by statistical modelling. In: *Proceedings of the*

*32nd International Conference on Data Engineering Workshops*; 2016 May 16-20; Helsinki, Finland. New Jersey: IEEE; 2016.

36. Gupta S, Singh SN, Kumar D. An empirical analysis of maternal health data: a case study of India. In: *Proceedings of the 2nd International Conference on Next Generation Computing Technologies*; 2016 Oct 14-16; Dehradun, India. New Jersey: IEEE; 2016.

37. Mai MV, Krauthammer M. Controlling testing volume for respiratory viruses using machine learning and text mining. *AMIA Annu Symp Proc*; 2017 Feb 10; 2016:1910-1919.

38. Sahle G. Ethiopic maternal care data mining: discovering the factors that affect postnatal care visit in Ethiopia. *Health Inf Sci Syst*; 2016 Dec; 4:4.

39. Elrazek A, Amer M, El-Hawary B, Salah A, Bhagavathula AS, Alborai M, Saab S. Prediction of HCV vertical transmission: what factors should be optimized using data miningcomputational analysis. *Liver Int*; 2017 Apr; 37(4):529-533.

40. Anilkumar NA, Gupta D, Khare S, Gopalkrishna DM, Jyotishi A. Characteristics and causes of malnutrition across Indian states: a cluster analysis based on Indian demographic and health survey data. In: *Proceedings of the 2017 International Conference on Advances in Computing, Communications and Informatics*; 2017 Sep 13-16; Udupi, India. New Jersey: IEEE; 2017.

41. Long E, Xu S, Liu Z, Wu X, Zhang X, Wang J et al. Construction and implications of structural equation modeling network for pediatric cataract: a data mining research of rare diseases. *BMC Ophthalmol*; 2017 May 19; 17(1):74.

42. Tesfaye B, Atique S, Elias N, Dibaba L, Shabbir SA, Kebede M. Determinants and development of a web-based child mortality prediction model in resource-limited

settings: a data mining approach. *Comput Methods Programs Biomed*; 2017 Mar; 140:45-51.

43. Sartorelli AP, Gomes DC, Cubas MR, Carvalho DR. Fatores que contribuem para a mortalidade infantil utilizando a mineração de dados. *Saude e Pesqui*; 2017 Jan/Abr; 10(1):33-41.

44. Chen Y, Kho AN, Liebovitz D, Ivory C, Osmundson S, Bian J, Malin BA. Learning bundled care opportunities from electronic medical records. *J Biomed Inform*; 2018 Jan; 77:1-10.

45. World Health Organization. *Survive and thrive: transforming care for every small and sick newborn. Key findings*. Geneva: World Health Organization; 2018 [cited 2019 Jan 30]. Available from: <https://apps.who.int/iris/bitstream/handle/10665/276655/WHO-FWC-MCA-18.11-eng.pdf?ua=1>.

46. Piatetsky-Shapiro G. Knowledge discovery in real databases: a report on the IJCAI-89 Workshop. *AI Magazines*; 1990 [cited 2018 Dec 20]; 11(5):37-54. Available from: <https://www.aaai.org/ojs/index.php/aimagazine/article/view/873/791>.

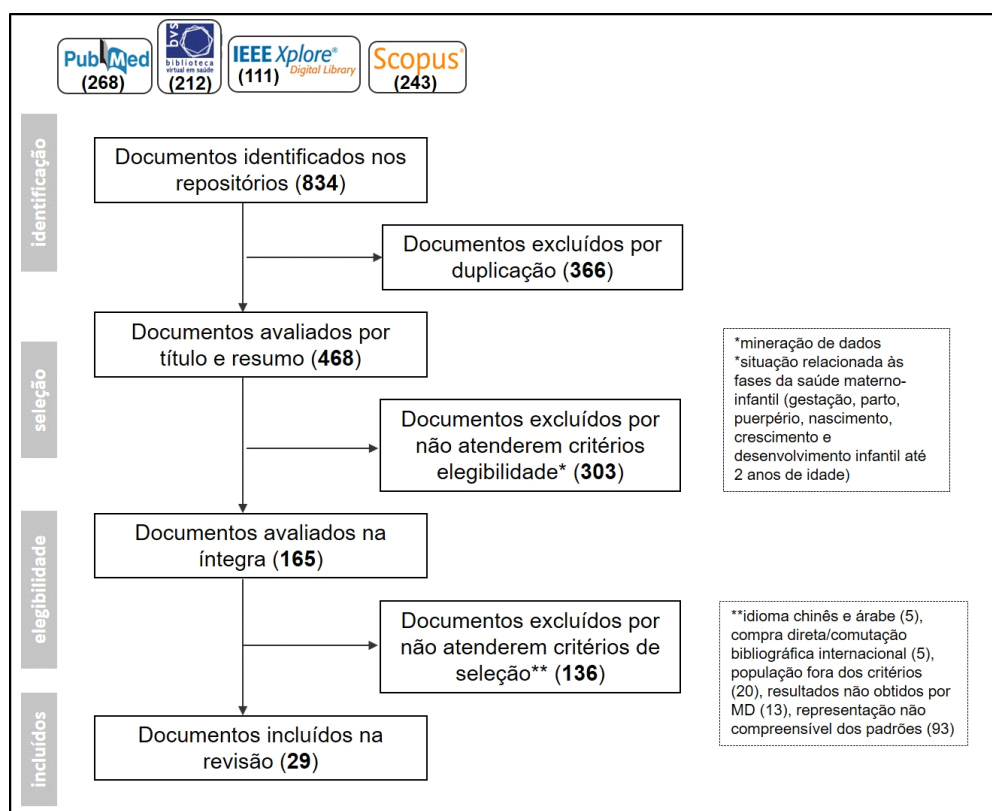
47. Liu H, Motoda H. *Feature selection for knowledge discovery and data mining*. Berlin: Springer; 2012.

48. Carvalho DR, Dallagassa M. Mineração de dados: aplicações, ferramentas, tipos de aprendizado e outros subtemas. *AtoZ Novas Práticas em Informação e Conhecimento*; 2014 Jul/Dez [acessado 2019 Mar 19]; 3(2):82-86. Disponível em: <http://www.atoz.ufpr.br>.

49. Sartorelli AP, Gomes DC, Cubas MR, Carvalho DR. Fatores que contribuem para mortalidade infantil utilizando a mineração de dados. *Rev Saude Pesq*; 2017 Jan/Abr; 10(1):33-41.

50. Neves BC, Braz MI. Interlocação entre saúde e ciência da informação: proposta para o diagrama multidisciplinar da Ci. *Inf Inf*, 2018 Set/Dez [acessado 2019 Mar 28]; 23(3):100-21. Disponível em: <http://www.uel.br/revistas/uel/index.php/informacao/article/view/26431/pdf>.

Figura 1 – Etapas da revisão



Quadro 1 – Características extraídas dos estudos incluídos, respectiva descrição e finalidade

Característica	Descrição	Finalidade
Título/autor/ano de publicação	Identificação do documento.	Inferir a atualidade do tema.
Área de publicação	Área de publicação: saúde, informática ou interdisciplinar.	Identificar a existência de interdisciplinaridade.
Desfecho	Situação ou problema de saúde objeto-alvo do estudo.	Verificar em quais situações tem sido utilizada a MD.
Temporalidade da base de dados	Quantidade de anos da base de dados.	Dimensionar o tamanho da base de dados e inferir se ele influencia a descoberta de novos conhecimentos.
Quantidade de registros	Número de registros da base de dados.	
Quantidade de variáveis	Número de variáveis da base de dados.	
Aderência à literatura	Associações encontradas corroboram, divergem ou são interessantes frente à literatura.	Verificar a contribuição da adoção de MD para a descoberta de conhecimentos novos e válidos.
DSS	Aproximação dos fatores de exposição associados ao desfecho com as camadas do modelo dos DSS.	Identificar a governabilidade e nível de intervenção necessário pelo Estado, de maneira a inferir a utilidade para a gestão, do conhecimento descoberto.



Quadro 2 – Estudos incluídos na revisão e características extraídas.

<b>Autor/ano</b>	<b>Área de publicação</b>	<b>Desfecho</b>	<b>Base de dados</b>	<b>Resultados - aderência à literatura e aproximação à camada do modelo dos DSS</b>
Goodwin et al., 2001 <sup>19</sup>	Saúde	Prematuridade	20 anos 19.970 registros 1.622 variáveis	CORROBORAÇÃO: individual, intermediário INTERESSANTE: individual
Silva et al., 2007 <sup>20</sup>	Informática	Eczema infantil	12,5 anos 620 registros Quant. variáveis não informado	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Orozova-Bekkevold et al., 2007 <sup>21</sup>	Saúde	Prematuridade	6 anos 92.235 registros 281 variáveis	INTERESSANTE: intermediário
Vianna et al., 2010 <sup>22</sup>	Saúde	Mortalidade infantil	5 anos 7.256 registros Quant. variáveis não informado	CORROBORAÇÃO: individual, intermediário
Loucoubar, et al., 2011 <sup>23</sup>	Saúde	Malária	18 anos 46.837 registros 34 variáveis	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Matharage et al., 2011 <sup>24</sup>	Informática	Natimortalidade	6 anos 215 registros Quant. variáveis não informado	CORROBORAÇÃO: individual INTERESSANTE: individual
Chen et al., 2011 <sup>25</sup>	Informática	Prematuridade	4 anos 910 registros Quant. variáveis não informado	CORROBORAÇÃO: individual, estilo de vida, intermediário
Belle et al., 2013 <sup>26</sup>	Informática	Aborto	Quant. anos não informado 1.435 registros Quant. variáveis não informado	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Kolyshkina et al., 2013 <sup>27</sup>	Informática	Vulnerabilidade infantil	6 anos Quant. registros não informado 45 variáveis	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Ariyadasa et al., 2013 <sup>16</sup>	Informática	Desnutrição infantil	2 anos 6.714 registros Quant. variáveis não informado	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Zacarias et al., 2013 <sup>28</sup>	Informática	Malária	3 anos Quant. registros não informado Quant. variáveis não informado	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Winald et al., 2014 <sup>29</sup>	Interdisciplinar	Duração de estadia em emergência pediátrica	3 anos 55.183 registros Quant. variáveis não informado	CORROBORAÇÃO: intermediário
Jalali et al., 2014 <sup>30</sup>	Interdisciplinar	Leucomalácia periventricular após cirurgia cardíaca	Quant. anos não informado 44 registros 36 variáveis	DIVERGÊNCIA: individual INTERESSANTE: intermediário

Huff et al., 2014 <sup>31</sup>	Saúde	Resistência a tetraciclina	3,8 anos 80.241 registros Quant. variáveis não informado	CORROBORAÇÃO: individual, intermediário INTERESSANTE: sociais e ambientais gerais
Jawad et al., 2015 <sup>32</sup>	Informática	Parto domiciliar	0,8 anos 5.248 registros 549 variáveis	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Le Meur et al., 2015 <sup>33</sup>	Saúde	Pré-natal	1 ano 2.518 registros Quant. variáveis não informado	CORROBORAÇÃO: individual, intermediário INTERESSANTE: intermediário
Ivancevic et al., 2015 <sup>34</sup>	Interdisciplinar	Cárie precoce da infância	Quant. anos não informado 341 registros 35 variáveis	CORROBORAÇÃO: individual, estilo de vida, intermediário INTERESSANTE: sociais e ambientais gerais
Sundararaman et al., 2016 <sup>35</sup>	Informática	Mortalidade materna	7 anos 2.688 registros 27 variáveis	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Gupta et al., 2016 <sup>36</sup>	Informática	Mortalidade materna	2 anos Quant. registros não informado 9 variáveis	CORROBORAÇÃO: individual, intermediário
Toti et al., 2016 <sup>15</sup>	Interdisciplinar	Asma	10 anos 20.959 registros Quant. variáveis não informado	CORROBORAÇÃO: ambientais gerais DIVERGÊNCIA: ambientais gerais INTERESSANTE: ambientais gerais
Mai et al., 2016 <sup>37</sup>	Interdisciplinar	<i>Overtesting</i>	3 anos 1.685 registros 400 variáveis	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Sahle, 2016 <sup>38</sup>	Interdisciplinar	Consulta puerperal	1 anos 6.558 registros Quant. variáveis não informado	INTERESSANTE: intermediário, sociais e ambientais gerais
Elrazek et al., 2016 <sup>39</sup>	Saúde	Hepatite C	1 ano 3.000 registros Quant. variáveis não informado	CORROBORAÇÃO: individual DIVERGÊNCIA: individual
Anilkumar et al., 2017 <sup>40</sup>	Informática	Desnutrição infantil	2 anos 41.304 registros 401 variáveis	Sem análise/discussão frente à literatura, enfoque na acurácia do modelo proposto
Long et al., 2017 <sup>41</sup>	Saúde	Catarata pediátrica	Quant. anos não informado 160 registros 15 variáveis	CORROBORAÇÃO: individual
Tesfaye et al., 2017 <sup>42</sup>	Interdisciplinar	Mortalidade na infância	1 ano 11.654 registros 21 variáveis	CORROBORAÇÃO: individual, estilo de vida, redes sociais, intermediário
Sartorelli, et al., 2017 <sup>43</sup>	Saúde	Mortalidade infantil	5 anos 266 registros 15 variáveis	CORROBORAÇÃO: individual, intermediário
Hu et al., 2017 <sup>14</sup>	Interdisciplinar	Retorno à emergência pediátrica	12 anos 125.940 registros Quant. variáveis não informado	CORROBORAÇÃO: individual, intermediário DIVERGÊNCIA: individual INTERESSANTE: intermediário

Chen et al., 2018 <sup>44</sup>	Interdisciplinar	Otimização de cuidados	0,3 anos 1.138.317 registros Quant. variáveis não informado	CORROBORAÇÃO: individual
------------------------------------	------------------	---------------------------	---	--------------------------

**APÊNDICE C – VARIÁVEIS UTILIZADAS NOS EXPERIMENTOS DE KDD, COM RESPECTIVA DESCRIÇÃO, VALORES, CAMADA APROXIMADA DOS DSS E EXPERIMENTO UTILIZADO, POR SIS**

Quadro 7 – Variáveis do SIM utilizadas nos experimentos de KDD, com respectiva descrição, valores, camada aproximada dos DSS e experimento utilizado

variável (rótulo atribuído)	descrição variável	Valores	camada DSS	Variáveis utilizadas no Experimento		
				1	2 e 3	4
ano do óbito	Ano de ocorrência do óbito. Derivado a partir da data do óbito (variável DTOBITO, excluída no pré-processamento)	Ano com quatro dígitos no formato aaaa.	-	.	✓	.
componente do óbito	Período conforme idade da criança no momento do óbito. Criada a partir da idade	NP = até 6 dias (inclusive). NT = 7-27 dias (inclusive). PN = 28d a 11meses (inclusive). 1-2 anos = 1 a 2 ans de idade.	idade, sexo e fatores hereditários	.	✓	.
idade da criança em meses completos	Idade da criança em meses no momento do óbito. Derivado a partir da diferença entre a data de nascimento e data do óbito. Calculado pela fórmula excel =DataDif(dtnasc;dtobito;"m").	0 a 24	idade, sexo e fatores hereditários	.	✓	.
idade da criança em dias	Idade da criança em dias no momento do óbito. Criada a partir da diferença entre a data de nascimento e data do óbito. Calculado pela fórmula excel =DataDif(dtnasc;dtobito;"d").	0 a 730	idade, sexo e fatores hereditários	.	✓	.
local ocorrência óbito	Local de ocorrência do óbito.	1 – hospital; 2 – outros estabelecimentos de saúde; 3 – domicílio; 4 – via pública; 5 – outros; 9 – ignorado.	cond. vida e trabalho	.	✓	.
CBO mãe no SIM	Código de ocupação da mãe.	Conforme tabela do CBO (Código Brasileiro de Ocupações).	cond. vida e trabalho	.	✓	.
óbito em relação ao parto	Informação como foi a morte em relação ao parto.	1 – antes; 2 – durante; 3 – depois; 9 – Ignorado.	idade, sexo e fatores hereditários	.	✓	.
assistência médica no óbito	Informação sobre assistência médica.	1 – sim; 2 – não; 9 – ignorado.	cond. vida e trabalho	.	✓	.
causa óbito Linha A	CIDs informados na Linha A da DO.	Lista CID-10.	idade, sexo e fatores hereditários cond. vida e trabalho cond. socioeconom, culturais e ambientais gerais	.	✓	.
causa óbito Linha B	CIDs informados na Linha B da DO.	Lista CID-10.	idade, sexo e fatores hereditários cond. vida e trabalho cond. socioeconom, culturais e ambientais gerais	.	✓	.
causa óbito Linha C	CIDs informados na Linha C da DO.	Lista CID-10.	idade, sexo e fatores hereditários	.	✓	.

			cond. vida e trabalho cond. socioeconom, culturais e ambientais gerais			
causa óbito Linha D	CIDs informados na Linha D da DO.	Lista CID-10.	idade, sexo e fatores hereditários cond. vida e trabalho cond. socioeconom, culturais e ambientais gerais	.	✓	.
causa óbito Linha II	CIDs informados na parte II da DO.	Lista CID-10.	idade, sexo e fatores hereditários cond. vida e trabalho cond. socioeconom, culturais e ambientais gerais	.	✓	.
causa básica óbito	Causa básica do óbito.	Lista CID-10.	idade, sexo e fatores hereditários cond. vida e trabalho cond. socioeconom, culturais e ambientais gerais	.	✓	.
evitabilidade do óbito	Classificação do óbito de acordo com a Classificação de Evitabilidade da Fundação SEADE. Criada a partir da variável CAUSABAS.	RI - Redutíveis por imunoprevenção. RACG - Redutíveis por adequado controle a gravidez. RAAP - Redutíveis por adequada atenção ao parto. RAPDTP - Redutíveis por ações de prevenção, diagnóstico e tratamento precoces. RPOS - Redutíveis através de parcerias com outros setores. NE - Não evitáveis. MD - Mal definidas. NCA - Não classificadas/outras. NO - Não óbito.	idade, sexo e fatores hereditários cond. vida e trabalho cond. socioeconom, culturais e ambientais gerais	✓	✓	✓
circunstância óbito não natural	Indicação de provável circunstância de morte quando não natural.	1 – acidente; 2 – suicídio; 3 – homicídio; 4 – outros; 9 – ignorado.	cond. vida e trabalho cond. socioeconom, culturais e ambientais gerais	.	✓	.

Fonte: A autora (2020).

Quadro 8 – Variáveis do SINASC utilizadas nos experimentos de KDD, com respectiva descrição, valores, camada aproximada dos DSS e experimento utilizado

variável (rótulo atribuído)	descrição variável	Valores	DSS	Variáveis utilizadas no Experimento		
				1	2 e 3	4
ano nascimento	Ano do nascimento, Criada da data de nascimento (variável DTNASC, excluída no pré-processamento).	Ano com quatro dígitos no formato aaaa.	-	✓	✓	.
vinculação SUS	Estabelecimento conforme assistência prestada pelo SUS. Criada a partir do código CNES do estabelecimento de saúde (variável CODESTAB).	1 - Fora Ctba. 2 - ignorado. 3 - não SUS. 4 - SUS.	cond. vida e trabalho	✓	✓	✓
código município nascimento	Município de ocorrência do nascimento.	Valores conforme código IBGE de municípios. Curitiba = 410690.	cond. socioeconom, culturais e ambientais gerais	✓	✓	.
nascimento Curitiba	criado. Nascimento em Curitiba.	CTBASIM - nascimento em Curitiba. CTBANAO - nascimento fora de Curitiba.	cond. socioeconom, culturais e ambientais gerais	✓	✓	.
local nascimento	Local de nascimento.	1 - Hospital; 2 – Outros estabelecimentos de saúde; 3 – Domicílio; 4 – Outros.	cond. vida e trabalho	✓	✓	✓

idade materna	Idade da mãe.	Idade em anos, formato aa.	idade, sexo e fatores hereditários	✓	✓	.
estado civil mãe	Situação conjugal da mãe.	1– Solteira; 2– Casada; 3– Viúva; 4– Separada judicialmente/divorciada; 5– União estável; 9– Ignorada.	redes sociais e comunitárias	✓	✓	.
escolaridade mãe em anos concluídos	Escolaridade da mãe, em anos de estudo concluídos.	1 –Nenhuma; 2 – 1 a 3 anos; 3 – 4 a 7 anos; 4 – 8 a 11 anos; 5 – 12 e mais; 9 – Ignorado.	cond. vida e trabalho	✓	✓	.
CBO mãe no SINASC	Código de ocupação da mãe.	Conforme tabela do CBO (Código Brasileiro de Ocupações).	cond. vida e trabalho	✓	✓	.
no. filhos nascidos vivos	Número de filhos vivos.	Número absoluto, no formato nn.	idade, sexo e fatores hereditários	✓	✓	.
no. perda fetal/aborto	Número de perdas fetais e abortos.	Número absoluto, no formato nn.	idade, sexo e fatores hereditários	✓	✓	.
bairro de residência	Bairro de residência.	Descrição do bairro conforme 2 últimos dígitos do código do bairro (variável CODBAIRES, três dígitos, excluída no pré-processamento)	cond. socioeconom, culturais e ambientais gerais	✓	✓	.
classificação de risco ao nascer	Classificação de risco realizada manualmente conforme critérios utilizados pela SMS Curitiba. Criada do código do bairro (variável CODBAIRES, excluída no pré-processamento).	CODBAIRES possui três dígitos. O primeiro dígito é relacionado ao risco: quando número PAR = não risco, quando ÍMPAR = risco.	idade, sexo e fatores hereditários cond. vida e trabalho	✓	✓	✓
DS de residência	Código DS de residência.	1 - SF, 2 - BV, 3 - BQ, 4 - PR, 5 - PN, 6 - CJ, 7 - MZ, 8 - BN, 9 - CIC, 10 - TQ.	cond. vida e trabalho	✓	✓	✓
duração gestação categorizada	Duração da gestação em semanas, categorizada.	1–Menos de 22 semanas; 2–22 a 27 semanas; 3– 28 a 31 semanas; 4– 32 a 36 semanas; 5– 37 a 41 semanas; 6– 42 semanas e mais; 9– Ignorado.	idade, sexo e fatores hereditários	✓	✓	.
gravidez	Tipo de gravidez.	1–Única; 2– Dupla; 3– Tripla ou mais; 9– Ignorado.	idade, sexo e fatores hereditários	✓	✓	.
tipo de parto	Tipo de parto.	1– Vaginal; 2– Cesário; 9– Ignorado.	cond. vida e trabalho	✓	✓	✓
consultas pré-natal categorizada	Número de consultas de pré-natal CATEGORIZADA.	1–Nenhuma; 2–de 1 a 3; 3– de 4 a 6; 4– 7 e mais; 9– Ignorado.	cond. vida e trabalho	✓	✓	✓
dia da semana nascimento	Dia da semana do nascimento. Derivada da data de nascimento (DTNASC, variável excluída no pré-processamento).	1-Domingo. 2-Segunda. 3-Terça. 4-Quarta. 5-Quinta. 6-Sexta. 7-Sábado.	-	✓	✓	.
turno nascimento	Turno de nascimento, considerando turno de hospital. Criada da hora do nascimento (variável HORANASC, excluída no pré-processamento).	7 às 12 - manhã. 14 às 18 - tarde. 19 às 6 - noite.	cond. vida e trabalho	✓	✓	✓
sexo	Sexo do nascido vivo.	M – Masculino; F – Feminino; I – ignorado.	idade, sexo e fatores hereditários	✓	✓	.
apgar 1º minuto	Apgar no 1º minuto.	Valores de 0 a 10, conforme escala de Apgar.	idade, sexo e fatores hereditários	✓	✓	.
apgar 5º minuto	Apgar no 5º minuto.	Valores de 0 a 10, conforme escala de Apgar.	cond. vida e trabalho	✓	✓	✓

recuperação apgar	Variação do Apgar no 1º e 5º minuto. Criada da diferença entre APGAR5 e APGAR1.	Valores de 0 a 10.	cond. vida e trabalho	✓	✓	.
apgar no 1º e 5º minuto	Recuperação de Apgar do 1º ao 5º minuto. Criada da junção de APGAR1 e APGAR5	Expressão aa_a'a', onde aa = valor do Apgar no 1º minuto e a'a' = valor do Apgar no 5º minuto.	cond. vida e trabalho	✓	✓	.
raça cor nascido	Tipo de raça e cor do nascido.	1-branca, 2-preta, 3-amarela, 4-par-da, 5-indígena.	idade, sexo e fatores hereditários	✓	✓	.
peso ao nascer	Peso ao nascer em gramas.	Valor numérico de 0 a 9999.	idade, sexo e fatores hereditários	✓	✓	.
malformação congênita	Identificação de anomalia identificada.	1– Sim; 2– Não; 9– Ignorado.	idade, sexo e fatores hereditários	✓	✓	.
código anomalia	Código da anomalia.	Conforme CID 10.	idade, sexo e fatores hereditários	✓	✓	.
escolaridade mãe ensino concluído	Escolaridade da mãe, classificação 2010.	0 – Sem escolaridade; 1 – Fundamental I (1ª a 4ª série); 2 – Fundamental II (5ª a 8ª série); 3 – Médio (antigo 2º Grau); 4 – Superior incompleto; 5 – Superior completo; 9 – Ignorado.	cond. vida e trabalho	✓	✓	.
escolaridade mãe série escolar concluída	Série escolar da mãe.	Valores de 1 a 8.	cond. vida e trabalho	✓	✓	.
raça cor mãe	Tipo de raça e cor da mãe.	1– Branca; 2– Preta; 3– Amarela; 4– Padra; 5-Indígena.	idade, sexo e fatores hereditários	✓	✓	.
nº gestação anterior	Número de gestações anteriores	Número absoluto, no formato nn.	idade, sexo e fatores hereditários	✓	✓	.
partos normais anteriores	Número de partos normais	Número absoluto, no formato nn.	cond. vida e trabalho	✓	✓	✓
partos cesáreos anteriores	Numero de partos cesáreos	Número absoluto, no formato nn.	cond. vida e trabalho	✓	✓	✓
idade pai	Idade do pai.	Idade do pai, no formato aa.	idade, sexo e fatores hereditários	✓	✓	.
duração gestação numérica	Número de semanas de gestação.	Número absoluto, no formato nn.	idade, sexo e fatores hereditários	✓	✓	.
método estimacão semanas gestação	Método utilizado para estimar semanas de gestação.	1-exame físico, 2-outro método, 9-ignorado.	cond. vida e trabalho	✓	✓	✓
classificação Intergrowth	Classificação do nascido vivo conforme percentil dos valores de referência Intergrowth. Criada a partir de peso (variável PESO) e idade gestacional (variável SEMAGESTAC). Para semana de gestação = 22 e 23, utilizado tabela de peso fetal. Tabela Intergrowth não abrange semanas de gestação = 19, 20, 21, 43, 44 e 45, que não foram classificados (em branco).	Percentil <10 = PIG (pequeno para a idade gestacional). Percentil entre 10-90 = AIG (adequado para a idade gestacional). Percentil >90 = GIG (grande para a idade gestacional).	idade, sexo e fatores hereditários	✓	✓	.
classificação Fenton	Classificação do nascido vivo conforme percentil dos	Percentil <10 = PIG (pequeno para a idade gestacional).	idade, sexo e fatores hereditários	✓	✓	.

	valores de referência Fenton. Criada a partir de peso (variável PESO) e idade gestacional (variável SEMAGESTAC). Utilizado ferramenta disponível no site <a href="https://www.ucalgary.ca/fenton/2013chart">https://www.ucalgary.ca/fenton/2013chart</a> . [tools - 2. Research Bulk Calculator]. Fenton não abrange semanas <22 semanas, que não foram classificados (em branco).	Percentil entre 10-90 = AIG (adequado para a idade gestacional). Percentil >90 = GIG (grande para a idade gestacional).				
consultas pré-natal nº absoluto	Número de consultas de pré-natal.	Número absoluto, no formato nn.	cond. vida e trabalho	✓	✓	.
mês início pré-natal	Mês de gestação em que iniciou o pré-natal	Mês no formato mm.	cond. vida e trabalho	✓	✓	✓
apresentação nascido vivo	Tipo de apresentação do RN.	1– Cefálico; 2–Pélvica ou podálica; 3– Transversa; 9– Ignorado.	idade, sexo e fatores hereditários	✓	✓	.
trabalho de parto induzido	Indução do trabalho de parto.	1– Sim; 2–Não; 3–Não se aplica; 9– Ignorado.	cond. vida e trabalho	✓	✓	✓
cesárea antes do trabalho de parto	Ocorrência da cesárea antes do início de trabalho de parto.	1– Sim; 2– Não; 3– Não se aplica; 9– Ignorado.	cond. vida e trabalho	✓	✓	✓
assistência profissional	Profissional que assistiu o nascimento.	1– Médico; 2– Enfermeira/obstetriz; 3– Parteira; 4– Outros; 9– Ignorado.	cond. vida e trabalho	✓	✓	✓
escolaridade mãe ensino concluído fragmentado	Escolaridade 2010 agregada.	00 – Sem Escolaridade; 01 – Fundamental I Incompleto; 02 – Fundamental I Completo; 03 – Fundamental II Incompleto; 04 – Fundamental II Completo; 05 – Ensino Médio Incompleto; 06 – Ensino Médio Completo; 07 – Superior Incompleto; 08 – Superior Completo; 09 – Ignorado; 10 – Fundamental I Incompleto ou Inespecífico; 11 – Fundamental II Incompleto ou Inespecífico; 12 – Ensino Médio Incompleto ou Inespecífico.	cond. vida e trabalho	✓	✓	.



grupo de Robson	Código do Grupo de Robson, gerado pelo sistema.	<p>1 - Nulíparas, ≥37 sem., feto único, cefálico, trabalho de parto espontâneo; 2 - Nulíparas, ≥37 sem., feto único, cefálico, parto induzido ou cesariana programada; 3 - Multipara (sem cesariana anterior), ≥37 sem., feto único, cefálico, trabalho de parto espontâneo; 4 - Multipara (sem cesariana anterior), ≥37 sem., feto único, cefálico, parto induzido ou cesariana programada; 5 - Multipara, ≥37 sem., feto único, cefálico, cesariana anterior; 6 - Nulíparas com feto em apresentação pélvica, qualquer idade gestacional; 7 - Multiparas com feto em apresentação pélvica, qualquer idade gestacional (inclui cesariana prévia); 8 - Gravidez múltipla (qualquer paridade, qualquer idade gestacional, inclui cesariana prévia); 9 - Feto em apresentação transversa/obliqua (qualquer paridade, qualquer idade gestacional, inclui cesariana prévia); 10 - Gravidez de &lt; 37 semanas, feto único, cefálico (qualquer paridade, inclui cesariana anterior)</p>	idade, sexo e fatores hereditários cond. vida e trabalho	✓	✓	✓
-----------------	---	--	--	---	---	---

Fonte: A autora (2020).

## APÊNDICE D – VARIÁVEIS NÃO UTILIZADAS NOS EXPERIMENTOS DE KDD, COM RESPECTIVA DESCRIÇÃO E JUSTIFICATIVA, POR SIS

Quadro 9 – Variáveis SIM não utilizadas nos experimentos de KDD, com respectiva descrição e justificativa.

variável (rótulo original)	descrição	justificativa
NUMERODV	Dígito verificador da DO	sem utilidade
CODESTCART	Código da UF do cartório	sem utilidade (cartório)
CODMUNCART	Código do município do cartório	sem utilidade (cartório)
CODCART	Código do cartório	sem utilidade (cartório)
NUMREGCART	Código de estabelecimento do cartório	sem utilidade (cartório)
DTREGCART	Data de registro no Cartório	sem utilidade (cartório)
TIPOBITO	Tipo óbito. 1 - fetal. 2 - não fetal	todos são 2 - não fetal
NUMSUS	número do cartão SUS	dado de identificação
NATURAL	naturalidade	código não há informação do que significa
CODMUNNATU	Código do município de naturalidade do falecido	todos registros em branco
NOME	Nome do falecido	dado de identificação
NOMEPAI	Nome do pai do falecido	dado de identificação
NOMEMAE	Nome da mãe do falecido	dado de identificação
ESTCIV	Situação conjugal: 1 – Solteiro; 2 – Casado; 3 – Viúvo; 4 – Separado, judicialmente/divorciado; 5 – União estável; 9 – Ignorado.	todos 1 (solteiro) ou 9 (ignorado). O que importa é estado civil da mãe, que está no SINASC
ESC	Escolaridade em anos. Valores: 1 – Nenhuma; 2 – de 1 a 3 anos; 3 – de 4 a 7 anos; 4 – de 8 a 11 anos; 5 – 12 anos e mais; 9 – Ignorado.	todos 1 (sem escolaridade), 1 com registro 2 (de 1 a 3 anos). O que importa é a escolaridade da mãe, que está no SINASC
ESC2010	Escolaridade 2010. Valores: 0 – Sem escolaridade; 1 – Fundamental I (1ª a 4ª série); 2 – Fundamental II (5ª a 8ª série); 3 – Médio (antigo 2º Grau); 4 – Superior incompleto; 5 – Superior completo; 9 – Ignorado.	todos os registros em branco
SERIESCFAL	Série escolar do falecido. Valores de 1 a 8.	todos os registros em branco
OCUP	Ocupação habitual e ramo de atividade	todos os registros em branco
CODESTRES	Código da UF de residência	todos registros 41
CODMUNRES	Código do município de residência	todos registros residência Curitiba
CODBAIRES	código do bairro de residência	redundante (já possui nome)
CODENDRES	código do endereço de residência	dado de identificação
ENDRES	endereço de residência	dado de identificação
CODREGRES	código da região de residência	todos com mesmo valor
NUMRES	número da residência	dado de identificação
COMPLRES	complemento do endereço de residência	dado de identificação
CEPRES	cep da residência	dado de identificação
CODPAISRES	código do país de residência	todos registros Brasil
CODESTOCOR	código de estabelecimento de ocorrência	trabalhar com código do estabelecimento
CODMUNOCOR	código município de ocorrência	trabalhar com código do estabelecimento
BAIOCOR	bairro de ocorrência	sem padrão
CODBAIOCOR	código do bairro de ocorrência	trabalhar com código do estabelecimento
ENDOCOR	endereço de ocorrência	trabalhar com código do estabelecimento
CODENDOCOR	código do endereço de ocorrência	trabalhar com código do estabelecimento
CODREGOCOR	código da região de ocorrência	trabalhar com código do estabelecimento
NUMENDOCOR	número do endereço de ocorrência	trabalhar com código do estabelecimento
COMPLOCOR	complemento do endereço de ocorrência	trabalhar com código do estabelecimento
CEPOCOR	cep do endereço de ocorrência	trabalhar com código do estabelecimento
CODDISOCOR	código do distrito da ocorrência	trabalhar com local e código do estabelecimento
ESCMAE2010	Escolaridade 2010. Valores: 0 – Sem escolaridade; 1 – Fundamental I (1ª a 4ª série); 2 – Fundamental II (5ª a 8ª série); 3 – Médio (antigo 2º Grau); 4 – Superior incompleto; 5 – Superior completo; 9 – Ignorado.	todos registros em branco

SERIESCMAE	Série escolar da mãe. Valores de 1 a 8.	todos registros em branco
SEMAGESTAC	Semanas de gestação	todos os registros em branco
TPMORTEOCO	Óbito MIF. Informar quando a morte ocorreu: 1 – na gravidez; 2 – no parto; 3 – no aborto; 4 – até 42 dias após o parto; 5 – de 43 dias a 1 ano após o parto; 8 – não ocorreu nestes períodos; 9 – ignorado.	todos registros em branco
OBITOGRAV	Óbito na gravidez: 1 – sim; 2 – não; 9 – ignorado.	todos registros de criança, nenhum de MIF
OBITOPUERP	Óbito no puerpério: 1 – Sim, até 42 dias após o parto; 2 – Sim, de 43 dias a 1 ano; 3 – Não; 9 – Ignorado.	todos registros de criança, nenhum de MIF
EXAME	Exame: 1 – sim; 2 – não; 9 – ignorado.	pós óbito, não caracteriza possível fator
CIRURGIA	Cirurgia: 1 – sim; 2 – não; 9 – ignorado.	pós óbito, não caracteriza possível fator
NECROPSIA	Confirmação do diagnóstico por necrópsia: 1 – sim; 2 – não; 9 – ignorado.	pós óbito, não caracteriza possível fator
DSTEMPO	tempo de duração dos CID's informados	trabalhar com causa básica
DSEXPLICA	Descrição da explicação das regras de seleção da causa básica	trabalhar com causa básica
MEDICO	nome do médico	dado identificação profissional
CRM	CRM do médico	dado identificação profissional
TPASSINA	sem descrição da variável no dicionário de dados fornecido	todos registros em branco, não descrito no dicionário de dados
COMUNSVOIM	Código do município do SVO ou do IML	todos registros em branco
CONTATO	Meio de contato do atestante (telefone, fax, email etc.).	dado identificação profissional
DTATESTADO	data do atestado: ddmmaaaa	pós óbito, não caracteriza possível fator
ACIDTRAB	Indicar se foi acidente de trabalho: 1 – sim; 2 – não; 9 – ignorado.	todos registros não ou ignorado
FONTE	Indicar a fonte da informação de morte não natural, conforme a tabela: 1 – boletim de ocorrência; 2 – hospital; 3 – família; 4 – outra; 9 – ignorado	pós óbito, não caracteriza possível fator
DSEVENTO	descrição sumária do acidente	campo texto, sem padronização
ENDACID	endereço do acidente	somente para causas externas
NUMEROLOTE	numero do lote	informação sistema
TPPOS	Óbito investigado: 1 – sim; 2 – não.	pós óbito, não caracteriza possível fator
DTINVESTIG	Data da investigação: dd mm aaaa	pós óbito, não caracteriza possível fator
LINHAA_O	CIDs informados na Linha originalmente da DO	trabalhar com causa básica pós investigação
LINHAB_O	CIDs informados na Linha originalmente da DO	trabalhar com causa básica pós investigação
LINHAC_O	CIDs informados na Linha originalmente da DO	trabalhar com causa básica pós investigação
LINHAD_O	CIDs informados na Linha originalmente da DO	trabalhar com causa básica pós investigação
LINHAII_O	CIDs informados na Linha originalmente da DO	trabalhar com causa básica pós investigação
CAUSABAS_O	CIDs informados na Linha originalmente da DO	trabalhar com causa básica pós investigação
DTCADASTRO	data do cadastro	informação sistema
ATESTANTE	Indica se o médico que assina atendeu o paciente: 1 – Sim; 2 – Substituto; 3 – IML; 4 – SVO; 5 – Outros.	pós óbito, não caracteriza possível fator
DESCACID	descrição do acidente	campo texto, sem padronização
CODENDACID	código de endereço do acidente	todos registros em branco
NUMENDACID	número do endereço do acidente	todos registros em branco
COMPLACID	complemento do endereço onde ocorreu o acidente	todos registros em branco
CEPACID	cep do endereço do acidente	todos registros em branco
CONFPEPO	sem descrição da variável no dicionário de dados fornecido	alguns registros com X, não descrito no dicionário de dados
CONFIDADE	sem descrição da variável no dicionário de dados fornecido	alguns registros com X, não descrito no dicionário de dados
CONFCAUSA	sem descrição da variável no dicionário de dados fornecido	alguns registros com X, não descrito no dicionário de dados
CONFCIDADE	sem descrição da variável no dicionário de dados fornecido	alguns registros com X, não descrito no dicionário de dados
CRITICA	sem descrição da variável no dicionário de dados fornecido	todos registros com mesmo valor, não descrito no dicionário de dados
CODINST	Código de configuração da instalação: – 1º caractere: nível de instalação (M – municipal; R – regional; E – estadual); – 2º e 3º caractere: UF de instalação; – 4º	informação sistema

	ao 9º caractere: código do município de instalação; – 10º ao 13º caractere: nº da máquina de instalação.	
STCODIFICA	Status de instalação: se codificadora (valor: S) ou não (valor: N)	informação sistema
CODIFICADO	Se estiver codificado (valor: S) ou não (valor: N)	informação sistema
VERSAOSIST	Versão do sistema	informação sistema
VERSAOSCB	Versão do seletor de causa básica	informação sistema
RETROALIM	Retroalimentação	informação sistema
FONTAINV	Fonte de investigação. Valores: 1 – Comitê de Morte Materna e/ou Infantil; 2 – Visita domiciliar / Entrevista família; 3 – Estabelecimento de Saúde / Prontuário; 4 – Relacionado com outros bancos de dados; 5 – S V O; 6 – I M L; 7 – Outra fonte; 8 – Múltiplas fontes; 9 – Ignorado.	pós óbito, não caracteriza possível fator
DTRECEBIM	Data do recebimento: dd mm aaaa	informação sistema
ATESTADO	CIDs informado no atestado	trabalhar com causa básica
CAUSABAS_R	causa básica resseleccionada	trabalhar com causa básica
DTRESSELE	Data da resseleção: dd mm aaaa	explicação seleção causa básica
STRESSELE	Status da resseleção: 1 – sim; 2 – não.	explicação seleção causa básica
EXPLICA_R	explicação das regras de resseleção de causa básica	explicação seleção causa básica
VRRESSELE	Versão de resseleção	explicação seleção causa básica
COMPARA_CB	Compara causa básica resseleccionada com a informada. Valores: IGUAL, DIFER.	explicação seleção causa básica
NRESSELE	descrição da explicação da não resseleção da causa básica.	explicação seleção causa básica
CB_PRE	Causa básica informada antes da resseleção (NRESSELE)	explicação seleção causa básica
NPROC	Códigos da explicação da não resseleção da causa básica. Valores: 1 – Depende de perguntas; 2 – Causa externa; 3 – Procedimento médico; 4 – Causa básica por CID de paralisia; 5 – Regra F; 6 – CID temporário não pode ser causa básica.	explicação seleção causa básica
DIFDATA	Diferença entre a data de óbito e data do recebimento original da DO ([DTOBITO] – [DTRECORIG])	explicação seleção causa básica
VERSCBPRE	Versão do SCB da resseleção	explicação seleção causa básica
DTRECORIG	data do recebimento original: dd mm aaaa	informação sistema
CAUSAMAT	causa externa associada a uma causa materna	mortalidade materna
ESCMAEAGR1	Escolaridade 2010 agregada. Valores: 00 – Sem Escolaridade; 01 – Fundamental I Incompleto; 02 – Fundamental I Completo; 03 – Fundamental II Incompleto; 04 – Fundamental II Completo; 05 – Ensino Médio Incompleto; 06 – Ensino Médio Completo; 07 – Superior Incompleto; 08 – Superior Completo; 09 – Ignorado; 10 – Fundamental I Incompleto ou Inespecífico; 11 – Fundamental II Incompleto ou Inespecífico; 12 – Ensino Médio Incompleto ou Inespecífico.	todos registros em branco
ESCMAEAGR2	Escolaridade em anos. Valores: 1 – Nenhuma; 2 – de 1 a 3 anos; 3 – de 4 a 7 anos; 4 – de 8 a 11 anos; 5 – 12 anos e mais; 9 – Ignorado	todos registros em branco
ESCFALAGR1	Escolaridade 2010 agregada. Valores: 00 – Sem Escolaridade; 01 – Fundamental I Incompleto; 02 – Fundamental I Completo; 03 – Fundamental II Incompleto; 04 – Fundamental II Completo; 05 – Ensino Médio Incompleto; 06 – Ensino Médio Completo; 07 – Superior Incompleto; 08 – Superior Completo; 09 – Ignorado; 10 – Fundamental I Incompleto ou Inespecífico; 11 – Fundamental II Incompleto ou Inespecífico; 12 – Ensino Médio Incompleto ou Inespecífico.	todos registros em branco
ESCFALAGR2	Escolaridade em anos. Valores: 1 – Nenhuma; 2 – de 1	todos registros em branco

	a 3 anos; 3 – de 4 a 7 anos; 4 – de 8 a 11 anos; 5 – 12 anos e mais; 9 – Ignorado	
STDOEPIDEM	Status de DO Epidemiológica. Valores: 1 – SIM; 0 – NÃO.	informação sistema
STDONOVA	Status de DO Nova. Valores: 1 – SIM; 0 – NÃO.	informação sistema
REGRA	Regra de seleção da causa básica	todos óbitos investigados e qualificados. Utilizado direto a causa básica

Fonte: A autora (2020).

Quadro 10 – Variáveis SINASC não utilizadas nos experimentos de KDD, com respectiva descrição e justificativa.

rótulo (original)	descrição	justificativa exclusão
NUMERODV	Dígito verificador da DNV	outra variável (numerodn)
PREFIXODN	Prefixo da DNV	outra variável (numerodn)
CODCART	Código do cartório	sem utilidade (cartório)
CODESTCART	Código de estabelecimento do cartório	sem utilidade (cartório)
CODMUNCART	Código do município do cartório	sem utilidade (cartório)
NUMREGCART	Número da Regional do Cartório	sem utilidade (cartório)
DTREGCART	Data de registro no Cartório	sem utilidade (cartório)
COMUNESTAB	Código do município do estabelecimento de saúde	outra variável (codmunnasc)
ENDNASC	endereço de nascimento	dado de identificação
BAINASC	bairro de nascimento	outra variável (estabelecimento de saúde)
CODBAINASC	código bairro de nascimento	outra variável (já existe variável bairro)
CODENDNASC	código de endereço do nascimento	todos registros em branco
COMPLNASC	complemento local nascimento	dado de identificação
NUMENDNASC	número do endereço de nascimento	dado de identificação
CEPNASC	cep do local de nascimento	outra variável (codmunnasc, codestab)
CODDISNASC	código do DS de nascimento	outra variável (codmunnasc, codestab)
CODREGOCOR	código regional de ocorrência do nascimento	outra variável (codmunnasc, codestab)
NOMEMAE	nome da mãe	dado de identificação
NUMSUSMAE	número SUS da mãe	dado de identificação
CONFIDADE	sem descrição da variável no dicionário de dados fornecido	não descrito no dicionário de dados
CODESTRES	código da UF de residência	todos registros com mesmo valor (41 - PR)
CODMUNRES	código do município de residência	todos registros com mesmo valor (410690 - Curitiba)
ENDRES	endereço de residência	dado de identificação
CODENDRES	código do endereço de residência	dado de identificação
COMPLRES	complemento do endereço de residência	dado de identificação
NUMRES	número da residência	dado de identificação
CEPRES	cep da residência	outra variável (baires)
CODPAISRES	código país de residência	todos registros com mesmo valor (país - Brasil)
CODREGRES	código regional de residência	todos registros com mesmo valor (2ª Regional)
HORANASC	hora de nascimento	trabalhado somente com data
DTCADASTRO	Data do cadastro da DN no sistema	sem utilidade (sistema)
CONFPEO	?	não descrito no dicionário de dados
NUMEROLOTE	número do lote	sem utilidade (sistema)
CRITICA	crítica	sem utilidade (sistema)
CODINST	código de instalação	sem utilidade (sistema)
VERSAOSIST	versão sistema	sem utilidade (sistema)
RETROALIM	retroalimentação	sem utilidade (sistema)
DTRECEBIM	data de recebimento da DNV	todos registros em branco
DIFDATA	diferença entre recebimento e digitação DNV	todos registros em branco
DTRECORIG	Data do 1º recebimento do lote, dada pelo Sisnet	todos registros em branco
NOMEPAI	nome do pai	dado de identificação
DTULTMENST	data da última menstruação (DUM): dd mm aaaa	outra variável (semagestac)
NOMERNASC	nome do recém-nascido	dado de identificação
NOMERESPO	Nome do responsável pelo preenchimento	sem utilidade (preenchimento)
TPFUNCRESP	Tipo de função do responsável pelo preenchimento. Valores: 1– Médico; 2– Enfermeiro; 3– Parteira; 4–	sem utilidade (preenchimento)

	Funcionário do cartório; 5– Outros.	
DSFUNCRESP	Descrição da função do responsável pelo preenchimento, se a função for “5– Outros”	sem utilidade (preenchimento)
CODPROF	Código da profissão, se a função for “5– Outros”.	sem utilidade (preenchimento)
PROFISSION	Nome do Profissional, se a função for “5– Outros”.	sem utilidade (preenchimento)
COREN	Nome do profissional enfermeiro	sem utilidade (preenchimento)
CODCOREN	Código do COREN	sem utilidade (preenchimento)
TPDOCRESP	Tipo do documento do responsável. Valores: 1-CNES; 2-CRM; 3COREN; 4-RG; 5-CPF.	sem utilidade (preenchimento)
NUDOCRESP	Número do documento do responsável pelo preenchimento.	sem utilidade (preenchimento)
ORGEMISSION	descrição do órgão emissor do documento do responsável pelo preenchimento	sem utilidade (preenchimento)
CODCRM	código do CRM do médico	dado de identificação
CODPROPREE	sem descrição da variável no dicionário de dados fornecido	todos registros em branco, não descrito no dicionário de dados do MS
DTDECLARAC	Data da declaração: dd mm aaaa	sem utilidade (sistema)
ESCMAEAGR2	sem descrição da variável no dicionário de dados fornecido	não descrito no dicionário de dados do MS
STDNEPIDEM	Status de DO Epidemiológica. Valores: 1 – SIM; 0 – NÃO.	todos registros com mesmo valor (0)
STDNNOVA	Status de DO Nova. Valores: 1 – SIM; 0 – NÃO.	sem utilidade (sistema)

Fonte: A autora (2020).

## APÊNDICE E – LINKAGE – DESCRIÇÃO DETALHADA E SCRIPT REALIZADO NO R

No intuito de encontrar o maior número possível de pares de óbito-nascido vivo, realizou-se *linkage* com três diferentes chaves de ligação: (1) número da DNV, (2) nome e último sobrenome da mãe + data de nascimento nascido vivo + sexo nascido vivo e (3) chave de ligação 2 com diferença de um caractere no nome e último sobrenome da mãe. A chave de ligação 3 objetivou captar pares com possíveis equívocos na escrita, seja por troca – Cristiane e Cristiana, por adição de um caractere – Cristiane e Cristianne ou por subtração de um caractere – Cristiane e Cristine. O *script* realizado consta na figura 21.

A partir dos conjuntos de dados iniciais SINASC (479.309 registros de nascidos vivos) e SIM (5.482 registros de óbitos), pareou-se 4.243 óbitos a nascidos vivos pelo número da DNV (chave de ligação 1), 660 pelo nome e último sobrenome mãe + data de nascimento + sexo (chave de ligação 2) e 129 pelo nome e último sobrenome mãe com distância 1 + data de nascimento + sexo (chave de ligação 3), totalizando 5.032 óbitos pareados a nascido vivo. Esses arquivos foram unidos e acrescidos dos registros de nascidos vivos que não foram pareados a nenhum óbito (474.293), dando origem ao conjunto de dados SIM-SINASC inicial com 479.325 registros. Quatrocentos e oitenta e nove óbitos (9,72%) foram descartados por não terem sido pareados a nenhum nascido vivo. A Figura X esquematiza o processo de *linkage* e seus resultados.

Figura 21 – Script do linkage realizado no R.®

```

sim_path <- "F:/SIM.xlsx"
sinasc_path <- "F:/SINASC.xlsx"
sim <- read.xlsx(sim_path, sheet = "1999_2017")
sinasc <- read.xlsx(sinasc_path)
sinasc$mae_nome <- stringi::stri_extract_first_words(sinasc$NOMEMAE)
sinasc$mae_sobrenome <- stringi::stri_extract_last_words(sinasc$NOMEMAE)
sim$mae_nome <- stringi::stri_extract_first_words(sim$NOMEMAE)
sim$mae_sobrenome <- stringi::stri_extract_last_words(sim$NOMEMAE)
a1 <- sim %>% inner_join(sinasc, by = c("NUMERODN"="NUMERODN"), suffix = c(".i", "")) %>%
select(-ends_with(".i"))
write.csv2(a1, "F:/a1.csv")
sim$DTNASC <- dmy(sim$DTNASC)
sinasc$DTNASC <- dmy(sinasc$DTNASC)
a2 <- sim %>% inner_join(sinasc, by=c("mae_nome"="mae_nome",
    mae_sobrenome"="mae_sobrenome", "DTNASC"="DTNASC", "SEXO"="SEXO"), keep=T,
    copy=T, suffix = c(".i", "")) %>% select(-ends_with(".i"))
write.csv2(a2, "F:/a2.csv")
a2.1 <- a2 %>% filter(! id_sim %in% a1$id_sim)
write.csv2(a2.1, "F:/a2.1.csv")
a3 <- sim %>% filter(!is.na(DTNASC) & !is.na(SEXO) & !is.na(mae_nome) & !is.na(mae_sobrenome))
%>%
    stringdist_left_join(
        sinasc %>% filter(!is.na(DTNASC) & !is.na(SEXO) & !is.na(mae_nome) & !is.na(mae_sobrenome)),
        by = c("mae_nome"="mae_nome", "mae_sobrenome"="mae_sobrenome"), max_dist=1) %>%
        filter(DTNASC.y == DTNASC.x & SEXO.x == SEXO.y)
write.csv2(a3, "F:/a3.csv")
a3.1t <- a3 %>% select(-ends_with(".x")) %>% rename_all(funs(str_replace(., ".y", "")))
write.csv2(a3.1t, "F:/a3.1t.csv")
a3.2 <- a3.1t %>% filter(! id_sim %in% c(a1$id_sim, a2.1$id_sim))
write.csv2(a3.2, "F:/a3.2.csv")
a_todos <- rbind(a1, a2.1, a3.2)
write.csv2(a_todos, "F:/a_todos.csv")
todos <- readr::read_csv2("a_todos.csv", col_names = T)
todos$X1 <- NULL
nao_sim <- sinasc %>% filter(!id_sinasc %in% a_todos$id_sinasc)
sinasc_repetidos <- a_todos %>% filter(id_sinasc %in% sinasc_repetidos) %>%
a_todos$DTNASC <- dmy(a_todos$DTNASC)
nao_sim$DTNASC <- ymd(nao_sim$DTNASC)
data_bind <- bind_rows(a_todos, nao_sim)
head(data_bind$DTNASC)
write_csv2(data_bind, "F:/join_todos.csv")

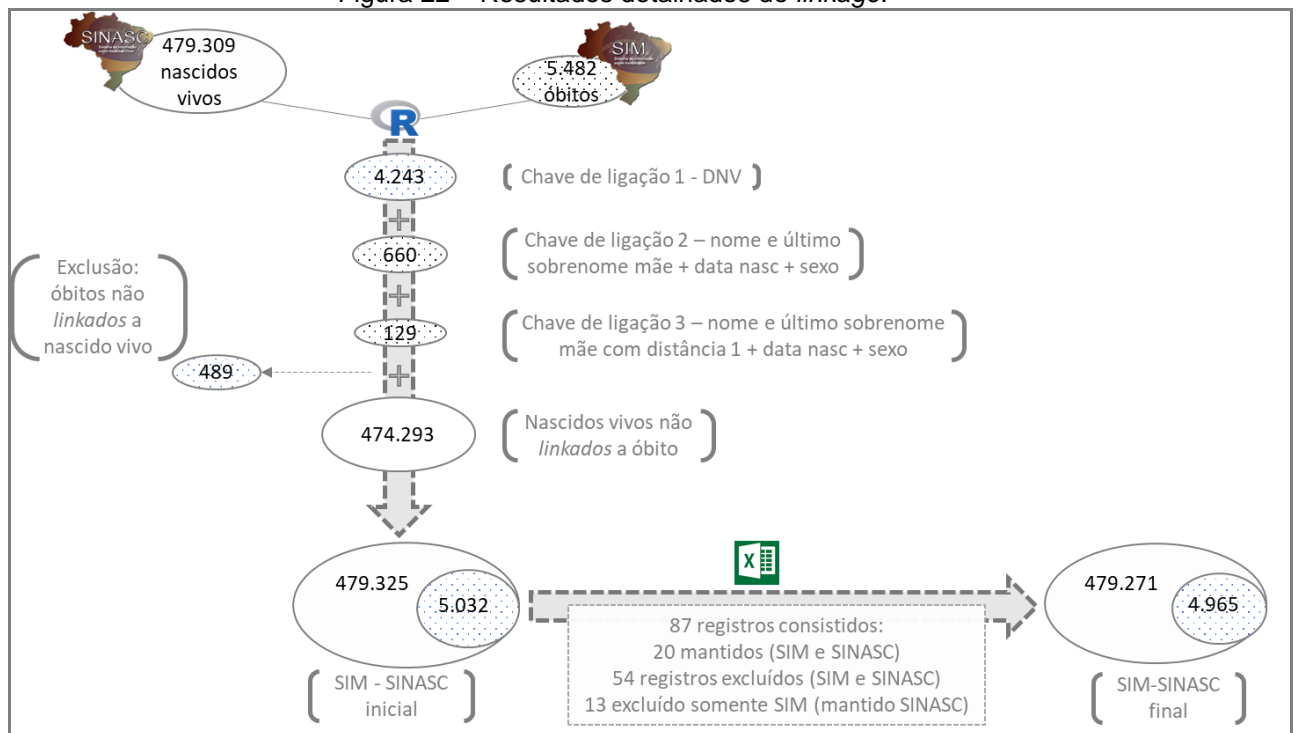
```

Fonte: A autora (2020).



O número maior de registros após *linkage* deveu-se à ligação do mesmo óbito a mais de um nascido vivo, como no caso de óbitos de gêmeos do mesmo sexo. As duplicidades e inconsistências foram verificadas no *Excel*<sup>®</sup> e, quando possível, corrigiu-se o registro a partir da verificação de outras variáveis como número de DNV e peso ao nascer, caso contrário, realizou-se a sua exclusão. Assim, 87 registros de óbito-nascido vivo foram consistidos no *Excel*<sup>®</sup> dos quais 20 foram mantidos por estarem corretos, 54 excluídos (46 por não ser possível identificar se o *linkage* estava correta e oito por estarem com data de óbito anterior à data de nascimento) e em 13 registros foram excluídos somente os dados relacionados ao óbito por terem sido pareados incorretamente. O resultado final do *linkage* foi o conjunto de dados contendo 479.271 nascidos vivos dos quais 4.965 foram a óbito (Figura 22).

Figura 22 – Resultados detalhados do *linkage*.



Fonte: A autora (2020).

## APÊNDICE F – ÁRVORES DE DECISÃO DO EXPERIMENTO 1, POR ANO

Figura 23 – Árvore de decisão do ano de 2000 do experimento 1.

Regra	SE	antecedente	ENTÃO	consequente	cobertura	erro
1		Peso ao nascer <=1450g E Apgar no 5ºminuto <=3		Óbito evitável por prevenção, diagnóstico e tratamento precoces	71.64	48.44
2		Peso ao nascer <=1450g E Apgar no 5ºminuto >3		Não óbito	302.98	111.04
3		Peso ao nascer >1450g		Não óbito	28931.39	217.99

Fonte: A autora (2020).

Figura 24 – Árvore de decisão do ano de 2001 do experimento 1.

Regra	SE	antecedente	ENTÃO	consequente	cobertura	erro
1		Peso ao nascer >1400g		Não óbito	26802.75	206.99
2		Peso ao nascer entre 936-1400g E Malformação congênita = não		Não óbito	166.14	47.01
3		Peso ao nascer entre 936-1400g E Malformação congênita = ignorado		Não óbito	0.01	
4		Peso ao nascer entre 936-1400g E Malformação congênita = sim E anos de estudo da mãe de 4-7 anos		Óbito evitável por prevenção, diagnóstico e tratamento precoces	2.0	1.0
5		Peso ao nascer entre 936-1400g E Malformação congênita = sim E anos de estudo da mãe de 8-11 anos		Óbito por causa não específica	3.0	0.0
6		Peso ao nascer entre 936-1400g E Malformação congênita = sim E anos de estudo da mãe >12 anos		Óbito evitável por parceria com outros setores	4.01	1.0
7		Peso ao nascer entre 936-1400g E Malformação congênita = sim E anos de estudo da mãe ignorado		Óbito evitável por adequado controle na gravidez	1.0	0.0
8		Peso ao nascer <=935g E Bairro de residência = Cajuru E Quantidade de filhos vivos <=3 E Quantidade de filhos mortos >0		Óbito evitável por prevenção, diagnóstico e tratamento precoces	5.0	
9		Peso ao nascer <=935g E Bairro de residência = Cajuru E Quantidade de filhos vivos >3		Óbito por causa não específica	2.0	
10		Peso ao nascer <=935g E Bairro de residência = Uberaba		Não óbito	3.0	1.0
11		Peso ao nascer <=935g E Bairro de residência = Parolim		Óbito evitável por prevenção, diagnóstico e tratamento	2.0	1.0

			precoces		
12		Peso ao nascer <=935g E Bairro de residência = Tatuquara	Óbito por causa não específica	2.0	1.0
13		Peso ao nascer <=935g E Bairro de residência = Centro	Óbito evitável por prevenção, diagnóstico e tratamento precoces	4.0	1.0
14		Peso ao nascer <=935g E Bairro de residência = Cristo Rei	Óbito evitável por adequada atenção ao parto	1.0	
15		Peso ao nascer <=935g E Bairro de residência = Sitio Cercado E Idade materna <=22	Não óbito	2.0	
16		Peso ao nascer <=935g E Bairro de residência = Sitio Cercado E Idade materna >22	Óbito por causa não específica	6.0	1.0
17		Peso ao nascer <=935g E Bairro de residência = Santa Felicidade	Óbito evitável por prevenção, diagnóstico e tratamento precoces	2.0	1.0
18		Peso ao nascer <=935g E Bairro de residência = Alto Boqueirão E CBO materno <=16300	Óbito evitável por prevenção, diagnóstico e tratamento precoces	3.0	1.0
19		Peso ao nascer <=935g E Bairro de residência = Alto Boqueirão E CBO materno >16300	Óbito por causa não específica	2.01	0.01
20		Peso ao nascer <=935g E Bairro de residência = Ignorado	Óbito evitável por adequada atenção ao parto	2.0	1.0
21		Peso ao nascer <=935g E Bairro de residência = Fazendinha	Óbito evitável por adequada atenção ao parto	2.0	1.0
22		Peso ao nascer <=935g E Bairro de residência = Orleans	Óbito evitável por prevenção, diagnóstico e tratamento precoces	2.0	
23		Peso ao nascer <=935g E Bairro de residência = Xaxim E anos de estudo da mãe de 4-7 anos	Não óbito	1.0	
24		Peso ao nascer <=935g E Bairro de residência = Xaxim E anos de estudo da mãe de 8-11 anos	Óbito por causa não específica	2.0	
25		Peso ao nascer <=935g	Óbito evitável por	2.0	

	E Bairro de residência = Xaxim	prevenção, diagnóstico e tratamento precoces		
26	E anos de estudo da mãe >12 anos Peso ao nascer <=935g E Bairro de residência = Novo Mundo	Óbito por causa não específica	2.0	
27	E Peso ao nascer <=935g E Bairro de residência = Santa Quiteria	Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	0.0
28	E Peso ao nascer <=935g E Bairro de residência = Capão da Imbuia	Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	0.0
29	E Peso ao nascer <=935g E Bairro de residência = Santa Candida E anos de estudo da mãe de 1-3 anos	Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	0.0
30	E Peso ao nascer <=935g E Bairro de residência = Santa Candida E anos de estudo da mãe de 4-7 anos	Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	0.0
31	E Peso ao nascer <=935g E Bairro de residência = Santa Candida E anos de estudo da mãe de 8-11 anos	Óbito evitável por adequado controle na gravidez	2.0	0.0
32	E Peso ao nascer <=935g E Bairro de residência = Santa Candida E anos de estudo da mãe >12 anos	Óbito evitável por adequado controle na gravidez	2.0	0.0
33	E Peso ao nascer <=935g E Bairro de residência = Santa Candida E anos de estudo da mãe ignorado	Óbito por causa não específica	1.0	0.0
34	E Peso ao nascer <=935g E Bairro de residência = CIC Pinheirinho	Óbito evitável por prevenção, diagnóstico e tratamento precoces	3.01	1.01
35	E Peso ao nascer <=935g E Bairro de residência = Água Verde	Óbito por causa não específica	1.0	
36	E Peso ao nascer <=935g E Bairro de residência = Campo Comprido	Óbito por causa não específica	3.0	2.0
37	E Peso ao nascer <=935g E Bairro de residência = Boa Vista	Óbito evitável por prevenção, diagnóstico e tratamento	2.0	1.0

		Quantidade de filhos mortos >0		precoces		
38		Peso ao nascer <=935g E Bairro de residência = Boa Vista E Quantidade de filhos mortos <=0		Não óbito	4.0	1.0
39		Peso ao nascer <=935g E Bairro de residência = Boqueirão		Óbito evitável por prevenção, diagnóstico e tratamento precoces	3.01	1.01
40		Peso ao nascer <=935g E Bairro de residência = Pinheirinho		Óbito evitável por adequado controle na gravidez	1.0	0.0
41		Peso ao nascer <=935g E Bairro de residência = Rebouças		Óbito evitável por prevenção, diagnóstico e tratamento precoces	5.0	1.0
42		Peso ao nascer <=935g E Bairro de residência = Jardim Social		Óbito evitável por adequada atenção ao parto	1.0	0.0
43		Peso ao nascer <=935g E Bairro de residência = Prado Velho		Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	
44		Peso ao nascer <=935g E Bairro de residência = CIC Portão E Quantidade de filhos vivos <=0		Óbito evitável por prevenção, diagnóstico e tratamento precoces	3.0	
45		Peso ao nascer <=935g E Bairro de residência = CIC Portão E Quantidade de filhos vivos >0		Óbito por causa não específica	3.0	2.0
46		Peso ao nascer <=935g E Bairro de residência = Abranches/Santa Efigênia		Óbito evitável por adequado controle na gravidez	1.0	
47		Peso ao nascer <=935g E Bairro de residência = Tingui		Óbito evitável por adequado controle na gravidez	1.0	
48		Peso ao nascer <=935g E Bairro de residência = São Francisco		Óbito evitável por prevenção, diagnóstico e tratamento precoces	2.0	1.0
49		Peso ao nascer <=935g E Bairro de residência = Butiatuvinha		Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	
50		Peso ao nascer <=935g E Bairro de residência = São Braz		Não óbito	1.0	
51		Peso ao nascer <=935g E Bairro de residência = CIC Santa Felicidade		Não óbito	1.0	
52		Peso ao nascer <=935g E Bairro de residência = Cabral		Óbito evitável por adequada atenção ao parto	1.0	
53		Peso ao nascer <=935g		Óbito evitável por	0.01	0.0

		E Bairro de residência = Bacacheri	parceria com outros setores		
54		Peso ao nascer <=935g E Bairro de residência = Pilarzinho	Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	
55		Peso ao nascer <=935g E Bairro de residência = Vila Hauer	Óbito por causa não específica	3.01	1.01
56		Peso ao nascer <=935g E Bairro de residência = Atuba	Não óbito	2.0	1.0
57		Peso ao nascer <=935g E Bairro de residência = Centro Cívico	Não óbito	1.0	
58		Peso ao nascer <=935g E Bairro de residência = Santo Inácio	Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	
59		Peso ao nascer <=935g E Bairro de residência = Vila Izabel	Não óbito	2.0	

Fonte: A autora (2020).

Figura 25 – Árvore de decisão do ano de 2002 do experimento 1.

Regra		antecedente		consequente	cobertura	erro
1		Peso ao nascer <=910g		Óbito evitável por adequado controle na gravidez	107.1	63.1
2		Peso ao nascer entre 911 - 1540g		Não óbito	253.23	54.02
3		Peso ao nascer >1540g E Malformação congênita = não		Não óbito	25781.53	146.97
4		Peso ao nascer >1540g E Malformação congênita = ignorado		Não óbito	5.99	
5	SE	Peso ao nascer entre 1541-2370g E Malformação congênita = sim E dia da semana do nascimento = quarta-feira	ENTÃO	Óbito por causa não específica	5.0	1.0
6		Peso ao nascer entre 1541-2370g E Malformação congênita = sim E dia da semana do nascimento = sexta-feira E Peso ao nascer <=1940		Óbito evitável por parceria com outros setores	2.0	0.0
7		Peso ao nascer entre 1541-2370g E Malformação congênita = sim E dia da semana do nascimento = sexta-feira E Peso ao nascer >1940		Óbito evitável por adequado controle na gravidez	2.0	0.0
8		Peso ao nascer entre 1541-2370g E		Óbito por causa não específica	2.0	

	Malformação congênita = sim E dia da semana do nascimento = sábado E Sexo masculino			
9	Peso ao nascer entre 1541-2370g E Malformação congênita = sim E dia da semana do nascimento = sábado E Sexo feminino	Óbito evitável por parceria com outros setores	3.0	1.0
10	Peso ao nascer entre 1541-2370g E Malformação congênita = sim E Dia da semana do nascimento = domingo E Consultas de pré-natal >=7	Óbito evitável por parceria com outros setores	3.01	0.01
11	Peso ao nascer entre 1541-2370g E Malformação congênita = sim E Dia da semana do nascimento = domingo E Consultas de pré-natal entre 4-6 consultas	Não óbito	2.0	1.0
12	Peso ao nascer entre 1541-2370g E Malformação congênita = sim E Dia da semana do nascimento = segunda-feira	Não óbito	3.0	1.0
13	Peso ao nascer entre 1541-2370g E Malformação congênita = sim E Dia da semana do nascimento = quinta-feira	Óbito por causa não específica	2.0	1.0
14	Peso ao nascer entre 1541-2370g E Malformação congênita = sim E Dia da semana do nascimento = terça-feira	Não óbito	3.0	
15	Peso ao nascer entre 1541-2370g E Malformação congênita = sim E Peso >2370g	Não óbito	147.13	19.0

Fonte: A autora (2020).

Figura 26 – Árvore de decisão do ano de 2003 do experimento 1.

Regra	antecedente	consequente	cobertura	erro
1	Peso ao nascer >1765g	Não óbito	24290.61	157.96
2	Peso ao nascer entre 916-1765g E Malformação congênita = não	Não óbito	364.22	48.02
3	Peso ao nascer entre 916-1765g E Malformação congênita = ignorado	Não óbito	0.05	
4	Peso ao nascer entre 916-1765g E Malformação congênita = sim E Apgar no 1º minuto <=3 E Sexo masculino	Óbito por causa não específica	5.0	0.0
5	Peso ao nascer entre 916-1765g E Malformação congênita = sim E Apgar no 1º minuto <=3 E Sexo feminino E Consultas de pré-natal >=7	Não óbito	2.0	0.0
6	Peso ao nascer entre 916-1765g E Malformação congênita = sim E Apgar no 1º minuto <=3 E Sexo feminino E Consultas de pré-natal entre 4-6	Óbito por causa não específica	2.01	0.01
7	Peso ao nascer entre 916-1765g E Malformação congênita = sim E Apgar no 1º minuto >3	Óbito evitável por parceria com outros setores	3.01	1.0
8	Peso ao nascer <=915g	Óbito evitável por adequado controle na gravidez	134.1	76.1

Fonte: A autora (2020).



Figura 27 – Árvore de decisão do ano de 2004 do experimento 1.

Regra		antecedente		consequente	cobertura	erro
1	SE	Peso ao nascer <= 865g	ENTÃO	Óbito evitável por adequado controle na gravidez	96.01	47.01
2		Peso ao nascer entre 866-1685g E Malformação congênita = não		Não óbito	372.05	53.02
3		Peso ao nascer entre 866-1685g E Malformação congênita = sim E Consultas de pré-natal entre 4-6 E Apgar no 1º minuto <=5		Óbito não evitável	3.0	0.0
4		Peso ao nascer entre 866-1685g E Malformação congênita = sim E Consultas de pré-natal entre 4-6 E Apgar no 1º minuto >5		Não óbito	3.0	
5		Peso ao nascer entre 866-1685g E Malformação congênita = sim E Consultas de pré-natal = 0		Não óbito	1.0	
6		Peso ao nascer entre 866-1685g E Malformação congênita = sim E Consultas de pré-natal >=7		Óbito não evitável	7.0	3.0
7		Peso ao nascer entre 866-1685g E Malformação congênita = sim E Consultas de pré-natal de 1-3		Não óbito	2.0	1.0
8		Peso ao nascer entre 866-1685g E Malformação congênita = ignorado		Não óbito	1.0	
9		Peso ao nascer >1685g		Não óbito	24637.94	157.98

Fonte: A autora (2020).

Figura 28 – Árvore de decisão do ano de 2005 do experimento 1.

Regra		antecedente		consequente	cobertura	erro
1	SE	Peso ao nascer <=850g	ENTÃO	Óbito evitável por adequado controle na gravidez	98.02	45.02
2		Peso ao nascer entre 851-1255		Não óbito	144.03	45.0
3		Peso ao nascer >1255		Não óbito	24194.95	154.0

Fonte: A autora (2020).

Figura 29 – Árvore de decisão do ano de 2006 do experimento 1.

Regra		antecedente	ENTÃO	consequente	cobertura	erro
1	SE	Peso ao nascer <=1545g E Malformação congênita = não E Peso ao nascer <=745g		Óbito evitável por adequado controle na gravidez	57.0	25.0
2		Peso ao nascer <=1545g E Malformação congênita = não E Peso ao nascer >745g		Não óbito	270.02	61.0
3		Peso ao nascer <=1545g E Malformação congênita = sim		Óbito não evitável	18.0	8.0
4		Peso ao nascer >1545g		Não óbito	23983.97	141.0

Fonte: A autora (2020).

Figura 30 - Árvore de decisão do ano de 2007 do experimento 1.

Regra	antecedente	consequente	cobertura	erro
1	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = não E Duração da gestação = 32 a 36 semanas	Não óbito	4.01	
2	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = não E Duração da gestação = 37 a 41 semanas	Não óbito	3.01	
3	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = não E Duração da gestação = 22 a 27 semanas	Óbito evitável por adequado controle na gravidez	81.0	33.0
4	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = não E Duração da gestação = 28 a 31 semanas	Não óbito	33.0	14.0
5	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = não E Duração da gestação <22 semanas	Óbito evitável por adequado controle na gravidez	5.0	
6	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = não E Duração da gestação >42 semanas	Não óbito	0.01	
7	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = não E Duração da gestação = ignorado	Não óbito	0.01	
8	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = sim	Óbito evitável por parceria com outros setores	2.0	
9	Peso ao nascer <=1325g E Peso ao nascer <=935g E Malformação congênita = ignorado	Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	
10	Peso ao nascer <=1325g E Peso ao nascer >935g E Malformação congênita = não E Duração da gestação = 32 a 36 semanas	Não óbito	127.02	23.0
11	Peso ao nascer >1325g E Malformação congênita = não	Não óbito	23595.75	110.0
12	Peso ao nascer >1325g E Malformação congênita = sim E Peso ao nascer <=2160	Óbito não evitável	24.01	13.01
13	Peso ao nascer >1325g E Malformação congênita = sim E Peso ao nascer >2160	Não óbito	110.18	11.0
14	Peso ao nascer >1325g E Malformação congênita = ignorado	Não óbito	8.01	

Fonte: A autora (2020).

Figura 31 – Árvore de decisão do ano de 2008 do experimento 1.

Regra	antecedente	consequente	cobertura	erro			
1	Peso ao nascer <=1390g E Peso ao nascer <=740g	Óbito evitável por adequado controle na gravidez	58.0	22.0			
2	Peso ao nascer <=1390g E Peso ao nascer >740g E Malformação congênita = não				Não óbito	205.0	58.0
3	Peso ao nascer <=1390g E Peso ao nascer >740g E Malformação congênita = sim				Óbito não evitável	6.0	2.0
4	Peso ao nascer >a390g	Não óbito	24725.0	151.0			

Fonte: A autora (2020).

Figura 32 – Árvore de decisão do ano de 2009 do experimento 1.

Regra	antecedente	consequente	cobertura	erro			
1	Peso ao nascer <=1510g E Peso ao nascer <=785g	Óbito evitável por adequado controle na gravidez	69.0	28.0			
2	Peso ao nascer <=1510g E Peso ao nascer >785g E Malformação congênita = não				Não óbito	260.0	41.0
3	Peso ao nascer <=1510g E Peso ao nascer >785g E Malformação congênita = sim E Apgar no 5º min <=7 E Consultas de pré-natal >=7				Óbito não evitável	4.0	1.0
4	Peso ao nascer <=1510g E Peso ao nascer >785g E Malformação congênita = sim E Apgar no 5º min <=7 E Consultas de pré-natal entre 4-6	Óbito evitável por parceria com outros setores	2.0				
5	Peso ao nascer <=1510g E Peso ao nascer >785g E Malformação congênita = sim E Apgar no 5º min >7	Não óbito	6.0	1.0			
6	Peso ao nascer <=1510g E Peso ao nascer >785g E Malformação congênita = ignorado	Não óbito	1.0				
7	Peso ao nascer >1510g	Não óbito	24213.0	113.0			

Fonte: A autora (2020).

Figura 33 – Árvore de decisão do ano de 2010 do experimento 1.

Regra		antecedente		consequente	cobertura	erro
1	SE	Peso ao nascer <=1630g E Malformação congênita = não E Peso ao nascer <=920g E CBO materno <=411045	ENTÃO	Não óbito	30.0	13.0
2		Peso ao nascer <=1630g E Malformação congênita = não E Peso ao nascer <=920g E CBO materno >411045		Óbito evitável por adequado controle na gravidez	79.01	31.01
3		Peso ao nascer <=1630g E Malformação congênita = não E Peso ao nascer >920		Não óbito	283.02	41.0
4		Peso ao nascer <=1630g E Malformação congênita = sim E Apgar no 1ºmin <=4		Óbito não evitável	9.0	1.0
5		Peso ao nascer <=1630g E Malformação congênita = sim E Apgar no 1ºmin >4 E Peso ao nascer <=1060g		Óbito não evitável	2.0	1.0
6		Peso ao nascer <=1630g E Malformação congênita = sim E Apgar no 1ºmin >4 E Peso ao nascer >1060g		Não óbito	4.0	
7		Peso ao nascer <=1630g E Malformação congênita = ignorado		Não óbito	3.0	
8		Peso ao nascer >1630g		Não óbito	24617.97	113.0

Fonte: A autora (2020).

Figura 34 – Árvore de decisão do ano de 2011 do experimento 1.

Regra		antecedente		consequente	cobertura	erro
1	<b>SE</b>	Peso ao nascer <=1510g E Apgar no 5ºminuto <=5 E Malformação congênita = sim	<b>ENTÃO</b>	Óbito não evitável	7.14	3.14
2		Peso ao nascer <=1510g E Apgar no 5ºminuto <=5 E Malformação congênita = não E Peso ao nascer <=855		Óbito evitável por adequado controle na gravidez	28.14	7.0
3		Peso ao nascer <=1510g E Apgar no 5ºminuto <=5 E Malformação congênita = não E Peso ao nascer >855 E Apgar no 1º minuto <=1		Não óbito	6.01	1.0
4		Peso ao nascer <=1510g E Apgar no 5ºminuto <=5 E Malformação congênita = não E Peso ao nascer >855 E Apgar no 1º minuto >1		Óbito evitável por adequado controle na gravidez	7.0	2.0
5		Peso ao nascer <=1510g E Apgar no 5ºminuto >5		Não óbito	305.76	66.86
6		Peso ao nascer >1510g		Não óbito	24515.97	127.0

Fonte: A autora (2020).

Figura 35 – Árvore de decisão do ano de 2012 do experimento 1.

Regra		antecedente		consequente	cobertura	erro
1	SE	Peso <=1235g E Malformação congênita = não E Peso ao nascer <=730g	ENTÃO	Óbito evitável por adequado controle na gravidez	78.0	34.0
2		Peso <=1235g E Malformação congênita = não E Peso ao nascer >730g		Não óbito	131.0	30.0
3		Peso <=1235g E Malformação congênita = sim E Apgar no 5º minuto <=7		Óbito não evitável	5.0	
4		Peso <=1235g E Malformação congênita = sim E Apgar no 5º minuto >7		Óbito evitável por adequado controle na gravidez	3.0	2.0
5		Peso <=1235g E Malformação congênita = ignorado E Turno nascimento = noite E Duração da gestação = 22-27 semanas		Óbito evitável por adequado controle na gravidez	2.0	
6		Peso <=1235g E Malformação congênita = ignorado E Turno nascimento = noite E Duração da gestação = 28-31 semanas		Não óbito	4.0	1.0
7		Peso <=1235g E Malformação congênita = ignorado E Turno nascimento = noite E Duração da gestação <22 semanas		Óbito evitável por adequado controle na gravidez	1.0	
8		Peso <=1235g E Malformação congênita = ignorado E Turno nascimento = tarde E Apgar no 1º minuto <=3		Óbito não evitável	2.0	
9		Peso <=1235g E Malformação congênita = ignorado E Turno nascimento = tarde E Apgar no 1º minuto >3		Não óbito	2.0	
10		Peso <=1235g E Malformação congênita = ignorado E Turno nascimento = manhã		Óbito evitável por adequado controle na gravidez	3.0	
11				Peso >1235g	Não óbito	24470.0

Fonte: A autora (2020).

Figura 36 – Árvore de decisão do ano de 2013 do experimento 1.

Regra		antecedente		consequente	cobertura	erro
1	SE	Peso <=1260g E Peso <=665g E Semanas gestação <=31 E Parto cesáreo anterior <=0	ENTÃO	Óbito evitável por adequado controle na gravidez	34.8	4.0
2		Peso <=1260g E Peso <=665g E Semanas gestação <=31 E Parto cesáreo anterior >0 E Quantidade de filhos mortos <=0		Óbito evitável por adequada atenção ao parto	4.0	1.0
3		Peso <=1260g E Peso <=665g E Semanas gestação <=31 E Parto cesáreo anterior >0 E Quantidade de filhos mortos >0		Óbito evitável por adequado controle na gravidez	2.0	1.0
4		Peso <=1260g E Peso <=665g E Semanas gestação >31 E Método para estimar semanas de gestação = exame físico		Óbito não evitável	2.17	1.17
5		Peso <=1260g E Peso <=665g E Semanas gestação >31 E Método para estimar semanas de gestação = outro método ou ignorado		Não óbito	6.35	0.35
6		Peso <=1260g E Peso >665g		Não óbito	199.0	54.0
7		Peso >1260g		Não óbito	24686.0	116.0

Fonte: A autora (2020).



Figura 37 – Árvore de decisão do ano de 2014 do experimento 1.

Regra	antecedente	consequente	cobertura	erro
1	Peso <=1105g E Malformação congênita = ignorado E recuperação de apgar <=0	Óbito evitável por adequada atenção ao parto	2.0	
1	Peso <=1105g E Malformação congênita = ignorado E recuperação de apgar >0 E Apgar 5º minuto <=7	Óbito evitável por adequado controle na gravidez	6.0	
2	Peso <=1105g E Malformação congênita = ignorado E recuperação de apgar <=0 E Apgar 5º minuto >7	Não óbito	6.0	
3	Peso <=1105g E Malformação congênita = sim E Idade materna <=31	Óbito evitável por parceria com outros setores	2.0	
4	Peso <=1105g E Malformação congênita = sim E Idade materna >31	Não evitável	3.0	
5	Peso <=1105g E Malformação congênita = não E Apgar no 5º minuto <=5	Óbito evitável por adequado controle na gravidez	43.26	20.26
6	Peso <=1105g E Malformação congênita = não E Apgar no 5º minuto >5 E Apgar no 1º minuto <=0	Óbito evitável por parceria com outros setores	2.01	0.01
7	Peso <=1105g E Malformação congênita = não E Apgar no 5º minuto >5 E Apgar no 1º minuto >0	Não óbito	118.72	37.0
8	Peso >1105g	Não óbito	24632.0	121.0

Fonte: A autora (2020).

Figura 38 – Árvore de decisão do ano de 2015 do experimento 1.

Regra	antecedente	consequente	cobertura	erro
1	Peso <=1500g E	Óbito evitável por adequado controle na gravidez	97.01	47.01
2	Peso <=1500g E Peso >780g	Não óbito	261.02	50.0
3	Peso >1500g E Apgar no 5ºminuto <=6 E Malformação congênita = não E Duração da gestação <=30 E Qtd óbito fetal/aborto <=0	Não óbito	3.0	2.0
4	Peso >1500g E Apgar no 5ºminuto <=6 E Malformação congênita = não E Duração da gestação <=30 semanas E Qtd óbito fetal/aborto >0	Óbito evitável por adequada atenção ao parto	2.0	
5	Peso >1500g E Apgar no 5ºminuto <=6 E Malformação congênita = não E Duração da gestação >30 semanas	Não óbito	134.24	13.0
6	SE Peso >1500g E Apgar no 5ºminuto <=6 E Malformação congênita = sim E Qtd parto cesáreo anterior <=0	ENTÃO Óbito evitável por parceria com outros setores	6.0	1.0
7	Peso >1500g E Apgar no 5ºminuto <=6 E Malformação congênita = sim E Qtd parto cesáreo anterior >0	Não evitável	2.0	0.0
8	Peso >1500g E Apgar no 5ºminuto <=6 E Malformação congênita = ignorado E Escolaridade materna <= superior completo	Óbito evitável por adequado controle na gravidez	2.0	1.0
	Peso >1500g E Apgar no 5ºminuto <=6 E Malformação congênita = ignorado E Escolaridade materna > superior completo	Não óbito	2.01	
8	Peso >1500g E Apgar no 5ºminuto >6	Não óbito	23996.72	76.0

Fonte: A autora (2020).

Figura 39 – Árvore de decisão do ano de 2016 do experimento 1.

Regra		antecedente		consequente	cobertura	erro
1	<b>SE</b>	Peso <=1615g E Duração da gestação <=31 semanas E Peso <=785g	<b>ENTÃO</b>	Óbito evitável por adequado controle na gravidez	75.32	33.02
2		Peso <=1615g E Duração da gestação <=31 semanas E Peso >785g		Não óbito	185.64	38.0
3		Peso <=1615g E Duração da gestação >31 semanas		Não óbito	140.14	9.7
4		Peso >1615g		Não óbito	22815.9	89.0

Fonte: A autora (2020).

Figura 40 – Árvore de decisão do ano de 2017 do experimento 1.

Regra	antecedente	consequente	cobertura	erro
1	Peso <= 1635g E Peso <=725g E classificação de risco ao nascer = risco	Óbito evitável por adequado controle na gravidez	75.0	31.0
2	Peso <= 1635g E Peso <=725g E classificação de risco ao nascer = não risco	Não óbito	5.0	1.0
3	Peso <= 1635g E Peso >725g	Não óbito	357.02	56.0
4	Peso >1635g E Malformação congênita = não	Não óbito	22153.69	37.0
5	Peso >1635g E Malformação congênita = sim E Apgar no 5ºminuto <=6	Óbito não evitável	18.01	10.01
6	Peso >1635g E Malformação congênita = sim E Apgar no 5ºminuto >6 E Peso ao nascer <=2125g E CBO materno <=999991	Óbito não evitável	4.0	0.0
7	Peso >1635g E Malformação congênita = sim E Apgar no 5ºminuto >6 E Peso ao nascer <=2125g E CBO materno >999991 E Peso ao nascer <=2005g	Não óbito	2.0	
8	Peso >1635g E Malformação congênita = sim E Apgar no 5ºminuto >6 E Peso ao nascer <=2125g E CBO materno >999991 E Peso ao nascer >2005g	Óbito evitável por parceria com outros setores	2.0	
9	Peso >1635g E Malformação congênita = sim E Apgar no 5ºminuto >6 E Peso ao nascer >2125g	Não óbito	75.18	2.0
10	Peso >1635g E Malformação congênita = ignorado	Não óbito	52.1	2.0

Fonte: A autora (2020).

## APÊNDICE G – ÁRVORES DE DECISÃO DO EXPERIMENTO 2

Figura 41 – Árvore de decisão do ano de 2002 do experimento 2.

Regra		antecedente		consequente	cobertura	erro
1	SE	idade da criança <=25 dias	ENTÃO	Óbito evitável por adequado controle na gravidez	222.0	139.0
2		idade da criança >25 dias		Óbito evitável por parceria com outros setores	104.0	60.0

Fonte: A autora (2020).

Figura 42 – Árvore de decisão do ano de 2003 do experimento 2.

Regra		antecedente		consequente	cobertura	erro
1	SE	classificação de risco ao nascer = NÃO RISCO	ENTÃO	Óbito evitável por parceria com outros setores	57.0	30.0
2		classificação de risco ao nascer = RISCO		Óbito evitável por adequado controle na gravidez	251.0	156.0

Fonte: A autora (2020).

Figura 43 – Árvore de decisão do ano de 2004 do experimento 2.

Regra		antecedente		consequente	cobertura	erro
1	SE	idade da criança <=0 mês completo	ENTÃO	Óbito evitável por adequado controle na gravidez	194.0	115.0
2		idade da criança >0 mês completo		Óbito evitável por parceria com outros setores	102.0	54.0

Fonte: A autora (2020).

Figura 44 – Árvore de decisão do ano de 2005 do experimento 2.

Regra		antecedente		consequente	cobertura	erro
1	SE	duração da gestação categorizada = 37-41 semanas	ENTÃO	Óbito evitável por parceria com outros setores	119.0	69.0
2		duração da gestação categorizada = 22-27 semanas		Óbito evitável por adequado controle na gravidez	82.0	27.0
3		duração da gestação categorizada = 28-31 semanas		Óbito evitável por adequado controle na gravidez	36.0	18.0
4		duração da gestação categorizada = <22 semanas		Óbito evitável por adequado controle na gravidez	4.0	1.0
5		duração da gestação categorizada = 32-36 semanas		Óbito evitável por parceria com outros setores	36.0	26.0

Fonte: A autora (2020).

Figura 45 – Árvore de decisão do ano de 2006 do experimento 2.

Regra		antecedente		consequente	cobertura	erro
1	SE	duração da gestação categorizada = 37-41 semanas	ENTÃO	Óbito evitável por parceria com outros setores	125.0	73.0
2		duração da gestação categorizada = 32-36 semanas		Óbito evitável por adequado controle na gravidez	45.0	31.0
3		duração da gestação categorizada = 22-27 semanas		Óbito evitável por adequado controle na gravidez	60.0	19.0
4		duração da gestação categorizada = 28-31 semanas		Óbito evitável por adequado controle na gravidez	30.0	13.0
5		duração da gestação categorizada = >42 semanas		Óbito evitável por parceria com outros setores	1.0	

Fonte: A autora (2020).

Figura 46 – Árvore de decisão do ano de 2008 do experimento 2.

Regra	SE	antecedente	ENTÃO	consequente	cobertura	erro
1		idade da criança <=0 mês completo		Óbito evitável por adequado controle na gravidez	170.0	78.0
2		idade da criança >0 mês completo		Óbito evitável por parceria com outros setores	90.0	49.0

Fonte: A autora (2020).

Figura 47 – Árvore de decisão do ano de 2009 do experimento 2.

Regra	SE	antecedente	ENTÃO	consequente	cobertura	erro
1		classificação de risco ao nascer = RISCO		Óbito evitável por adequado controle na gravidez	167.0	89.0
2		classificação de risco ao nascer = NÃO RISCO		Óbito evitável por parceria com outros setores	45.0	20.0

Fonte: A autora (2020).

Figura 48 – Árvore de decisão do ano de 2010 do experimento 2.

Regra	SE	antecedente	ENTÃO	consequente	cobertura	erro
1		classificação de risco ao nascer = RISCO		Óbito evitável por adequado controle na gravidez	194.0	102.0
2		classificação de risco ao nascer = NÃO RISCO		Óbito evitável por parceria com outros setores	40.0	17.0
3		classificação de risco ao nascer = IGNORADO		Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	

Fonte: A autora (2020).

Figura 49 – Árvore de decisão do ano de 2011 do experimento 2.

Regra	SE	antecedente	ENTÃO	consequente	cobertura	erro
1		idade da criança <=56 dias		Óbito evitável por adequado controle na gravidez	161.0	85.0
2		idade da criança >56 dias		Óbito evitável por parceria com outros setores	74.0	47.0

Fonte: A autora (2020).

Figura 50 – Árvore de decisão do ano de 2012 do experimento 2.

Regra	SE	antecedente	ENTÃO	consequente	cobertura	erro
1		classificação de risco ao nascer = RISCO		Óbito evitável por adequado controle na gravidez	178.0	94.0
2		classificação de risco ao nascer = NÃO RISCO		Óbito evitável por parceria com outros setores	45.0	24.0

Fonte: A autora (2020).

Figura 51 – Árvore de decisão do ano de 2014 do experimento 2.

Regra	SE	antecedente	ENTÃO	consequente	cobertura	erro
1		classificação de risco ao nascer = RISCO		Óbito evitável por adequado controle na gravidez	172.79	104.65
2		classificação de risco ao nascer = NÃO RISCO		Óbito evitável por parceria com outros setores	36.21	20.04

Fonte: A autora (2020).

Figura 52 – Árvore de decisão do ano de 2015 do experimento 2.

Regra	antecedente	consequente	cobertura	erro
1	idade da criança <=2 meses completos E classificação de risco ao nascer = RISCO	Óbito evitável por adequado controle na gravidez	160.0	71.0
2	idade da criança <=2 meses completos E classificação de risco ao nascer = NÃO RISCO E grupo de Robson <=8	Óbito evitável por parceria com outros setores	20.0	5.0
3	idade da criança <=2 meses completos E classificação de risco ao nascer = NÃO RISCO E grupo de Robson >8 E componente do óbito = NEOPRECOCE	Óbito evitável por adequada atenção ao parto	2.0	
4	idade da criança <=2 meses completos E classificação de risco ao nascer = NÃO RISCO E grupo de Robson >8 E componente do óbito = NEOTARDIO	Óbito evitável por prevenção, diagnóstico e tratamento precoces	2.0	1.0
5	idade da criança <=2 meses completos E local ocorrência do óbito = HOSPITAL E gestação categorizada = 37-41 semanas	Óbito evitável por prevenção, diagnóstico e tratamento precoces	23.0	11.0
6	idade da criança <=2 meses completos E local ocorrência do óbito = HOSPITAL E gestação categorizada = 28-311 semanas	Óbito evitável por prevenção, diagnóstico e tratamento precoces	1.0	
7	idade da criança <=2 meses completos E local ocorrência do óbito = HOSPITAL E gestação categorizada = 33-36 semanas	Óbito evitável por parceria com outros setores	8.0	2.0
8	idade da criança <=2 meses completos E local ocorrência do óbito = DOMICÍLIO	Óbito evitável por parceria com outros setores	3.0	
9	idade da criança <=2 meses completos E local ocorrência do óbito = OUTROS ESTAB DE SAÚDE	Óbito evitável por parceria com outros setores	4.0	1.0

Fonte: A autora (2020).

Figura 53 – Árvore de decisão do ano de 2016 do experimento 2.

Regra	antecedente	consequente	cobertura	erro
1	duração da gestação numérica <=32 semanas	Óbito evitável por adequado controle na gravidez	106.62	34.54
2	duração da gestação numérica >32 semanas	Óbito evitável por parceria com outros setores	91.38	41.38

Fonte: A autora (2020).

Figura 54 – Árvore de decisão do ano de 2017 do experimento 2.

Regra		antecedente		consequente	cobertura	erro
1	SE	duração da gestação numérica <=31 semanas	ENTÃO	Óbito evitável por adequado controle na gravidez	100.33	34.17
2		duração da gestação numérica >31 semanas		Óbito não evitável	71.67	46.25

Fonte: A autora (2020).



ANEXO A – DNV



República Federativa do Brasil  
Ministério da Saúde  
1ª VIA - SECRETARIA DE SAÚDE

**Declaração de Nascido Vivo**

<b>I</b>	<b>Identificação do Recém-nascido</b>	
	1 Nome do Recém-nascido	
	2 Data e hora do nascimento Data: _____ Hora: _____ 3 Sexo <input type="checkbox"/> M - Masculino <input type="checkbox"/> F - Feminino <input type="checkbox"/> I - Ignorado	
<b>II</b>	<b>Local da ocorrência</b>	
	4 Peso ao nascer _____ em gramas 5 Índice de Apgar _____ 1º minuto _____ 5º minuto _____ 6 Detectada alguma anomalia ou defeito congênito? Caso afirmativo, usar o bloco anomalia congênita para descrevê-las 1 <input type="checkbox"/> Sim 2 <input type="checkbox"/> Não 9 <input type="checkbox"/> Ignorado	
	7 Local da ocorrência <input type="checkbox"/> Hospital <input type="checkbox"/> Domicílio <input type="checkbox"/> Ignorado <input type="checkbox"/> Outros estab. saúde <input type="checkbox"/> Outros <input type="checkbox"/> 9 8 Estabelecimento _____ Código CNES _____ 9 Endereço da ocorrência, se fora do estab. ou da resid. da Mãe (rua, praça, avenida, etc) _____ Número _____ Complemento _____ 10 CEP _____ 11 Bairro/Distrito _____ Código _____ 12 Município de ocorrência _____ Código _____ 13 UF _____	
<b>III</b>	<b>Mãe</b>	
	14 Nome da Mãe _____ 15 Cartão SUS _____ 16 Escolaridade (última série concluída) Nível _____ Série _____ 17 Ocupação habitual (Informar anterior, se aposentada/desempregada) Código CBO 2002 _____ 0 <input type="checkbox"/> Sem escolaridade 3 <input type="checkbox"/> Médio (antigo 2º grau) <input type="checkbox"/> Ignorado 1 <input type="checkbox"/> Fundamental I (1ª a 4ª série) 4 <input type="checkbox"/> Superior incompleto 2 <input type="checkbox"/> Fundamental II (5ª a 8ª série) 5 <input type="checkbox"/> Superior completo 9 18 Data nascimento da Mãe _____ 19 Idade (anos) _____ 20 Naturalidade da Mãe _____ Município / UF (se estrangeiro informar País) _____ 21 Situação conjugal 1 <input type="checkbox"/> Solteira 4 <input type="checkbox"/> Separada judicialmente/divorciada 2 <input type="checkbox"/> Casada 5 <input type="checkbox"/> União estável 3 <input type="checkbox"/> Viúva 9 <input type="checkbox"/> Ignorada 22 Raça / Cor da Mãe 1 <input type="checkbox"/> Branca 4 <input type="checkbox"/> Parda 2 <input type="checkbox"/> Preta 5 <input type="checkbox"/> Indígena 3 <input type="checkbox"/> Amarela Residência da Mãe 23 Logradouro _____ Número _____ Complemento _____ 24 CEP _____ 25 Bairro/Distrito _____ Código _____ 26 Município _____ Código _____ 27 UF _____	
	28 Nome do Pai _____ 29 Idade do Pai _____	
<b>IV</b>	<b>Pai</b>	
	30 Histórico gestacional ■ Nº gestações anteriores _____ ■ Nº de partos vaginais _____ ■ Nº de cesáreas _____ ■ Nº de nascidos vivos _____ ■ Nº de perdas fetais / abortos _____	
	31 Data da Última Menstruação (DUM) ____/____/____ 32 Nº de semanas de gestação, se DUM ignorada _____ Método utilizado para estimar 1 <input type="checkbox"/> Exame Físico 2 <input type="checkbox"/> Outro método 9 <input type="checkbox"/> Ignorado 33 Número de consultas de pré-natal _____ 34 Mês de gestação em que iniciou o pré-natal _____ 35 Tipo de gravidez 1 <input type="checkbox"/> Única 2 <input type="checkbox"/> Dupla 3 <input type="checkbox"/> Tripla ou mais 9 <input type="checkbox"/> Ignorado 36 Apresentação 1 <input type="checkbox"/> Cefálica 2 <input type="checkbox"/> Pélvica ou Podálica 3 <input type="checkbox"/> Transversa 9 <input type="checkbox"/> Ignorado 37 O Trabalho de parto foi induzido? 1 <input type="checkbox"/> Sim 2 <input type="checkbox"/> Não 9 <input type="checkbox"/> Ignorado 38 Tipo de parto 1 <input type="checkbox"/> Vaginal 2 <input type="checkbox"/> Cesáreo 9 <input type="checkbox"/> Ignorado 39 Cesárea ocorreu antes do trabalho de parto iniciar? 1 <input type="checkbox"/> Sim 2 <input type="checkbox"/> Não 3 <input type="checkbox"/> Não se aplica 9 <input type="checkbox"/> Ignorado 40 Nascimento assistido por 1 <input type="checkbox"/> Médico 2 <input type="checkbox"/> Enfermeira/Obstetiz 3 <input type="checkbox"/> Parteira 4 <input type="checkbox"/> outros 9 <input type="checkbox"/> Ignorado	
<b>V</b>	<b>Gestação e parto</b>	
	41 Descrever todas as anomalias ou defeitos congênitos observados	
	42 Data do preenchimento _____ 43 Nome do responsável pelo preenchimento _____ 44 Função 1 <input type="checkbox"/> Médico 2 <input type="checkbox"/> Enfermeiro 3 <input type="checkbox"/> Parteira 4 <input type="checkbox"/> Func. Cartório 5 <input type="checkbox"/> Outros (descrever) _____ 45 Tipo documento 1 <input type="checkbox"/> CNES 2 <input type="checkbox"/> CRM 3 <input type="checkbox"/> COREN 4 <input type="checkbox"/> RG 5 <input type="checkbox"/> CPF _____ 46 Nº do documento _____ 47 Órgão emissor _____	
<b>VI</b>	<b>Anomalia congênita</b>	
	48 Cartório _____ Código _____ 49 Registro _____ 50 Data _____	
	51 Município _____ 52 UF _____	

**ATENÇÃO: ESTE DOCUMENTO NÃO SUBSTITUI A CERTIDÃO DE NASCIMENTO**  
O Registro de Nascimento é obrigatório por lei.  
Para registrar esta criança, o pai ou responsável deverá levar este documento ao cartório de registro civil.

ANEXO B – DO

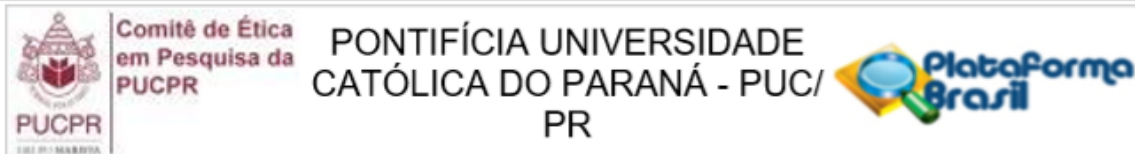

**República Federativa do Brasil**  
**Ministério da Saúde**  
 1ª VIA - SECRETARIA DE SAÚDE

**Declaração de Óbito**

<b>I</b>	1 Tipo de óbito		2 Data do óbito		Hora		3 Cartão SUS		4 Naturalidade	
	1 <input type="checkbox"/> Fetal 2 <input type="checkbox"/> Não fetal								MUNICÍPIO (UF) DE RESIDÊNCIA ANTERIOR (opcional)	
<b>II</b>	5 Nome do Falecido									
	6 Nome do Pai					7 Nome da Mãe				
<b>III</b>	8 Data de nascimento		9 Idade		10 Sexo		11 Raça/Cor		12 Situação conjugal	
	13 Escolaridade (última série concluída)		14 Matão (antigo 2º grau)		15 Superior incompleto		16 Superior completo		17 Ocupação habitual	
<b>IV</b>	18 Logradouro (rua, praça, avenida etc.)									
	19 Bairro/Distrito		20 Código		21 Município de residência		22 Código		23 UF	
<b>V</b>	24 Local de ocorrência do óbito		25 Estabelecimento		26 Código CNES					
	27 Endereço da ocorrência, se fora do estabelecimento ou da residência (rua, praça, avenida, etc.)									
<b>VI</b>	28 Bairro/Distrito		29 Código		30 Município de ocorrência		31 Código		32 UF	
	<b>PREENCHIMENTO EXCLUSIVO PARA ÓBITOS FETAIS E DE MEMORES DE 1 ANO - INFORMAÇÕES SOBRE A MÃE</b> 33 Idade (anos)    34 Escolaridade (última série concluída)    35 Sexo    36 Ocupação habitual 37 Número de filhos vivos    38 Nº de semanas de gestação    39 Tipo de gravidez    40 Tipo de parto    41 Morte em relação ao parto    42 Peso ao nascer    43 Número da Declaração de Nascimento Vivo									
<b>VII</b>	<b>ÓBITO DE MULHER EM IDADE FÉRTIL</b> 44 A morte ocorreu: 1 No gravidez, 2 No parto, 3 De 43 dias a 1 ano após o parto, 4 No parto, 5 Não ocorreu nestes períodos.									
	<b>ASSISTÊNCIA MÉDICA</b> 45 Recebeu assistência médica durante a doença que ocasionou a morte? 1 Sim, 2 Não, 3 Ignorado.									
<b>VIII</b>	<b>CAUSAS DA MORTE</b> PARTE I: Descrição ou estado clínico que ocorreu imediatamente a morte. CAUSAS ANTERIORES: Estados pré-existentes, se existentes, que produziram a causa acima registrada, mencionando-se em último lugar a causa letal.									
	ANOTE SOMENTE UM DIAGNÓSTICO POR LINHA. Devido ao como consequência de: a) _____ b) _____ c) _____ d) _____ PARTE II: Outras condições significativas que contribuíram para a morte, e que não estariam, porém, na cadeia acima.									
<b>IX</b>	46 Nome do Médico		47 CRM		48 Óbito atestado por Médico		49 Município e UF de SVO ou IML			
	50 Meio de contato (telefone, fax, e-mail etc.)		51 Data do atestado		52 Assinatura					
<b>X</b>	<b>PROVÁVEIS CIRCUNSTÂNCIAS DE MORTE NÃO NATURAL (informações de caráter estritamente epidemiológico)</b> 53 Tipo: 1 Acidente, 2 Suicídio, 3 Homicídio, 4 Outras.									
	54 Descrição sucinta do evento, incluindo o tipo de local de ocorrência.									
<b>XI</b>	55 Se a ocorrência for em via pública, ANOTAR O ENDEREÇO: 56 Logradouro (rua, praça, avenida, etc.)									
	57 Cartório		58 Código		59 Registro		60 Data			
<b>XII</b>	61 Município		62 Código		63 UF					
	64 Declaração					65 Testemunhas				

Modelo C-110 - 7ª edição 01/2010

## ANEXO C – PARECER CONSUBSTANCIADO DO CEP DA PUCPR



### PARECER CONSUBSTANCIADO DO CEP

#### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** COOSMIC Coorte de Saúde Materno-Infantil de Curitiba

**Pesquisador:** Simone Tetu Moysés

**Área Temática:** Genética Humana:

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP.);

**Versão:** 1

**CAAE:** 90052118.9.0000.0020

**Instituição Proponente:** Pontifícia Universidade Católica do Parana - PUCPR

**Patrocinador Principal:** Financiamento Próprio

#### DADOS DO PARECER

**Número do Parecer:** 2.672.385

#### Apresentação do Projeto:

O cuidado com a saúde de mães e crianças durante os primeiros anos de vida é considerado uma estratégia chave para o combate a iniquidades

em saúde, reduzindo a carga de doenças e de mortes preveníveis e/ou evitáveis. Deve constituir, portanto, prioridade em políticas de saúde pública

e de proteção social em qualquer contexto populacional, independentemente de tempo e lugar. Determinantes socioambientais e fatores biológicos

desde o período fetal e a primeira infância, em especial nos primeiros 1000 dias, afetam a vida e a saúde com repercussões percebidas na fase

adulta e na terceira idade. Estudos longitudinais de coorte oferecem oportunidade para a produção de conhecimento dinâmico sobre iniquidades em

saúde, diagnóstico de riscos e chances individuais e comunitárias, identificação de doenças e pistas de causalidade, além da avaliação de impacto

de estratégias de cuidado. A produção de conhecimento epidemiológico relevante em saúde materno-infantil, por meio da análise complexa dos

dados disponíveis no sistema de informação de serviços de saúde, e com base no acompanhamento de gestantes e seus filhos durante os primeiros

mil dias (até os dois anos de idade), pode apoiar a definição de políticas e ações intersetoriais de

**Endereço:** Rua Imaculada Conceição 1155

**Bairro:** Prado Velho

**CEP:** 80.215-901

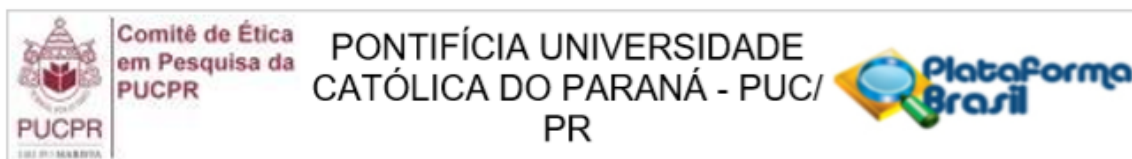
**UF:** PR

**Município:** CURITIBA

**Telefone:** (41)3271-2103

**Fax:** (41)3271-2103

**E-mail:** nep@pucpr.br



Continuação do Parecer: 2.672.385

saúde pública que impactem positivamente a proteção e a saúde de mães e crianças. Considerando a expertise de pesquisadores da PUCPR e de profissionais da área técnica da Secretaria Municipal da Saúde, bem como a existência de estrutura necessária para o desenvolvimento de uma pesquisa operacional multidisciplinar, surge como oportuna a proposta de estruturação de um estudo de coorte mista, com gestantes e crianças até os dois anos de idade em Curitiba vinculadas à Rede Mãe Curitiba, alinhado à oferta paralela de serviços especializados para atendimento dessas mulheres e seus filhos.

#### **Objetivo da Pesquisa:**

Objetivo Primário:

Investigar a saúde de gestantes e crianças nos primeiros 1000 dias de vida, vinculadas à Rede Mãe Curitiba.

Objetivo Secundário:

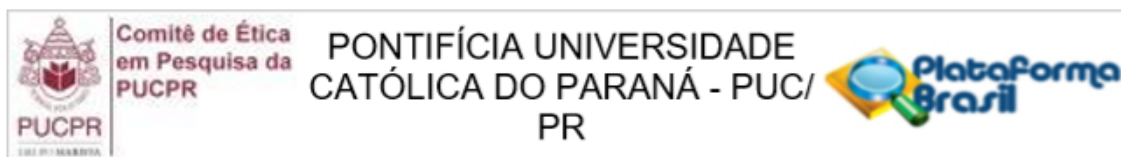
1. Explorar as condições de saúde e seus determinantes, na população de gestantes e crianças até dois anos. 2. Analisar desfechos na saúde de gestantes e crianças até dois anos e fatores associados tais como determinantes biológicos, comportamentais, ambientais, sociais e econômicos. 3. Identificar fatores preditores de eventos de saúde/doença/cuidado ao longo dos primeiros 1000 dias de vida das crianças. 4. Investigar vulnerabilidades que acometem o grupo materno-infantil pesquisado. 5. Promover o acompanhamento e cuidado em saúde da população do estudo, por meio de abordagem multidisciplinar. 6. Avaliar o impacto das intervenções multidisciplinares implementadas durante o estudo, sobre a condição de saúde de gestantes e crianças até dois anos. 7. Apoiar o planejamento e gestão de políticas de atenção materno-infantil em Curitiba, ampliando a qualificação de equipes de saúde e de pesquisadores para a atenção materno-infantil. 8. Criar um centro de excelência na atenção materno-infantil vinculado à Rede Mãe Curitiba, como referência para populações com características similares, seja no Brasil ou em outros países.

#### **Avaliação dos Riscos e Benefícios:**

Riscos:

Alguns desconfortos ou riscos podem ser esperados para a participação das gestantes. Isto pode

Endereço: Rua Imaculada Conceição 1155  
 Bairro: Prado Velho CEP: 80.215-901  
 UF: PR Município: CURITIBA  
 Telefone: (41)3271-2103 Fax: (41)3271-2103 E-mail: nep@pucpr.br



Continuação do Parecer: 2.672.385

incluir constrangimentos para responder a algumas perguntas do questionário, ou algum desconforto durante a realização dos exames, como cansaço para realização de algumas medidas corporais. Para diminuir tais desconfortos e riscos serão tomadas as seguintes medidas: em caso de constrangimento para responder perguntas do questionário, será garantido o direito das gestantes em não responder ou interromper sua participação a qualquer momento; os exames serão realizados em local apropriado, em ambiente confortável e privado, conduzidos por examinadores treinados para agilizar os procedimentos necessários.

**Benefícios:**

O benefício direto para as participantes será a ampliação do conhecimento sobre sua saúde durante a gestação e, em uma fase posterior do estudo, acompanhar o desenvolvimento e a saúde de seu(s) filho(s) até os dois anos de idade, com o suporte de uma equipe de profissionais de odontologia, psicologia, nutrição, fisioterapia, farmácia e educação física. Além disso, benefícios indiretos também poderão ser esperados como o aumento

**Comentários e Considerações sobre a Pesquisa:**

Pesquisa muito bem estrutura com todos os documentos anexados corretamente e muito bem organizada.

**Considerações sobre os Termos de apresentação obrigatória:**

Todos documentos obrigatórios anexados corretamente

**Recomendações:**

sem comentários adicionais

**Conclusões ou Pendências e Lista de Inadequações:**

Aprovado

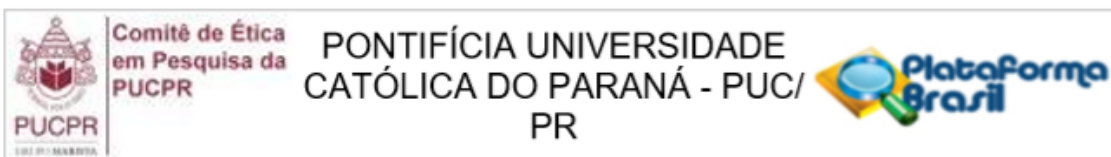
**Considerações Finais a critério do CEP:**

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
----------------	---------	----------	-------	----------

Endereço: Rua Imaculada Conceição 1155  
 Bairro: Prado Velho CEP: 80.215-901  
 UF: PR Município: CURITIBA  
 Telefone: (41)3271-2103 Fax: (41)3271-2103 E-mail: nep@pucpr.br





Continuação do Parecer: 2.672.385

Informações Básicas do Projeto	PB_INFORMAÇÕES_BASICAS_DO_PROJETO_1139359.pdf	19/05/2018 17:32:23		Aceito
Projeto Detalhado / Brochura	PROJETOCOOSMICmai2018.pdf	19/05/2018 17:31:31	Simone Tetu Moysés	Aceito
Investigador				
Declaração de Instituição e Infraestrutura	AutorizacaoLEM.pdf	19/05/2018 17:30:45	Simone Tetu Moysés	Aceito
Outros	RequerimentodeApreciacaodePesquisa.pdf	18/05/2018 19:24:10	Simone Tetu Moysés	Aceito
Outros	POPQuestionariofrequenciaalimentar.pdf	18/05/2018 19:23:24	Simone Tetu Moysés	Aceito
Outros	POPPressaoarterial.pdf	18/05/2018 19:23:09	Simone Tetu Moysés	Aceito
Outros	POPExperienciadecarieindiceCPOD.pdf	18/05/2018 19:22:52	Simone Tetu Moysés	Aceito
Outros	POPEdema.pdf	18/05/2018 19:22:36	Simone Tetu Moysés	Aceito
Outros	POPCondicaoperiodontal.pdf	18/05/2018 19:22:17	Simone Tetu Moysés	Aceito
Outros	POPColetadematerialbiologico.pdf	18/05/2018 19:22:00	Simone Tetu Moysés	Aceito
Outros	POPAntropometria.pdf	18/05/2018 19:21:44	Simone Tetu Moysés	Aceito
Outros	Manualdoentrevistador.pdf	18/05/2018 19:20:31	Simone Tetu Moysés	Aceito
Outros	POPAlteracoesestomatologicas.pdf	18/05/2018 19:19:36	Simone Tetu Moysés	Aceito
Outros	Instrumentoavaliacaoclinica.pdf	18/05/2018 19:18:46	Simone Tetu Moysés	Aceito
Outros	QuestionarioGestantesPreNatal.pdf	18/05/2018 19:18:18	Simone Tetu Moysés	Aceito
Declaração de Pesquisadores	TermodenaocustoSMS.pdf	18/05/2018 19:17:14	Simone Tetu Moysés	Aceito
Declaração de Pesquisadores	TCUD.pdf	18/05/2018 19:16:46	Simone Tetu Moysés	Aceito
Declaração de Instituição e Infraestrutura	DeclaracaocienciaeinteresseSMS.pdf	18/05/2018 19:16:23	Simone Tetu Moysés	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLEmenores18anospaisresponsavelCOOSMICmai2018.pdf	18/05/2018 19:15:36	Simone Tetu Moysés	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLEparticipantes16_18_anos_COOSMICmai2018.pdf	18/05/2018 19:14:32	Simone Tetu Moysés	Aceito

Endereço: Rua Imaculada Conceição 1155

Bairro: Prado Velho

CEP: 80.215-901

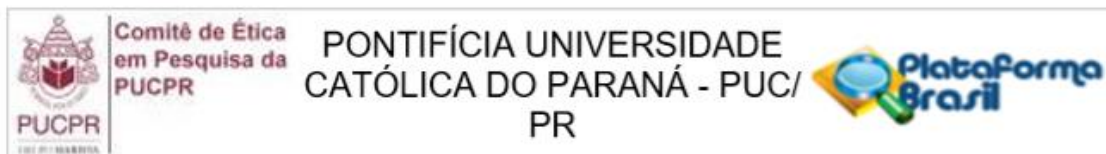
UF: PR

Município: CURITIBA

Telefone: (41)3271-2103

Fax: (41)3271-2103

E-mail: nep@pucpr.br



Continuação do Parecer: 2.672.385

TCLE / Termos de Assentimento / Justificativa de Ausência	TCLEparticipantesmaiores18_anosCOO SMICmai2018.pdf	18/05/2018 19:14:04	Simone Tetu Moysés	Aceito
Folha de Rosto	Folhaderosto.pdf	18/05/2018 19:09:00	Simone Tetu Moysés	Aceito

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

CURITIBA, 24 de Maio de 2018

---

**Assinado por:**  
**NAIM AKEL FILHO**  
**(Coordenador)**

Endereço: Rua Imaculada Conceição 1155  
 Bairro: Prado Velho CEP: 80.215-901  
 UF: PR Município: CURITIBA  
 Telefone: (41)3271-2103 Fax: (41)3271-2103 E-mail: nep@pucpr.br

## ANEXO D – PARECER CONSUBSTANCIADO DO CEP DA SMS DE CURITIBA

SECRETARIA MUNICIPAL DA  
SAÚDE DE CURITIBA - SMS



### PARECER CONSUBSTANCIADO DO CEP

Elaborado pela Instituição Coparticipante

#### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** COOSMIC Coorte de Saúde Materno-Infantil de Curitiba

**Pesquisador:** Simone Tetu Moysés

**Área Temática:** Genética Humana;

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP:);

**Versão:** 1

**CAAE:** 90052118.9.3001.0101

**Instituição Proponente:** Secretaria Municipal da Saúde de Curitiba

**Patrocinador Principal:** Financiamento Próprio

#### DADOS DO PARECER

**Número do Parecer:** 2.728.771

#### Apresentação do Projeto:

Projeto protocolado no CEP-SMS/Curitiba sob nº 59-2018.

Estudos longitudinais de coorte oferecem oportunidade para a produção de conhecimento dinâmico, histórico e/ou prospectivo, sobre situações de saúde, exposições, riscos e fatores associados, bem como iniquidades em saúde e chances individuais e comunitárias, identificação de doenças e pistas de causalidade, além da avaliação de impacto de estratégias de cuidado. A produção de conhecimento epidemiológico relevante em saúde materno-infantil, por meio da análise complexa dos dados disponíveis no sistema de informação de serviços de saúde, e com base no acompanhamento de gestantes e seus filhos até os dois anos de idade, pode apoiar a definição de políticas e ações intersetoriais de saúde pública que impactem positivamente a proteção e a saúde de mães e crianças. Considerando a expertise de pesquisadores da PUCPR e de profissionais da área técnica da Secretaria Municipal da Saúde, bem como a existência de estrutura necessária para o desenvolvimento de uma pesquisa operacional multidisciplinar, surge como oportuna a proposta de estruturação de um estudo de coorte mista, com gestantes e crianças até os dois anos de idade em Curitiba, alinhado à oferta paralela de serviços especializados para atendimento dessas mulheres e seus filhos.

#### Objetivo da Pesquisa:

Objetivo geral:

**Endereço:** Rua Atilio Bório, 680

**Bairro:** Cristo Rei

**CEP:** 80.050-250

**UF:** PR

**Município:** CURITIBA

**Telefone:** (41)3360-4961

**Fax:** (41)3360-4965

**E-mail:** etica@sms.curitiba.pr.gov.br



SECRETARIA MUNICIPAL DA  
SAÚDE DE CURITIBA - SMS



Continuação do Parecer: 2.728.771

Investigar a saúde de gestantes e crianças nos primeiros 1000 dias de vida, vinculadas à Rede Mãe Curitibana.

Objetivos específicos:

1. Explorar as condições de saúde e seus determinantes, na população de gestantes e crianças até dois anos.
2. Analisar desfechos na saúde de gestantes e crianças até dois anos e fatores associados tais como determinantes biológicos, comportamentais, ambientais, sociais e econômicos.
3. Identificar fatores preditores de eventos de saúde/doença/cuidado ao longo dos primeiros 1000 dias de vida das crianças.
4. Investigar vulnerabilidades que acometem o grupo materno-infantil pesquisado.
5. Promover o acompanhamento e cuidado em saúde da população do estudo, por meio de abordagem multidisciplinar.
6. Avaliar o impacto das intervenções multidisciplinares implementadas durante o estudo, sobre a condição de saúde de gestantes e crianças até dois anos.
7. Apoiar o planejamento e gestão de políticas de atenção materno-infantil em Curitiba, ampliando a qualificação de equipes de saúde e de pesquisadores para a atenção materno-infantil.
8. Criar um centro de excelência na atenção materno-infantil vinculado à Rede Mãe Curitibana, como referência para populações com características similares, seja no Brasil ou em outros países.

**Avaliação dos Riscos e Benefícios:**

**Riscos:** Alguns desconfortos ou riscos podem ser esperados para a participação das gestantes. Isto pode incluir constrangimentos para responder a algumas perguntas do questionário, ou algum desconforto durante a realização dos exames, como cansaço para realização de algumas medidas corporais. Para diminuir tais desconfortos e riscos serão tomadas as seguintes medidas: em caso de constrangimento para responder perguntas do questionário, será garantido o direito das gestantes em não responder ou interromper sua participação a qualquer momento; os exames serão realizados em local apropriado, em ambiente confortável e privativo, conduzidos por examinadores treinados para agilizar os procedimentos necessários.

**Benefícios:-**

O benefício direto para as participantes será a ampliação do conhecimento sobre sua saúde durante a gestação e, em uma fase posterior do estudo, acompanhar o desenvolvimento e a saúde de seu(s) filho(s) até os dois anos de idade, com o suporte de uma equipe de profissionais de odontologia, psicologia, nutrição, fisioterapia, farmácia e educação física. Além disso, benefícios indiretos também poderão ser esperados como o aumento da qualidade da atenção em saúde

**Endereço:** Rua Atilio Bório, 680

**Bairro:** Cristo Rei

**CEP:** 80.050-250

**UF:** PR

**Município:** CURITIBA

**Telefone:** (41)3360-4961

**Fax:** (41)3360-4965

**E-mail:** [etica@sms.curitiba.pr.gov.br](mailto:etica@sms.curitiba.pr.gov.br)

SECRETARIA MUNICIPAL DA  
SAÚDE DE CURITIBA - SMS

Continuação do Parecer: 2.726.771

oferecida pela Rede Mãe Curitibana, contribuindo para a melhoria da saúde de gestantes e crianças de Curitiba.

**Comentários e Considerações sobre a Pesquisa:**

O projeto recebeu parecer técnico de viabilidade pela coordenação do Programa Mãe Curitibana, bem como respaldo político-gerencial das autoridades competentes na SMS-Curitiba. Projeto relevante para o desenvolvimento de boas práticas em Curitiba, permitindo desenvolver acompanhamento com metodologia robusta e sólida base teórica.

**Considerações sobre os Termos de apresentação obrigatória:**

Termos apresentados estando em conformidade às recomendações da Res. CNS 466/12.

**Recomendações:**

Em cumprimento à Resolução CNS 466/12, este Comitê de Ética em Pesquisa (CEP) deverá receber relatórios parciais sobre o andamento do estudo, bem como o relatório completo ao final do estudo. Eventuais notificações, ou modificações que gerem emendas ao protocolo original, devem ser apresentadas prontamente, identificando a parte do protocolo a ser modificado e as suas justificativas. Salientamos a necessidade de entrar previamente em contato com nossas Unidades ou Equipes, de posse do Termos de Aprovação da Pesquisa, para agendar as atividades necessárias com as Chefias locais.

**Conclusões ou Pendências e Lista de Inadequações:**

Projeto encontra-se em adequação à Res. CNS 466/12.

**Considerações Finais a critério do CEP:**

O Colegiado do CEP/SMS-Curitiba ratifica o parecer do(s) relator(es).

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Projeto Detalhado / Brochura Investigador	PROJETOCOOSMICmai2018.pdf	19/05/2018 17:31:31	Simone Tetu Moysés	Aceito
Outros	RequerimentodeApreciacaodePesquisa.pdf	18/05/2018 19:24:10	Simone Tetu Moysés	Aceito
Outros	POPQuestionariofrequeciaalimentar.pdf	18/05/2018 19:23:24	Simone Tetu Moysés	Aceito
Outros	POPPressaoarterial.pdf	18/05/2018 19:23:09	Simone Tetu Moysés	Aceito
Outros	POPExperienciadecarieindiceCPOD.	18/05/2018	Simone Tetu	Aceito

**Endereço:** Rua Atílio Bório, 680

**Bairro:** Cristo Rei

**CEP:** 80.050-250

**UF:** PR

**Município:** CURITIBA

**Telefone:** (41)3360-4961

**Fax:** (41)3360-4965

**E-mail:** etica@sms.curitiba.pr.gov.br

SECRETARIA MUNICIPAL DA  
SAÚDE DE CURITIBA - SMS

Continuação do Parecer: 2.726.771

Outros	pdf	19:22:52	Moysés	Aceito
Outros	POPEdema.pdf	18/05/2018 19:22:38	Simone Tetu Moysés	Aceito
Outros	POPCondiçaooperiodontal.pdf	18/05/2018 19:22:17	Simone Tetu Moysés	Aceito
Outros	POPColetadematerialbiologico.pdf	18/05/2018 19:22:00	Simone Tetu Moysés	Aceito
Outros	POPAntropometria.pdf	18/05/2018 19:21:44	Simone Tetu Moysés	Aceito
Outros	Manualdoentrevistador.pdf	18/05/2018 19:20:31	Simone Tetu Moysés	Aceito
Outros	POPAlteracoesestomatologicas.pdf	18/05/2018 19:19:38	Simone Tetu Moysés	Aceito
Outros	Instrumentoavaliacaoclinica.pdf	18/05/2018 19:18:48	Simone Tetu Moysés	Aceito
Outros	QuestionarioGestantesPreNatal.pdf	18/05/2018 19:18:18	Simone Tetu Moysés	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLEmenores18anospaisresponsavelCOOSMICmai2018.pdf	18/05/2018 19:15:38	Simone Tetu Moysés	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLEparticipantes16_18_anos_COOSMICmai2018.pdf	18/05/2018 19:14:32	Simone Tetu Moysés	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLEparticipantesmaiores18_anosCOOSMICmai2018.pdf	18/05/2018 19:14:04	Simone Tetu Moysés	Aceito

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

CURITIBA, 21 de Junho de 2018

Assinado por:  
**SAMUEL JORGE MOYSÉS**  
(Coordenador)

**Endereço:** Rua Atílio Bório, 680**Bairro:** Cristo Rei**CEP:** 80.050-250**UF:** PR**Município:** CURITIBA**Telefone:** (41)3360-4961**Fax:** (41)3360-4965**E-mail:** etica@sms.curitiba.pr.gov.br