

**UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE MEDICINA
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE**

**CRIAÇÃO DE UM MODELO MATEMÁTICO BASEADO EM MÚLTIPLAS
VARIÁVEIS AMBIENTAIS PARA ESTIMAR MORTALIDADE EM SAÚDE PÚBLICA**

ANAÍSA FILMIANO ANDRADE LOPES

**UBERLÂNDIA/MG
2022**

ANAÍSA FILMIANO ANDRADE LOPES

**CRIAÇÃO DE UM MODELO MATEMÁTICO BASEADO EM MÚLTIPLAS
VARIÁVEIS AMBIENTAIS PARA ESTIMAR MORTALIDADE EM SAÚDE PÚBLICA**

**Tese de Doutorado apresentada ao
Programa de Pós-Graduação em Ciências
da Saúde da Faculdade de Medicina da
Universidade Federal de Uberlândia, como
requisito parcial para a obtenção do título
de Doutora em Ciências da Saúde.**

Área de Concentração: Ciências da Saúde

**Orientador: Prof. Dr. Morun Bernardino
Neto**

**UBERLÂNDIA/MG
2022**

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU
com dados informados pelo(a) próprio(a) autor(a).

L864
2022

Lopes, Anaísa Filmiano Andrade, 1992-
Criação de um modelo matemático baseado em múltiplas
variáveis ambientais para estimar mortalidade em saúde
pública [recurso eletrônico] / Anaísa Filmiano Andrade
Lopes. - 2022.

Orientador: Morun Bernardino Neto.
Tese (Doutorado) - Universidade Federal de Uberlândia,
Pós-graduação em Ciências da Saúde.
Modo de acesso: Internet.
Disponível em: <http://doi.org/10.14393/ufu.te.2023.30>
Inclui bibliografia.
Inclui ilustrações.

1. Ciências médicas. I. Bernardino Neto, Morun ,1974-
(Orient.). II. Universidade Federal de Uberlândia. Pós-
graduação em Ciências da Saúde. III. Título.

CDU: 61

Bibliotecários responsáveis pela estrutura de acordo com o AACR2:
Gizele Cristine Nunes do Couto - CRB6/2091
Nelson Marcos Ferreira - CRB6/3074



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
 Coordenação do Programa de Pós-Graduação em Ciências da Saúde
 Av. Pará, 1720, Bloco 2H, Sala 11 - Bairro Umarama, Uberlândia-MG, CEP 38400-902
 Telefone: (34) 3225-8628 - www.ppcsafamed.ufu.br - ppcsaf@famed.ufu.br



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciências da Saúde				
Defesa de:	Tese de Doutorado Nº 021/PPCSA				
Data:	27.10.2022	Hora de início:	08:00h	Hora de encerramento:	11:00h
Matrícula do Discente:	11813CSD006				
Nome do Discente:	Anaísa Filmiano Andrade Lopes				
Título do Trabalho:	Criação de um modelo matemático baseado em múltiplas variáveis ambientais para estimar mortalidade em saúde pública				
Área de concentração:	Ciências da Saúde				
Linha de pesquisa:	1: EPIDEMIOLOGIA DA OCORRÊNCIA DE DOENÇAS E AGRAVOS À SAÚDE				
Projeto de Pesquisa de vinculação:	EPIDEMIOLOGIA E PREVENÇÃO DOS AGRAVOS EXTERNOS À SAÚDE				

Reuniu-se em web conferência pela plataforma Google Meet, em conformidade com a PORTARIA Nº 36, DE 19 DE MARÇO DE 2020 da COORDENAÇÃO DE APERFEIÇOAMENTO DE PESSOAL DE NÍVEL SUPERIOR - CAPES, pela Universidade Federal de Uberlândia, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciências da Saúde, assim composta: Profs. Drs. Marco Aurélio Kondracki de Alcântara (USP), Teresa Cristina Brazil de Paiva (USP), Jean Ezequiel Limongi (UFU), Winston Kleiber de Almeida Bacelar (UFU) e Morun Bernardino Neto (USP), orientador da candidata.

Iniciando os trabalhos, o presidente da mesa, Prof. Dr. Morun Bernardino Neto, apresentou a Comissão Examinadora e a candidata, agradeceu a presença do público, e concedeu a Discente a palavra para a exposição do seu trabalho. A duração da apresentação da Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(às) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovada.

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Marco Aurélio Kondracki de Alcântara, Usuário Externo**, em 27/10/2022, às 11:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.



Documento assinado eletronicamente por **Jean Ezequiel Limongi, Professor(a) do Magistério Superior**, em 27/10/2022, às 11:59, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.



Documento assinado eletronicamente por **Teresa Cristina Brazil de Paiva, Usuário Externo**, em 27/10/2022, às 12:03, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.



Documento assinado eletronicamente por **Morun Bernardino Neto, Usuário Externo**, em 27/10/2022, às 17:00, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.



Documento assinado eletronicamente por **Winston Kleiber de Almeida Bacelar, Professor(a) do Magistério Superior**, em 03/11/2022, às 15:28, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do Decreto nº 8.539, de 8 de outubro de 2015.



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **4031733** e o código CRC **D711B9FA**.

FOLHA DE APROVAÇÃO

ANAÍSA FILMIANO ANDRADE LOPES

CRIAÇÃO DE UM MODELO MATEMÁTICO BASEADO EM MÚLTIPLAS VARIÁVEIS AMBIENTAIS PARA ESTIMAR MORTALIDADE EM SAÚDE PÚBLICA

Presidente da banca: Prof. Dr. Morun Bernardino Neto

Tese de doutorado apresentada ao Programa de Pós-Graduação em Ciências da Saúde da Faculdade de Medicina da Universidade Federal de Uberlândia como requisito parcial para obtenção do título Doutora em Ciências da Saúde.

Área de concentração: Ciências da Saúde

Banca Examinadora

Titular: Prof^o Dr^o Jean Ezequiel Limongi

Instituição: Universidade Federal de Uberlândia

Titular: Prof^o Dr^o Winston Kleiber de Almeida Bacelar

Instituição: Universidade Federal de Uberlândia

Titular: Prof^o Dr^o Marco Aurélio Kondracki de Alcântara

Instituição: Universidade de São Paulo

Titular: Prof^a Dr^a Teresa Cristina Brazil de Paiva

Instituição: Universidade de São Paulo

Suplente: Prof^o Dr^o Paulo Cezar Mendes

Instituição: Universidade Federal de Uberlândia

Suplente: Prof^o Dr^o Fabiano Fernandes Bargas

Instituição: Universidade de São Paulo

DEDICATÓRIA

Aos sanitaristas, à comunidade acadêmica, médicos, enfermeiros, agentes de saúde, governantes e a todos que tem comprometimento com o meio ambiente e com a saúde pública.

AGRADECIMENTOS

Primeiramente, agradeço a Deus por tudo, pelo dom da vida, por permitir que eu chegasse até aqui, pelo amparo e força concedida durante todos os momentos. Obrigada meu Deus por me fazer acreditar, ter fé e por colocar pessoas tão iluminadas no meu caminho.

Aos meus pais Renata Maria Filmiano e Jeverson Andrade Lopes que são minha base, que se dedicaram e nunca mediram esforços para me proporcionar uma boa educação. Tudo o que sou e conquistei devo a vocês! Ao meu irmão Alisson Filmiano Andrade pelos incentivos fornecidos em meus estudos, pelo suporte necessário e por todos os valores que construímos juntos.

À minha pequena filha Cecília Andrade Magalhães que chegou durante o doutorado e que despertou em mim uma força que eu jamais imaginava ter! Obrigada minha filha por ser minha mais pura fonte de ternura e amor.

Ao meu esposo Lucas Silva Magalhães pelo apoio, conselhos, pelas palavras de conforto nas horas difíceis e pelas ajudas com nossa filha durante os momentos em que precisei. Obrigada por seu amor, paciência e companheirismo constante.

Aos meus avôs paternos: José Joaquim Lopes e Ismalia Andrade Lopes, e maternos: Adair Filmiano e Maria Célia Cardoso, obrigada pelas suas bênçãos, por serem meu suporte e me guiarem no caminho do bem. Vocês são meus exemplos de vida e como é bom tê-los comigo!

Ao tio Cleber Andrade e à tia Patrícia Filmiano sou grata por me auxiliarem e estarem ao meu lado em tudo que precisei. Vocês são muito especiais! E aos demais familiares, agradeço pelo estímulo e pela força que me foram proporcionados.

Aos meus sogros Waldete Magalhaes e Adenor Silva pelo apoio e pelos dias que passaram comigo me ajudando enquanto eu estudava e escrevia a tese. Nunca irei esquecer-me de todas as ajudas fornecidas por vocês.

Meus sinceros agradecimentos ao professor Morun Bernardino Neto, pela oportunidade de orientação, pela sua sabedoria e humildade em ensinar. Sou grata pelos

direcionamentos dados a essa pesquisa, por todos os esclarecimentos de dúvidas, pelos valiosos ensinamentos de estatística para a área da saúde e pela confiança depositada em mim.

Obrigada também a todos os professores que fazem parte do corpo docente do Programa de Pós-Graduação em Ciências da Saúde - PPCSA/UFU. Gratidão pelas proveitosas aulas, pelos conhecimentos compartilhados, pelas experiências e aprendizados concedidos no meio acadêmico.

À coordenadora da pós-graduação professora Yara Cristina de Paiva Maia pelas contribuições fornecidas na elaboração dos artigos científicos resultantes desse estudo. E às secretarias Gisele de Melo Rodrigues e Viviane Pires por toda a atenção e informes dados quando precisei.

Aos professores Jean Ezequiel Limongi e Teresa Cristina Brazil de Paiva pelas valiosas contribuições, auxílios e sugestões à pesquisa realizadas durante o exame de qualificação.

Aos amigos que a vida me deu Lorryne Britto e Luís Paulo Pires por todas as ajudas concedidas, pela disponibilidade de estarem presentes sempre quando precisei e pelos incentivos fornecidos. E aos novos amigos que a pós-graduação me presenteou e que compartilharam comigo trocas de conhecimentos, experiências, ajudas e conselhos, principalmente à Lorena Vargas e à Flávia Araújo. O caminho se torna leve com as mãos amigas!

À Universidade Federal de Uberlândia e à Faculdade de Medicina FAMED/UFU por me acolher e por proporcionar o meu ensino gratuito e de qualidade. Quanta honra ter sido discente dessa instituição e vivenciar tantos aprendizados!

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) por me conceder uma bolsa de estudos durante o doutorado.

Enfim, meu eterno reconhecimento a todos que contribuíram para o desenvolvimento dessa tese e que torceram pela realização dessa conquista! Gratidão a cada um de vocês!

“É hora de encerrar ciclos e terminar capítulos para que outros possam surgir!”

“O conhecimento da ciência começa quando podemos medir o que está se falando e expressar isso em números”.

Lord Kelvin

RESUMO

Introdução: Um dos desafios que a gestão pública vem enfrentando é o de estruturar sistemas de vigilância voltados para a mudança de contextos ambientais que representem situações de risco e desfechos críticos à saúde humana. **Objetivo:** Diante disso, essa pesquisa teve como objetivo criar um modelo matemático baseado em múltiplas variáveis ambientais capaz de estimar mortalidade em saúde pública. **Métodos:** Para tal, foi feito um levantamento, seleção e organização de múltiplas variáveis pautados no modelo Força Motriz/Pressão/Situação/Exposição/Efeito (FPSEE) recomendado pela Organização Mundial da Saúde. A partir da escolha das variáveis ambientais, foram empregados os seguintes métodos estatísticos de análise multivariada: Análise Fatorial Exploratória (AFE), a fim de encontrar sua estrutura latente e variáveis marcadoras que foram, por fim, usadas para estimar o melhor modelo preditor de mortalidade em saúde pública, por meio da técnica de Análise de Regressão Linear Múltipla (RLM) do tipo *Stepwise*. Todas as análises estatísticas foram processadas pelo software IBM-SPSS *Statistics*, versão 22.0. O banco de dados original foi composto por 853 observações que se referem aos municípios de Minas Gerais, região Sudeste, Brasil e os dados foram obtidos em sistemas públicos de informações virtuais para o ano de 2017. **Resultados:** Com base nos fundamentos teóricos subjacentes, foram selecionadas, a princípio, 130 variáveis para análise, agrupadas em 14 grupos. A partir do modelo FPSEE identificou-se que 19,2% das variáveis foram classificadas como Força Motriz; 6,9% como Pressão; 14,6% como Situação; 21,5% como Exposição e 37,7% como Efeito à Saúde. Após a revisão bibliográfica e a verificação das suposições teóricas e estatísticas da AFE foram excluídas 54 variáveis como decorrência de informações repetidas, natureza da escala, casos omissos e correlações parciais, restando 76 variáveis apropriadas para a análise fatorial. A matriz de correlações de *Spearman* (ρ) apresentou 54,73% de correlações lineares significantes ($\alpha < 0,05$), percentual que aumenta para 59,17% quando consideradas as correlações significantes a nível $\alpha < 0,10$. A fatorabilidade das variáveis foi confirmada pelo teste de esfericidade de Bartlett (p -valor $< 0,001$) e a medida de adequação da amostra de Kaiser-Meyer-Olkin (KMO) com resultado igual a 0,952. A partir da matriz de cargas fatoriais rotacionada (varimax) e com base nos resultados convergentes dos testes *Scree Plot* e o percentual de variância explicada, foram extraídos 5 fatores que, juntos, explicam 59,78% da variância total dos dados. O primeiro fator foi rotulado como socioambiental; o segundo como vulnerabilidade social; o terceiro como qualidade do ar; o quarto como mortalidade e o quinto como agropecuária. As variáveis marcadoras foram respectivamente: número de óbitos por neoplasia; percentual de pessoas inscritas no Cadastro Único sem abastecimento de água adequado; concentração de NO_2 ; taxa de mortalidade por homicídio e por fim, percentual de cobertura por floresta plantada e percentual de cobertura vegetal natural. A variável com maior carga fatorial em cada fator e a variável com a segunda maior carga fatorial do quinto fator foram selecionadas para a estimação do melhor modelo preditor de mortalidade por meio de Regressão Linear Múltipla *Stepwise*. O melhor modelo matemático encontrado pelo método de RLM ($R_{adj}^2 = 0,126$, p -valor $< 0,001$) foi $Y = 7,655 + (-0,289 X_1) + (0,132 X_2) + (-0,109 X_3$), no qual, a variação da variável dependente (taxa bruta de mortalidade) é prevista pelas variáveis ambientais: X_1 = percentual de cobertura vegetal natural ($\beta = -0,289$; p -valor = 0,000), X_2 = taxa de homicídio ($\beta = 0,132$; p -valor = 0,000) e X_3 = percentual de cobertura por floresta plantada ($\beta = -0,109$; p -valor = 0,001). **Conclusão:** Por meio da AFE foram identificados 5 fatores e deles obtidas 6 variáveis marcadoras capazes de representar todo o conjunto inicial de variáveis com a menor perda de informação. A partir das variáveis selecionadas pela AFE, foi possível obter um modelo preditor de mortalidade e determinar quais são as variáveis ambientais que melhor explicam o comportamento de mortalidade em saúde pública. Ao esclarecer as inter-relações entre as variáveis ambientais e a saúde pública pode-se subsidiar tomada de decisão em gestão pública e mitigação dos desfechos críticos na saúde humana.

Palavras-chaves: Saúde Ambiental. Saúde Pública. Análise Fatorial Exploratória. Regressão Linear Múltipla. Análise Estatística Multivariada.

ABSTRACT

Introduction: One of the challenges that public management has been facing is to structure surveillance systems aimed at changing environmental contexts that represent risk situations and critical outcomes for human health. **Objective:** That said, this research aimed to create a mathematical model based on multiple environmental variables capable of estimating mortality in public health. **Methods:** To this end, a survey was carried out, selection and organization of multiple variables was carried out based on the Driving Force/Pressure/Situation/Exposure/Effect (FPSEE) model recommended by the World Health Organization. From the choice of environmental variables, the following statistical methods of multivariate analysis were used: Exploratory Factor Analysis (EFA), in order to find its latent structure and marker variables that were, finally, used to estimate the best mortality predictor model in public health, using the Stepwise Multiple Linear Regression Analysis technique. All statistical analyzes were processed using the IBM-SPSS Statistics software, version 22.0. The original database consisted of 853 observations that refer to the municipalities of Minas Gerais, southeastern region, Brazil and the data were obtained from public virtual information systems for the year 2017. **Results:** Based on the underlying theoretical foundations, 130 variables were initially selected for analysis, grouped into 14 groups. From the FPSEE model, it was identified that 19.23% of the variables were classified as Driving Force; 6.9% as Pressure; 14.6% as Status; 21.5% as Exposure and 37.7% as Health Effect. After reviewing the literature and verifying the theoretical and statistical assumptions of the AFE, 54 variables were excluded as a result of repeated information, nature of the scale, missing cases and partial correlations, leaving 76 variables suitable for factor analysis. The Spearman correlation matrix (ρ) showed 54.73% of significant linear correlations ($\alpha < 0.05$), a percentage that increases to 59.17% when considering significant correlations at the level $\alpha < 0.10$. The factorability of the variables was confirmed by the Bartlett sphericity test (p-value < 0.001) and the Kaiser-Meyer-Olkin (KMO) sample adequacy measure with a result equal to 0.952. From the rotated factor loading matrix (varimax) and based on the convergent results of the Scree Plot tests and the percentage of explained variance, 5 factors were extracted that, together, explain 59.78% of the total variance of the data. The first factor was labeled as socioenvironmental; the second as social vulnerability; the third as air quality; the fourth as mortality and the fifth as agrilivestock. The marker variables were, respectively: number of deaths from cancer; percentage of people enrolled in the Single Registry without adequate water supply; NO₂ concentration; homicide mortality rate and finally, planted forest cover and natural vegetation cover. The variable with the highest factor loading in each factor and the variable with the second highest factor loading in the fifth factor were selected for the estimation of the best predictor model of mortality through Stepwise Multiple Linear Regression. The best model mathematical found by the RLM method ($R^2_{adj} = 0.126$, p-value < 0.001) was $Y = 7.655 + (-0.289 X_1) + (0.132 X_2) + (-0.109 X_3)$, in which the variation of the variable dependent (gross mortality rate) is predicted by environmental variables: X_1 = percentage of natural vegetation cover ($\beta = -0.289$; p-value = 0.000), X_2 = homicide rate ($\beta = 0.132$; p-value = 0.000) and X_3 = percentage of coverage by planted forest ($\beta = -0.109$; p-value = 0.001). **Conclusion:** Through the EFA, 5 factors were identified and from them 6 marker variables were obtained capable of representing the entire initial set of variables with the least loss of information. From the variables selected by the AFE, it was possible to obtain a predictor model of mortality and determine which environmental variables best explain the behavior of mortality in public health. By clarifying the interrelationships between environmental variables and public health, it is possible to support decision-making in public management and mitigation of critical outcomes in human health.

Keywords: Environmental health. Public health. Exploratory Factor Analysis. Multiple Linear Regression. Multivariate Statistical Analysis.

LISTA DE ILUSTRAÇÕES

FUNDAMENTAÇÃO TEÓRICA

Figura 01-	Estrutura organizacional da CGVAM	24
Figura 02-	Fatores de risco não biológicos e suas doenças e agravos	25
Figura 03-	Fatores ambientais de risco biológico	26
Figura 04-	Programas de Vigilância em Saúde Ambiental relacionados aos fatores de risco não biológicos	27
Figura 05-	Pirâmide de Informações (Dados, Indicadores e Índices)	32
Figura 06-	Modelo de Organização dos Indicadores: Pressão - Estado - Resposta	39
Figura 07-	Modelo de Organização dos Indicadores: Pressão - Estado - Impacto – Resposta	40
Figura 08-	Modelo de organização de indicadores FPSEEA	41
Quadro 01-	Propriedades desejáveis dos indicadores	34
Quadro 02-	Lista dos principais Sistemas de Informações em Saúde	37
Quadro 03-	Lista dos principais Sistemas de Informações Ambientais	37

ARTIGO 01

Quadro 01-	Descrição das variáveis de saúde ambiental, fator e classificação no modelo FPSEE	61
------------	---	----

ARTIGO 02

Figura 01-	Fluxograma da seleção das variáveis para estudo	80
Figura 02-	Gráfico de <i>Scree Plot</i>	83

ARTIGO 03

Figura 01-	Análises estatísticas multivariadas usadas	102
Figura 02-	Gráfico de Dispersão de Resíduos	104

LISTA DE TABELAS

ARTIGO 01

Tabela 1 -	Número de fatores, n amostral, tipo e natureza das variáveis selecionadas	65
Tabela 2-	Tabela 02: Estatística Descritiva (variável número de óbitos por neoplasia)	66
Tabela 3 -	Tabela 03: Estatística Descritiva (variável Cadastro Único sem abastecimento de água adequado)	67
Tabela 4 -	Tabela 04: Estatística Descritiva (variável concentração de Dióxido de Nitrogênio - NO ₂)	68
Tabela 5 -	Tabela 05: Estatística Descritiva (variável Taxa de mortalidade por homicídio)	68
Tabela 6 -	Tabela 06: Estatística Descritiva (variável cobertura vegetal natural)	69
Tabela 7 -	Resumo de processamento do caso, fator (qualidade do solo, contaminantes ambientais e substâncias químicas)	70

ARTIGO 02

Tabela 1 -	Identificação das variáveis excluídas e incluídas na AF	81
Tabela 2-	Teste de <i>KMO</i> e <i>Bartlett</i>	82
Tabela 3 -	Variância Total Explicada	82
Tabela 4 -	Fatores extraídos e suas respectivas variáveis	84
Tabela 5 -	Fatores (F1-F5), variáveis marcadoras e suas respectivas cargas fatoriais	85

ARTIGO 03

Tabela 1 -	Correlações de Pearson entre as variáveis independentes e Taxa bruta de mortalidade	104
Tabela 2-	Avaliação da Colinearidade no melhor modelo obtido pelo método <i>Stepwise</i> de RLM	105
Tabela 3 -	Variáveis incluídas e excluídas em cada modelo preditor da Taxa bruta de mortalidade	106
Tabela 4-	Capacidade preditiva do modelo	107

LISTA DE ABREVIATURAS E SIGLAS

AFC	Análise Fatorial Confirmatória
AFE	Análise Fatorial Exploratória
ASISAST	Área de Análise de Situação de Saúde Ambiental e Saúde do Trabalhador
BVS	Biblioteca Virtual em Saúde
CGVAM	Coordenação Geral de Vigilância em Saúde Ambiental
CMMAD	Comissão Mundial sobre o Meio Ambiente e Desenvolvimento
CNUMAD	Conferência das Nações Unidas sobre Meio Ambiente e Desenvolvimento
COFAB	Coordenação de Vigilância e Controle dos Fatores de Risco Biológicos
CONAB	Coordenação de Vigilância e Controle dos Fatores de Risco Não Biológicos
COPASAD	Conferência Pan Americana de Saúde e Ambiente e Desenvolvimento
CTI	Comitê Temático Interdisciplinar
CVE	Centro de Vigilância Epidemiológico
DATASUS	Departamento de Informática do SUS
DCNT	Doenças Crônicas não Transmissíveis
DCV	Doenças Cerebrovasculares
DDA	Doenças Diarréicas Agudas
DECS	Descritores em Ciências da Saúde
DIC	Doenças Isquêmica do Coração
DOMA	Doenças Ocasionadas pelo Meio Ambiente
DPOC	Doença Pulmonar Obstrutiva Crônica
FEAM	Fundação Estadual do Meio Ambiente de Minas Gerais
FGV	Fundação Getúlio Vargas
FIOCRUZ	Fundação Oswaldo Cruz
FJP	Fundação João Pinheiro
FPSEEA	Força Motriz, Pressão, Situação, Exposição, Efeito, Ação
FUNASA	Fundação Nacional de Saúde
IBAMA	Instituto Brasileiro de Meio Ambiente e Recursos Renováveis
IBGE	Instituto Brasileiro de Geografia e Estatística
IDB	Indicadores e Dados Básicos
IDS	Indicadores de Desenvolvimento Sustentável
IGAM	Instituto da Gestão das Águas de Minas Gerais
INEP	Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira
INPE	Instituto Nacional de Pesquisas Espaciais

IPEA	Instituto de Pesquisa Econômica Aplicada
IRA	Infecções Respiratórias Agudas
ISA	Indicadores de Saúde Ambiental
MEE	Modelagem de Equações Estruturais
MMA	Ministério do Meio Ambiente
MS	Ministério da Saúde
OCDE	Organização para a Cooperação e Desenvolvimento Econômico
ONU	Organização das Nações Unidas
OMS	Organização Mundial de Saúde
OPAS	Organização Pan-Americana de Saúde
PER	Pressão, Estado, Resposta
PEIR	Pressão, Estado, Impacto, Resposta
PIB	Produto Interno Bruto
PISAST	Painel de Informações em Saúde Ambiental e Saúde do Trabalhador
PNAD	Pesquisa Nacional por Amostra de Domicílios
PNMCSAÚDE	Plano Nacional de Mudanças Climáticas e Saúde
PNSB	Pesquisa Nacional de Saneamento Básico
PNUMA	Programa das Nações Unidas para o Meio Ambiente
RIPSA	Rede Interagencial de Informações para a Saúde
SAI	Sistema de Informações Ambulatoriais
SIA/SUS	Sistema de Informações Ambulatoriais do Sistema Único de Saúde
SIAB	Sistema de Informações da Atenção Básica
SIAM	Sistema de Informações Ambientais no Mercosul
SIBEA	Sistema Brasileiro de Informações sobre Educação Ambiental
SIBSA	Simpósio Brasileiro de Saúde Ambiental
SIGEPRO	Sistema de Georreferenciamento de Projetos
SIH/SUS	Sistema de Informações Hospitalares do Sistema Único de Saúde
SIM	Sistema de Informação Sobre Mortalidade
SINAN	Sistema de Informações sobre Agravos de Notificação
SINASC	Sistema de Informação Sobre Nascidos Vivos
SINIS	Sistema Nacional de Informações sobre Saneamento
SINITOX	Sistema Nacional de Informação Tóxico Farmacológica
SINVAN	Sistema de Informação de Vigilância Alimentar e Nutricional
SINVSA	Subsistema Nacional de Vigilância em Saúde Ambiental
SIS	Sistema de Informações em Saúde

SISÁGUA	Sistema de Informação de Vigilância da Qualidade da Água
SISAM	Sistema de Informações Ambientais Integrados a Saúde
SVS	Secretária de Vigilância em Saúde
SUS	Sistema Único de Saúde
UNESCO	Organização das Nações Unidas para Educação, Ciência e Cultura
VIGIAGUA	Vigilância em Saúde Ambiental Relacionada à Qualidade da Água
VIGIAPP	Vigilância Ambiental em Saúde Relacionada aos Acidentes Envolvendo Produtos Perigosos
VIGIAR	Vigilância em Saúde Ambiental Relacionada à Qualidade do Ar
VIGIDESASTRES	Vigilância em Saúde Ambiental Relacionada aos Riscos Decorrentes dos Desastres Naturais
VIGIFIS	Vigilância em Saúde Ambiental Relacionada a Fatores Físicos
VIGIPEQ	Vigilância em Saúde de Populações Expostas a Contaminantes Químicos
VIGIQUIM	Vigilância em Saúde Ambiental Relacionada à Segurança Química
VIGISOLO	Vigilância em Saúde Ambiental Relacionada às Populações Expostas a Solos Contaminados

SUMÁRIO

1. INTRODUÇÃO	17
1.1. Considerações Iniciais	21
2. FUNDAMENTAÇÃO TEÓRICA	22
2.1. O CONTEXTO: SAÚDE/MEIO AMBIENTE, SAÚDE AMBIENTAL E VIGILÂNCIA EM SAÚDE AMBIENTAL.....	22
2.2. ESTRUTURA ORGANIZACIONAL DO SUBSISTEMA NACIONAL DE VIGILÂNCIA EM SAÚDE AMBIENTAL (SINVSA)	24
2.3. COORDENAÇÃO GERAL DE VIGILÂNCIA EM SAÚDE AMBIENTAL (CGVAM)	26
2.3.1. VIGIÁGUA	26
2.3.2. VIGIPEQ	27
2.3.2.1. VIGIAR	27
2.3.2.2. VIGISOLO	27
2.3.2.3. VIGIQUIM.....	28
2.3.3. VIGIDESASTRES	28
2.3.3.1. VIGIFIS	29
2.3.3.2. VIGIAPP	29
2.4. DEFINIÇÕES BÁSICAS: DADO, INDICADOR E ÍNDICE	30
2.5. INDICADORES AMBIENTAIS <i>VERSUS</i> INDICADORES DE SAÚDE	31
2.6. INDICADORES DE SAÚDE AMBIENTAL	32
2.6.1. Critérios para o desenvolvimento de indicadores de saúde ambiental	32
2.7. FONTES DE DADOS: NACIONAIS, ESTADUAIS E MUNICIPAIS E SISTEMAS DE INFORMAÇÃO: PRINCIPAIS CARACTERÍSTICAS	35
2.7.1. Sistemas de Informações em Saúde	36
2.7.2. Sistemas de Informações Ambientais	36
2.7.3. Sistema Nacional de Informação em Saúde Ambiental	37
2.8. MODELOS DE ORGANIZAÇÃO DOS INDICADORES	37
2.8.1. Modelo Pressão - Estado - Resposta - PER	38
2.8.2. Modelo Pressão - Estado - Impacto - Resposta - PEIR	39
2.8.3. Modelo Força Motriz - Pressão - Situação - Exposição - Efeito - Ação - FPSEEA	40
2.9. CONCEITOS BÁSICOS ESTATÍSTICOS: ANÁLISE MULTIVARIADA DE DADOS	42

2.9.1. A variável estatística	43
2.9.2. Tipos de Escalas de Medida	43
2.9.2.1. Escalas de medidas não-métricas	43
2.9.2.2. Escalas de medidas métricas	44
2.9.3. Validade e Confiabilidade	45
2.9.4. Testes de Significância	45
2.10. CLASSIFICAÇÃO E TÉCNICAS MULTIVARIADAS UTILIZADAS.....	47
2.10.1. Análise Fatorial	48
2.10.2. Regressão Linear Múltipla	50
3. OBJETIVOS	51
3.1. Objetivo Geral.....	51
3.2. Objetivos Específicos	51
4. RESULTADOS	52
4.1. Artigo 1	52
4.2. Artigo 2.....	72
4.3. Artigo 3.....	95
5. CONSIDERAÇÕES FINAIS	113
REFERÊNCIAS	114
ANEXO 1	118

1. INTRODUÇÃO

A urbanização brasileira intensificada na segunda metade do século XX e o crescimento populacional e econômico observado nas últimas décadas ocasionou um cenário de condições ambientais insatisfatórias. Um dos desafios que a gestão pública vem enfrentando é o de estruturar sistemas de vigilância e monitoramento voltados para a mudança de contextos socioambientais que representem situações de risco à saúde humana. É diante dessa problemática que o monitoramento dos determinantes da saúde pública em municípios, estados e países vêm sendo debatidos no meio acadêmico e governamental.

A abordagem interdisciplinar de Saúde Única (*One Health*) definida pela Organização Mundial da Saúde (OMS) demonstra a interligação da saúde humana, animal e ambiental representada por um complexo sistema biológico e social. É necessário um olhar integrado entre os diferentes componentes da tríade humano/animal/ambiente para prevenir doenças e proteger a saúde da população (OMS, *et. al.*, 2022).

Algumas doenças e agravos à saúde possuem consistentes vínculos com o meio ambiente, tais como: doenças microbiológicas de origem alimentar, intoxicações humanas por agrotóxicos, doenças de veiculação hídrica (febre tifóide, cólera, hepatite, amebíase, giardíase, esquistossomose, ascaridíase, leptospirose), doenças transmitidas por vetores e animais (malária, dengue, febre amarela, zika vírus, leishmaniose, raiva, toxoplasmose, doença de Chagas), Doenças Diarréicas Agudas (DDA) relacionadas ao saneamento ambiental inadequado, doenças cardiovasculares e Infecções Respiratórias Agudas (IRA) causadas pela poluição do ar (SALDIVA, 2020); (ARRUDA e HELLER, 2022); (GODOY et al., 2021); (LOPES e ALBUQUERQUE, 2018).

De acordo com a Organização Mundial da Saúde (OMS), ocorrem anualmente 7 milhões de mortes prematuras atribuídas à poluição do ar no mundo (OMS, 2021). Em 2016, a poluição foi responsável por cerca de 58% de mortes prematuras por Doenças Cerebrovasculares (DCV) e Doenças Isquêmica do Coração (DIC); 18% por Doença Pulmonar Obstrutiva Crônica (DPOC) e infecção respiratória aguda; e 6% por câncer de pulmão, traquéia e brônquios (OMS, 2019).

No Brasil, dados do estudo ‘Saúde Brasil’ (2018) realizado pelo Ministério da Saúde, constataram que de 2006 para 2016 houve um aumento de 14% na carga de mortalidade em decorrência da poluição atmosférica. Nesse período, o número de óbitos por Doenças Crônicas não Transmissíveis (DCNT) atribuídas à poluição do ar passou de 38.782 em 2006

para 44.228 mortes em 2016 (BRASIL, 2019). As Doenças Isquêmicas do Coração (DIC) foram responsáveis pelo maior número de mortes, sendo seguidas pelas doenças cerebrovasculares (BRASIL, 2019). Também foi constatado que mulheres expostas a contaminantes do ar durante a gestação são susceptíveis a desenvolverem efeitos adversos no crescimento e desenvolvimento do feto (OMS, 2021).

Além das vidas perdidas, a poluição atmosférica também é responsável por custos diretos na saúde pública. As estimativas são elevadas, apenas no ano de 2018, as despesas com internações relacionadas à problemas respiratórios ultrapassaram R\$ 1,3 bilhão (BRASIL, 2019).

Portanto, é necessário um olhar integrado sobre as questões de saúde que seja capaz de compreender as complexas inter-relações entre a carga de doenças, pressões ambientais e exposição à múltiplos fatores, assim como o papel desempenhado pelas vulnerabilidades sociais – uma vez que muitos riscos ambientais atingem de forma mais grave populações de baixa renda, aquelas que moram em áreas de risco como encostas, áreas de enchentes ou de maior poluição, que dependem ou moram perto de lixões. Ainda fazem parte dessa complexa rede de relações, questões emergentes como as mudanças climáticas, zoonoses, alterações do padrão de consumo e o uso de produtos químicos (SOLOMON e ROCQUE, 2019).

Diante dos efeitos das condições ambientais sobre saúde humana é que os estudos sobre a saúde ambiental vêm se desenvolvendo na tentativa de monitorar os efeitos adversos à saúde. A saúde ambiental envolve os aspectos da saúde humana, que são influenciados por fatores ambientais como físicos, químicos, biológicos e sociais (ALENCAR, *et.al.*, 2020).

Diante dessa conjuntura, fazem-se necessárias políticas públicas voltadas para o objetivo de proteger as condições de sobrevivência das populações e garantir ambientes saudáveis (LOPES, BERNARDINO NETO, 2020). Nesse sentido, os Indicadores de Saúde Ambiental (ISA) são considerados medidas-síntese que possibilitam a avaliação da saúde por meio do ambiente, a partir deles têm-se uma compreensão dessa inter-relação, sua possível evolução, e conseqüentemente permite seu monitoramento. Os ISA começaram a ser definidos em meados da década de 1990, pela Organização Mundial da Saúde – OMS e pela Organização Pan-Americana da Saúde – OPAS (BRASIL, 2011). Os indicadores já existentes eram aplicáveis ao meio ambiente, mas não aos aspectos sanitários, pois careciam de dois elementos-chave: a exposição a fatores ambientais e seu impacto sobre a saúde (BRASIL, 2011).

Nesse cenário, essa pesquisa teve a teoria científica subjacente como principal embasamento para levantamento das múltiplas variáveis ambientais e, como orientação, o Guia Básico para a Construção de Indicadores em Saúde Ambiental - realizado no âmbito da Rede Interagencial de Informações para a Saúde (RIPSA) (BRASIL, 2011).

Para a definição dos fatores foram considerados os eixos de saúde ambiental propostos pela OMS e apresentados no documento: “*Environmental Health Indicators: Development of a Methodology for the WHO European Region*” (OMS, 2001), tais como: qualidade do ar, habitação, resíduos, solos contaminados, água para consumo, segurança alimentar, substâncias químicas, violência e ambiente de trabalho. Ademais, foram considerados os eixos dos Programas de Vigilância em Saúde Ambiental, de competência da Coordenação Geral de Vigilância em Saúde Ambiental (CGVAM) criada pelo Ministério da Saúde, tais como: água, ar, solo, desastres naturais e contaminantes (BRASIL, 2013).

Destarte, a organização da matriz de indicadores foi pautada no último modelo de análise desenvolvido pela OMS (Força Motriz/Pressão/Situação/Exposição/Efeito/Ação /FPSEEA) (ARANTES e PEREIRA, 2017). A sua vantagem em relação aos modelos anteriores, como o modelo Pressão/Estado/Resposta (PER) e o Pressão/Estado/Impacto/Resposta (PEIR), é que ele permite uma melhor flexibilidade na análise das inter-relações dos diferentes níveis da matriz, e ao mesmo tempo, amplia o domínio da saúde ambiental (exposição/efeito) (BRASIL, 2011).

Outro importante fundamento desta pesquisa refere-se às análises estatísticas multivariadas de dados. Embora os indicadores sejam uma representação simplificada da realidade, suas análises múltiplas nem sempre são operações simples. Um dos problemas que se tem é de ordem conceitual e está relacionado ao arco de abrangência dos ISA, uma vez que acomodar aleatoriamente a maioria dos indicadores traz em si o risco de construção de cenários não realísticos e de difícil compreensão (OMS, 2022). No entanto, quando examinados corretamente por meio de análises estatísticas adequadas, são considerados importantes instrumentos que podem ser valiosos para gestores públicos, pesquisadores e profissionais de saúde nas tomadas de decisões.

Nesse sentido, esse trabalho tem a seguinte pergunta norteadora: Como analisar simultaneamente múltiplos indicadores/variáveis de saúde ambiental de modo que sejam capazes de inferir efeitos críticos à saúde pública? Para tal, foram empregados levantamento de dados oficiais, estudo de bases conceituais/ teóricas e métodos estatísticos de análise multivariada de dados: Análise Fatorial Exploratória (AFE) e a Regressão Linear Múltipla

(RLM). A AFE teve como principal objetivo avaliar o comportamento das variáveis, determinar seus constructos latentes por meio das relações entre elas e, assim, permitir a redução do conjunto original de variáveis. A Regressão Linear Múltipla teve a capacidade de encontrar a melhor combinação linear entre múltiplas variáveis ambientais capaz de maximizar a previsibilidade da variável desfecho.

Portanto, esse trabalho possui relevante importância uma vez que ao identificar o constructo de variáveis ambientais relacionadas à riscos à saúde pública, bem como aquelas que são melhores preditoras do desfecho óbito, poderá subsidiar a tomada de decisão de forma responsável, coletiva, integrada e territorializada por governantes e agentes de saúde. O que emerge dessa discussão sobre tomada de decisão é a demanda por evidências que possam orientar políticas e estratégias eficazes em saúde.

Para executar ambas as análises estatísticas, foram usadas informações oficiais de 853 municípios de Minas Gérias, região Sudeste, Brasil. A escolha desses municípios se deve à sua marca da heterogeneidade, como destaca Queiroz (2001), o estado de Minas Gerais é um retrato-síntese do Brasil e uma das regiões mais heterogêneas do país, com algumas regiões desenvolvidas e outras que oferecem condições de vida insatisfatórias à sua população. Ademais, essa escolha garante também o número de observações suficiente para a significância estatística dos testes realizados. Assim sendo, essa pesquisa possui não apenas validade interna verificada estatisticamente, como também validade externa (extensão) uma vez que, os resultados encontrados podem ser generalizados para populações além dessa amostra e em outros países.

A originalidade deste trabalho fica evidente na análise simultânea de múltiplas variáveis de saúde ambiental, na identificação da estrutura latente de relações entre elas e na determinação de um modelo matemático capaz de identificar desfechos críticos sobre a saúde pública a partir do uso de análises estatísticas multivariadas de dados, especificamente a técnica de Modelagem de Equações Estruturais (combinação da AFE e RLM), estudo nunca antes feito até o momento atual!

A iniciativa e os resultados encontrados nessa tese apresentam relevância, originalidade e validade, sendo instrumento valioso para futuras pesquisas além de corroborar como material de apoio à administração pública das esferas municipais à federal, à iniciativa privada, à comunidade acadêmica, profissionais de saúde responsáveis por tomadas de decisões em saúde pública e à sociedade civil organizada que tenha interesse em mitigar os riscos e danos em saúde humana.

1.1. Considerações Iniciais

Esta tese está organizada de acordo com o modelo do Programa de Pós-Graduação em Ciências da Saúde (PPCSA-UFU), que permite a apresentação dos resultados da pesquisa na forma de artigos científicos.

A estrutura da tese segue as seções: **introdução** com explanação geral do tema, fundamentos do estudo, pergunta norteadora, justificativa, relevância, extensão e originalidade; **fundamentação teórica** com abordagens sobre o contexto saúde, meio ambiente, vigilância epidemiológica e Programas de Vigilância em Saúde Ambiental. Ademais, apresentam-se as definições básicas de dado, indicador e índice com ênfase para indicadores ambientais e de saúde e os critérios para o desenvolvimento de indicadores de saúde ambiental.

Após as contextualizações e definições importantes sobre a temática partiu-se para a análise dos sistemas de informações existentes para a obtenção dos dados, com destaque para o Departamento de Informática do Sistema Único de Saúde (DATASUS), Instituto Brasileiro de Geografia e Estatística (IBGE), Instituto Nacional de Pesquisas Espaciais (INPE), Sistema de Informações Ambientais integrado à Saúde Ambiental (SISAM). Posteriormente, tem-se o levantamento dos modelos existentes de organização de múltiplas variáveis com destaque para o modelo Força Motriz- Pressão- Situação- Exposição- Efeito- Ação (FPSEEA).

Finalizando a abordagem teórica, são apresentados os conceitos básicos estatísticos de variável estatística, escalas de medida, validade, confiabilidade, testes de significância e os métodos estatísticos de análise multivariada (Análise Fatorial Exploratória e Regressão Linear Múltipla). A terceira seção refere-se aos **objetivos** da pesquisa (objetivo geral e específicos) e na quarta seção são abordados os **resultados** adquiridos a partir do desenvolvimento dessa investigação que procederam em três artigos científicos, sendo:

1º artigo: Levantamento, seleção e organização de indicadores de saúde ambiental: Uma análise estatística exploratória e descritiva;

2º artigo: Meio Ambiente e Saúde Humana: O uso da Análise Fatorial Exploratória – AFE como estatística de análise de múltiplas variáveis;

3º artigo: O uso da Regressão Linear Múltipla - RLM como modelo de explicação das variáveis ambientais que inferem desfechos críticos na saúde pública (óbitos).

Por fim, são apresentadas as **conclusões**, com considerações sobre o objetivo da pesquisa, limitações encontradas e sugestões para trabalhos futuros.

2. FUNDAMENTAÇÃO TEÓRICA

2.1. O CONTEXTO: SAÚDE/MEIO AMBIENTE, SAÚDE AMBIENTAL E VIGILÂNCIA EM SAÚDE AMBIENTAL

Foi na Conferência das Nações Unidas sobre Meio Ambiente e Desenvolvimento (CNUMAD ou Rio-92), realizada em 1992 no Rio de Janeiro, que a relação entre o meio ambiente e os impactos na saúde ganhou forte apoio nacional e internacional. A Rio-92 foi um marco pela aprovação da agenda 21 - documento que apresenta orientações articuladas para a saúde, meio ambiente e desenvolvimento sustentável. Em 1995, a Organização Pan Americana da Saúde (OPAS) realizou a Conferência Pan Americana de Saúde e Ambiente e Desenvolvimento (COPASAD) estabelecendo as bases estratégicas, pelos países membros, de criar ações que envolvessem o equilíbrio do meio ambiente e a saúde (OPAS, 2018).

No final do século XIX, a medicina e a saúde pública foram revolucionadas com a descoberta dos microorganismos causadores de doenças infecciosas, marcando o advento do modelo biomédico. Esse modelo definiu a doença como um desequilíbrio no funcionamento do organismo humano, portanto, a ênfase estava no meio interno (corpo humano) e na força dos germes ou microorganismos causarem doença (REZENDE, MORAES e PERINI, 2018).

Com o declínio das doenças infecciosas nos países desenvolvidos e com o avanço das doenças crônico-degenerativas surgiu um novo modelo chamado de processual ou história natural da doença. Esse modelo prevê que o meio ambiente (externalidade) contém fatores de risco para a saúde humana. Portanto, o modelo processual se constitui um avanço quando comparado com o biomédico, uma vez que reconhece que fatores ambientais relacionados ao indivíduo interferem no desenvolvimento e repercussão da doença (REZENDE, MORAES e PERINI, 2018).

No entanto, uma abordagem mais abrangente do processo saúde-doença é explicada no modelo sistêmico. Ele reconhece que os fatores de riscos podem ser físicos, químicos, biológicos, sociais, econômicos e ambientais e estão relacionados sinergicamente - de forma que mudanças em um determinado elemento provocam mudanças nos demais. Assim, a interação desses fatores é o que determina a ocorrência do dano (COSTA *et.al.*, 2020).

Nesse sentido, foi só na metade do século XX que se estruturou uma área específica para tratar a questão saúde/meio ambiente, denominada de saúde ambiental (ALENCAR, *et.al.*, 2020). A saúde ambiental é uma área da saúde pública que se refere à teoria e prática de

controlar e evitar os fatores do meio ambiente que, potencialmente, possam prejudicar a saúde das presentes e futuras gerações (OMS, 2022). O Ministério da Saúde também a caracteriza sob duas dimensões principais: promoção da saúde e avaliação de risco voltada para as adversidades ambientais que potencialmente geram agravos e enfermidades (BRASIL, 2007).

Diante desse contexto, a principal iniciativa no âmbito do Ministério da Saúde, foi estruturar uma área de Vigilância em Saúde Ambiental na Fundação Nacional de Saúde (FUNASA) e que, no fim da década de 1990, passou a ser competência da Secretaria de Vigilância em Saúde (SVS), especificamente da Coordenação Geral de Vigilância em Saúde Ambiental (CGVAM) (BRASIL, 2018).

A Vigilância em Saúde Ambiental consiste em um conjunto de ações e serviços que propiciam o conhecimento e a detecção de mudanças nos fatores determinantes e condicionantes do meio ambiente que interferem na saúde humana, com a finalidade de recomendar e adotar medidas de promoção à saúde, prevenção e monitoramento dos fatores de riscos relacionados às doenças ou agravos à saúde (BRASIL, p.4, 2018).

No Brasil, a estruturação da Vigilância em Saúde Ambiental se consolidou com o surgimento da SVS a partir do decreto federal nº4726/2003 e pelo Subsistema Nacional de Vigilância em Saúde Ambiental (SINVSA) instituído pela instrução normativa nº 01 de 07 de março de 2005 (BRASIL, 2003); (BRASIL, 2005).

A SVS é responsável pelas ações de vigilância, prevenção e controle de doenças transmissíveis, análise epidemiológica e vigilância de doenças não transmissíveis, vigilância de doenças de condições crônicas e infecções sexualmente transmissíveis, vigilância em saúde ambiental e saúde do trabalhador, assim como articulação estratégica de vigilância em saúde e emergências em saúde pública (BRASIL, 2022). Por sua vez, cabe ao SINVSA um conjunto de ações para detecção de qualquer alteração nos fatores determinantes e condicionantes do meio ambiente que possam propiciar efeitos na saúde humana (BRASIL, 2005).

Portanto, a atuação da Vigilância em Saúde Ambiental requer um trabalho multiprofissional e interinstitucional, sendo necessário considerar a utilização e desenvolvimento de métodos para análise, detecção e controle dos riscos ambientais que geram doenças e agravos.

No que tange a vigilância epidemiológica, nota-se que tem ampliado seu escopo desde as doenças transmissíveis e imunopreveníveis até as doenças crônicas não transmissíveis, violência, acidentes de trabalho, intoxicações e mais recentemente, para a população expostas

aos fatores de risco ambientais que impactam a saúde humana (BRASIL, 2018). Portanto, percebe-se uma transição epidemiológica em que a preocupação com os indivíduos não é apenas com a doença em si (ênfase biológica), mas sim com um conjunto de fatores que engloba aspectos sociais, econômicos e ambientais (COSTA *et.al.*, 2020).

Com a introdução da computação eletrônica e o raciocínio estatístico, as investigações epidemiológicas passaram a ser cada vez mais diversificadas indo para além das doenças infecciosas e associando vários aspectos que englobam o processo saúde/doença, sendo possível trabalhar com amplos bancos de dados (OMS, 2022).

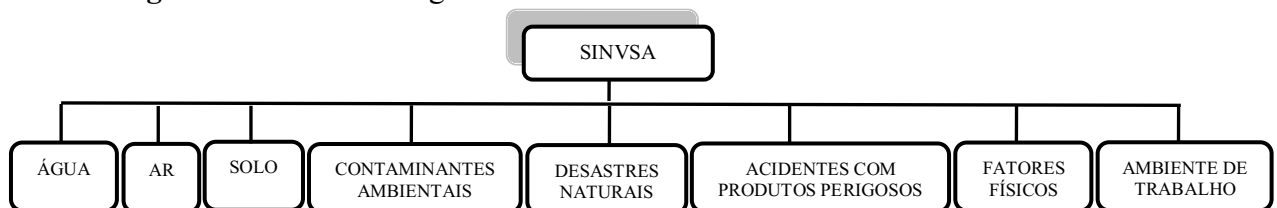
Em suma, a Vigilância Epidemiológica em Saúde Ambiental visa deter o conhecimento da situação de saúde a partir dos aspectos ambientais analisando a morbidade e mortalidade referente às três principais características: população, espaço e tempo. Essa análise pode ser feita a partir da obtenção de dados e informações disponibilizados nos bancos de dados virtuais tais como: DATASUS, IBGE, SISAM, INPE (BRASIL, 2019).

Portanto, dentre as ferramentas fundamentais para análise da Vigilância em Saúde Ambiental pode-se destacar o georeferenciamento de dados que permite a elaboração de mapas de risco de um determinado local, estudos epidemiológicos que permitem analisar os fatores de risco ambientais, o uso de indicadores/variáveis, sistemas de informação e métodos estatísticos de análise (BRASIL, 2019).

2.2. ESTRUTURA ORGANIZACIONAL DO SUBSISTEMA NACIONAL DE VIGILÂNCIA EM SAÚDE AMBIENTAL (SINVSA)

O Subsistema Nacional de Vigilância em Saúde Ambiental (SINVSA), implementado pelo Ministério da Saúde tem a finalidade de adotar e recomendar medidas que visem a promoção da saúde, prevenção e controle de fatores do ambiente (Figura 01) que estão relacionados a agravos e doenças (BRASIL, 2005).

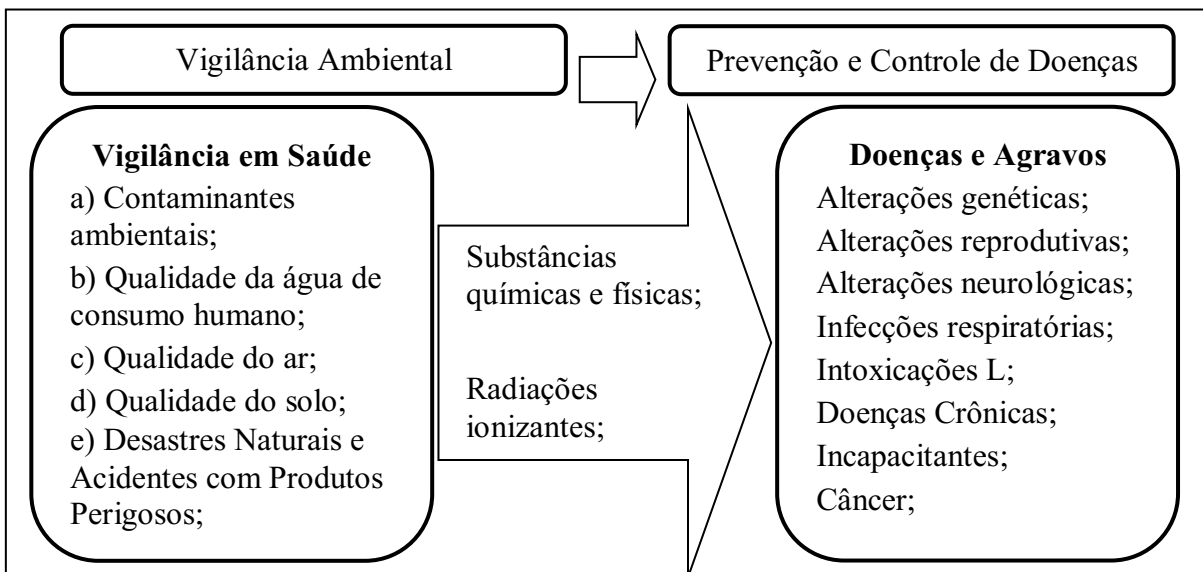
Figura 01- Estrutura Organizacional do SINVSA



Fonte: BRASIL, 2005. Organização: LOPES, A. F. A.; 2020.

As áreas de concentração focadas pelo SINVSA referem-se às atividades relativas aos fatores não biológicos do meio ambiente sendo: água para consumo humano, qualidade do ar, do solo, desastres naturais, contaminantes ambientais e substâncias químicas, desastres naturais, acidentes com produtos perigosos, fatores físicos associados aos riscos de doenças - alterações genéticas, reprodutivas, neurológicas, respiratórias, intoxicações, doenças crônicas e ambiente de trabalho (BRASIL, 2005) (Figura 02).

Figura 02- Fatores de risco ambientais e suas doenças e agravos



Fonte: BRASIL, 2013. Organização: LOPES, A. F. A., 2020.

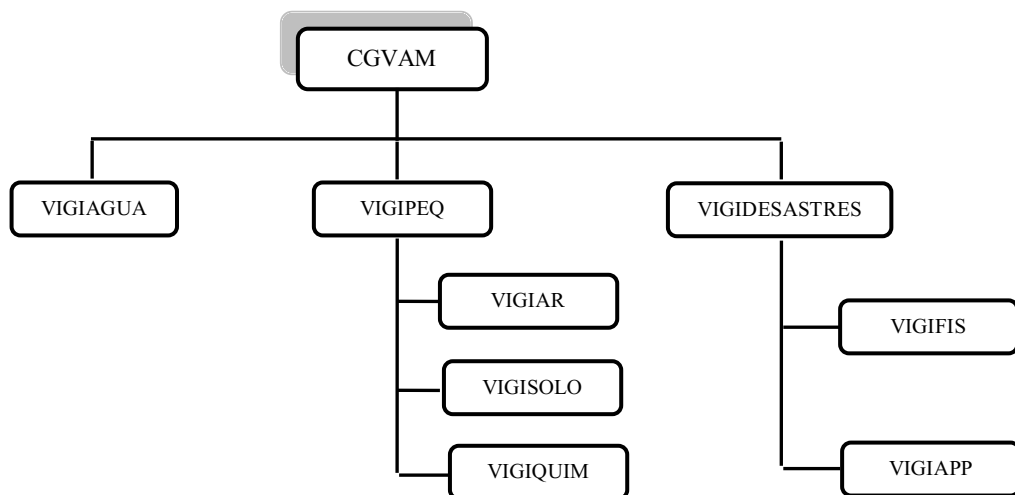
A vigilância dos contaminantes ambientais requer o mapeamento das áreas de risco, (água para consumo humano, atmosfera ou solo) a fim de identificar as fontes de contaminação e minimizar os riscos (doenças e agravos) que podem ser provocados por agrotóxicos, chumbo, mercúrio, benzeno, amianto e outros (BRASIL, 2013).

A vigilância da qualidade da água refere-se à sua potabilidade para consumo humano a fim de evitar que a população adoça pela presença de patógenos ou contaminantes. Por sua vez, o monitoramento da qualidade do ar visa detectar as áreas de risco de poluição identificando as substâncias químicas e os agentes físicos que provocam efeitos adversos à saúde. Em relação à vigilância da qualidade do solo, o objetivo principal é identificar as áreas (superfície ou subsolo terrestre) que se encontram contaminados, principalmente por resíduos. E o monitoramento de desastres naturais, antropogênicos e acidentes com produtos perigosos referem-se às situações de emergências relacionadas às enchentes, secas, incêndios e acidentes químicos que envolvam riscos para a saúde humana ou para o ambiente (BRASIL, 2013).

2.3. COORDENAÇÃO DE GERAL DE VIGILÂNCIA EM SAÚDE AMBIENTAL (CGVAM)

As principais áreas de concentração da Coordenação Geral de Vigilância em Saúde Ambiental são: Vigilância da Qualidade da Água para Consumo Humano, Vigilância em Saúde de Populações Expostas a Substâncias Químicas; Vigilância em Saúde de Populações Expostas a Poluentes Atmosféricos e Vigilância em Saúde Ambiental relacionada aos riscos decorrentes de Desastres (BRASIL, 2022). Portanto, foram estruturados os programas de vigilância em saúde ambiental: VIGIÁGUA; VIGIPEQ - VIGIAR, VIGISOLO, VIGIQUIM; e VIGIDESASTRES – VIGIFIS e VIGIAPP (BRASIL, 2013) (Figura 04).

Figura 04- Programas de Vigilância em Saúde Ambiental



Fonte: BRASIL, 2013. **Organização:** LOPES, A. F. A., 2020.

2.3.1. VIGIAGUA

A Vigilância em Saúde Ambiental Relacionada à Qualidade da Água para Consumo Humano (VIGIÁGUA) consiste no conjunto de ações adotadas continuamente pelas autoridades de saúde pública para garantir à população o acesso à água potável em quantidade suficiente e de acordo com os critérios de potabilidade determinado pela legislação vigente (BRASIL, 2004). A ingestão de água contaminada com patógenos como vírus, bactérias, protozoários, helmintos ocasionam-se a muitas enfermidades, tais como: cólera, diarreia, amebíase, giardíase, hepatite A, por isso a sua qualidade para consumo humano é uma prioridade na área da saúde pública.

A VIGIAGUA visa melhorar as condições sanitárias de abastecimento de água e monitorar sistematicamente a sua qualidade para reduzir a morbimortalidade por doenças e agravos de transmissão hídrica, além disso, apoia o desenvolvimento de ações de educação em saúde e mobilização social (BRASIL, 2013). Outro objetivo de sua atuação é coordenar o Sistema de Informação de Vigilância da Qualidade da Água (SISAGUA).

2.3.2. VIGIPEQ

A Vigilância em Saúde de Populações Expostas a Contaminantes Químicos (VIGIPEQ) tem sua ênfase nas populações expostas a riscos ambientais e estabelece ações tendo como objetivo adotar medidas de atenção integral de populações que são expostas a contaminantes químicos no que se refere aos componentes: poluição atmosférica (VIGIAR), áreas contaminadas (VIGISOLO) e substâncias químicas prioritárias - agrotóxicos, benzeno, chumbo, amianto e mercúrio (VIGIQUIM) (BRASIL, 2013).

2.3.2.1. VIGIAR

A Vigilância em Saúde Ambiental Relacionada à Qualidade do Ar (VIGIAR) tem como objetivo proteger a população exposta aos contaminantes atmosféricos. O aumento do risco de morbimortalidade da população devido principalmente às doenças cardiovasculares e respiratórias tem sido relacionado aos poluentes do ar e um sério problema de saúde pública (SALDIVA, et al., 2020).

O VIGIAR foi concebido para fornecer subsídios às políticas nacionais e locais de saúde dessa população exposta, identificando e avaliando os efeitos agudos e crônicos (RADICCHI e LEMOS, 2009).

2.3.2.2. VIGISOLO

A Vigilância em Saúde Ambiental Relacionada às Populações Expostas a Solos Contaminados (VIGISOLO) refere-se à exposição dos indivíduos às situações de risco que se derivam da contaminação química natural ou antrópica do solo. Ao se notificar áreas contaminadas por substâncias químicas devem ser realizadas ações para proteger a saúde como descrição da área contaminada e da população exposta, cadastro da área contaminada no Sistema de Informação do Solo (SISSOLO), capacitação dos profissionais do Sistema Único de Saúde e informar a população, sobre os riscos à saúde decorrente da contaminação ambiental na área (BRASIL, 2013).

2.3.2.3. VIGIQUIM

A Vigilância em Saúde Ambiental Relacionada à Segurança Química (VIGIQUIM) visa trabalhar com a prevenção de acidentes químicos e a comunicação de risco à população potencialmente exposta a eles, tais como: amianto, benzeno, agrotóxicos, mercúrio e chumbo (BRASIL, 2020).

As especificidades do VIGIQUIM buscam controlar e detectar os agravos relacionados às substâncias químicas, identificando e acompanhando a população exposta. Os agrotóxicos, por exemplo, podem afetar a saúde humana de modo direto (manuseio ou contato por ambientes contaminados) e indireto (contaminação da biota de áreas próximas às plantações agrícolas). A intoxicação e a mortalidade por acidentes químicos são considerados sérios agravantes para a saúde, sendo que a contaminação por essas substâncias também se revelam, após certo período, cancerígenos e disruptores endócrinos (BRASIL, 2020).

2.3.3. VIGIDESASTRES

A Vigilância em Saúde Ambiental Relacionada aos Riscos Decorrentes dos Desastres Naturais (VIGIDESASTRES) tem como objetivo o envolvimento do setor de saúde para reduzir os impactos dos desastres sejam eles de origem natural ou tecnológica. Portanto, dentre seus objetivos de atuação estão os desastres naturais (inundações, seca e estiagem, deslizamentos, vendavais, geadas, granizos, dentre outros), a vigilância em saúde associada aos fatores radionucleares - VIGIFIS e aos acidentes com produtos perigosos - VIGIAPP (BRASIL, 2020).

O estudo e avaliação dos desastres é uma das funções essenciais da saúde pública, pois provocam dentre outros fatores um número inesperado de óbitos, lesões ou traumatismos e enfermidades que comprometem o comportamento emocional e social das comunidades e saturam o sistema de saúde. Os desastres associados às inundações, por exemplo, além de causar afogamentos podem aumentar os casos de morbimortalidade por doenças diarréicas agudas e infecções intestinais, infecções cutâneas, leptospirose, hepatite A, tétano acidental, acidentes por animais peçonhentos, doenças transmitidas por vetores, reservatórios e hospedeiros, dentre outras (ANEXO 01) (BRASIL, 2013).

2.3.3.1. VIGIFIS

A Vigilância em Saúde Ambiental Relacionada a Fatores Físicos (VIGIFIS) visa proteger a saúde da população decorrente da exposição de radiações eletromagnéticas advindas, por exemplo, da energia elétrica, meios de telecomunicações, diagnóstico e terapia medicinais (BRASIL, 2020).

É função do VIGIFIS identificar e conceituar um conjunto de indicadores e determinar limites máximos de emissão, delimitar a presença humana nas diferentes faixas de referência para áreas de risco, segundo gênero, idade, tempo de permanência, condição clínico-sanitária, assim como propor normas a fim de dar suporte às ações de vigilância ambiental em saúde para os fatores físicos (BRASIL, 2020).

2.3.3.2. VIGIAPP

A Vigilância Ambiental em Saúde Relacionada aos Acidentes Envolvendo Produtos Perigosos (VIGIAPP) visa identificar e mapear riscos dos efeitos a saúde humana decorrentes da exposição aos produtos perigosos (BRASIL, 2020). Esses produtos são considerados como toda substância ou mistura de substâncias que, em razão das suas propriedades químicas, físicas ou toxicológicas, isoladas ou combinadas, constitui perigo para a saúde e o ambiente (RADICCHI e LEMOS, 2009).

A periculosidade intrínseca de certos produtos associada à probabilidade de acidentes agrava o risco de exposição humana seja de forma direta (trabalhadores) ou indireta (contaminação do ambiente) que pode levar a efeitos agudos ou crônicos, com uma diversidade de sinais e sintomas, e ao óbito (BRASIL, 2020). Portanto, cabe à VIGIAPP um conjunto de ações que vise prevenir a ocorrência de mazelas das populações expostas ou sob risco de exposição a esses produtos. Diante disso, o próximo item aborda os conceitos essenciais para entendimento da inter-relação dado, indicador e índice no âmbito dos agravos ambientais que interferem na saúde.

2.4. DEFINIÇÕES BÁSICAS: DADO, INDICADOR E ÍNDICE

Uma das características da criação de indicadores e índices constitui a agregação de dados e informações com vista a simplificar a complexidade inerente às inter-relações entre problemas ambientais e de saúde (BRASIL, 2011).

Na construção dos indicadores, o dado é considerado um valor quantitativo, ou seja, um número bruto que ainda não passou por nenhum tratamento estatístico (BRASIL, 2011). O dado precisa passar por um processo mais refinado de sistematização, pois por si só não consegue representar um determinado fenômeno. Essa utilidade dos dados permite a construção de um indicador, como a taxa de exposição aos agrotóxicos para cada cem mil habitantes, representando um recorte da realidade (BRASIL, 2011).

Assim como o próprio termo indica, o conceito de indicador é originário do latim, *indicare* que significa estimar, apontar, anunciar, comunicar, avaliar, descobrir. Sua utilização é para deixar mais perceptível a indicação de um estado ou fenômeno não imediatamente detectável apenas pelos dados isolados (ARANTES e PEREIRA, 2017). Os indicadores são instrumentos para simplificar, quantificar e analisar informações, de maneira a comunicar os resultados a diversos grupos de usuários (ARANTES e PEREIRA, 2017).

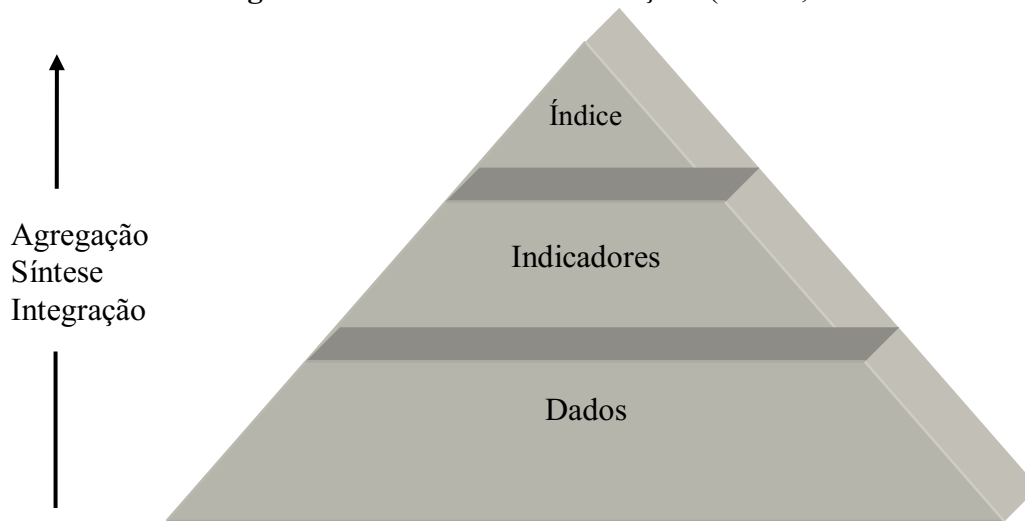
Para tanto, o indicador é constituído por um conjunto de dados ou variáveis que, submetidos a operações estatísticas informam sobre uma determinada situação. E a capacidade de sintetizar informações (medida-síntese) que os tornam ferramentas essenciais nos processos de gestão (OMS, 2022).

No setor saúde, nota-se que os indicadores utilizados envolvem aspectos demográficos, socioeconômicos, mortalidade, morbidade, recursos, cobertura e outros. Já no setor ambiental, alguns indicadores são relativamente mais recentes, e vem sendo realizados esforços internacionais e no Brasil para a produção de indicadores que permitam o monitoramento da dimensão ambiental. Em geral, envolvem aspectos relacionados à água, ar, solo, desastres naturais e outros. E o principal objetivo da construção desses indicadores é o de estruturar sistemas que articulem o monitoramento da situação ambiental aos riscos à saúde (ARANTES e PEREIRA, 2017).

O índice, por sua vez, é a agregação de dois ou mais indicadores, gerando uma equação expressa, que uma vez calculada revela um valor que define uma condição que se quer medir, ou seja, são estruturados em um modelo matemático (OPAS, 2018). Os índices são considerados indicadores compostos e são medidas-síntese por sinalizarem por meio de um valor, facilitando assim a comunicabilidade em decorrência do grande número de dados e informações (OPAS, 2018).

A figura abaixo refere à pirâmide de função que mostra como as informações devem ser agrupadas e condensadas para a obtenção de indicadores e, posteriormente de índices.

Figura 05 - Pirâmide de Informações (Dados, Indicadores e Índices)



Fonte: BRASIL, 2011. Organização: LOPES, A. F. A., 2020.

Os dados brutos (primários) estão na base da pirâmide, seguidos pelos indicadores e pelo índice no topo. A formulação de indicadores é feita por meio da análise dos dados primários obtidos, que formarão os indicadores que, por sua vez agrupados, darão origem aos índices (FERREIRA, 2011). Ressalta-se que os índices nos proporcionam respostas quantitativas, já os indicadores podem ser qualitativos ou quantitativos.

E um dos desafios da saúde pública é o de interpretar e mensurar o sistema holístico de indicadores que envolvem a inter-relação entre os aspectos ambientais e os efeitos na saúde humana, portanto, esse trabalho utilizou a análise estatística multivariada para trabalhar com esse número elevado de dados (853 observações e a princípio, 130 variáveis).

2.5. INDICADORES AMBIENTAIS *VERSUS* INDICADORES DE SAÚDE

Os indicadores ambientais se referem à situação ambiental, ou seja, informam, principalmente sobre a qualidade do ambiente (água, ar e solo), com relação aos aspectos físico-químicos e das condições de biodiversidade (ARANTES e PEREIRA, 2017).

Estes indicadores descrevem de forma isolada apenas o ambiente, seja ele natural, ocupacional, urbano ou rural com ênfase para a gestão ambiental sem mostrar a inter-relação com a saúde. Eles têm como finalidade auxiliar na avaliação da efetividade das ações na área ambiental, tais como: atmosfera e mudanças climáticas, biodiversidade e florestas, oceanos e mares, qualidade ambiental, recursos hídricos, terra e solos (MMA, 2014).

Por conseguinte, os indicadores de saúde refletem o estado da saúde humana, ou seja, estão associados às doenças específicas (infecciosas, parasitárias, respiratórias, etc) relacionadas às características da população como idade, sexo, escolaridade (BRASIL, 2011).

Devido ao avanço do conceito de saúde e seus determinantes sociais, a avaliação do estado de saúde passou a ser ampla, além dos dados sobre mortalidade e sobrevivência, passou-se também a avaliar dados sobre morbidade, incapacidade, serviços, qualidade da atenção e fatores ambientais – não apenas o ambiente meramente natural, mas também o antrópico. Desse modo, avaliados em conjunto passam a refletir a situação sanitária de uma população (RIPSA, 2008).

Diante do exposto, surgem os indicadores em Saúde Ambiental ao integrarem os aspectos ambientais para a compreensão e análise da saúde pública.

2.6. INDICADORES DE SAÚDE AMBIENTAL

Os indicadores de saúde ambiental podem ser definidos como uma medida que integra a relação ambiente e saúde, ou seja, são capazes de sintetizar, de maneira compreensível e relevante os fatores de risco ambientais à saúde humana. Com o uso de ferramentas, como a estatística, epidemiologia e técnicas computacionais, eles podem ser estudados e sintetizados em informações compreensíveis para a tomada de decisões (ARANTES e PEREIRA, 2017).

2.6.1. Critérios para o desenvolvimento de indicadores de saúde ambiental

Alguns critérios são relevantes para que os indicadores de saúde ambiental sejam aceitos como medida de referência para os gestores e tomadores de decisões. Para Arantes e Pereira (2017), mesmo sendo considerada uma representação simplificada da realidade, sua construção não é tarefa simples, pois envolvem uma gama de dados, informações e dimensões que devem ser analisadas com cautela.

Portanto, alguns critérios são fundamentais para seu planejamento e construção, tais como: relevância social, validade, confiabilidade, cobertura, sensibilidade, especificidade, inteligibilidade de sua construção, comunicabilidade, periodicidade de atualização, factibilidade para obtenção, desagregabilidade e historicidade (BRASIL, 2011). O quadro a seguir apresenta a descrição de cada um desses critérios.

Quadro 01- Propriedades desejáveis dos indicadores

CRITÉRIO	DESCRIÇÃO
Relevância Social	A relevância social diz respeito a sua utilidade, legitimando seu emprego no processo de análise, formulação e implementação de políticas. Além de ser facilmente compreendido e aplicável por potenciais usuários;
Validade	É a capacidade de medir o que se pretende e ser representativo da realidade;
Confiabilidade	Está relacionada com a qualidade dos dados usados em sua construção, assim como à credibilidade da instituição que produz as estatísticas;
Cobertura	Refere-se à cobertura espacial e populacional do indicador;
Sensibilidade	Expressa a capacidade de mostrar se ocorreram mudanças significativas nos fatores que afetam as condições sociais, ambientais e de saúde ao longo do tempo;
Especificidade	Diz respeito ao grau de associação existente entre os dados utilizados para a construção de um indicador. Se houver baixa associação entre os indicadores constitutivos do índice, este pode não ser específico o suficiente para o monitoramento desejado;

Inteligibilidade de sua construção	Relacionada com a transparência da metodologia empregada na sua construção;
Comunicabilidade	O indicador deve ser facilmente compreendido para que decisões políticas sejam efetivas;
Periodicidade de atualização	Os indicadores devem ser atualizados e disponibilizados com certa regularidade;
Factibilidade para obtenção	A obtenção dos dados deve ser viável, ou seja, ao alcance dos usuários (acessíveis e aplicáveis);
Desagregabilidade	É a capacidade de um indicador se relacionar a grupos populacionais de interesse, espaços geográficos ou vulnerabilidades sociais específicas;
Historicidade	Refere-se à capacidade de um indicador de dispor séries históricas, comparando valores do presente a situações do passado;

Fonte: BRASIL, 2011. Adaptação e Organização: LOPES, A. F. A., 2020.

É importante salientar que um bom indicador deve buscar contemplar a maioria desses critérios, no entanto, existem limitações - como a falta de recursos, a necessidade de compatibilidade com outros indicadores e as dificuldades metodológicas - que fazem com que todas essas características nem sempre sejam alcançadas.

2.7. FONTES DE DADOS NACIONAIS, ESTADUAIS E MUNICIPAIS E SISTEMAS DE INFORMAÇÃO: PRINCIPAIS CARACTERÍSTICAS

O sistema Brasileiro de estatística é alimentado por órgãos orientados para o levantamento e sistematização de um conjunto de dados, tais como o Instituto Brasileiro de Geografia e Estatística (IBGE); o Ministério da Saúde (MS) que mantém o DATASUS; o Ministério do Meio Ambiente (MMA); o Ministério da Educação que mantém o Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP), a Fundação Getúlio Vargas (FGV) e o Instituto de Pesquisas Econômicas e Aplicadas (IPEA) (BRASIL, 2013).

O IBGE é o órgão central e coordenador do Sistema Estatístico Nacional e entre algumas das principais pesquisas realizadas pela área da Saúde são os censos demográficos, as contagens populacionais, a Pesquisa Nacional por Amostra de Domicílios (PNAD), a Pesquisa Nacional de Saneamento Básico (PNSB) e a Pesquisa de Assistência Médico-Sanitária (MAS) (BRASIL, 2011).

As agências estaduais e municipais de estatísticas também apresentam processos de produção, compilação e disseminação das estatísticas públicas para subsidiar a definição de políticas públicas em diversas áreas. Na esfera estadual, estão sob a responsabilidade das secretarias de planejamento ou de órgãos específicos a elas vinculados, em Minas Gerais, por exemplo, têm-se destaque para a fundação João Pinheiro. No âmbito municipal, a principal aplicação das estatísticas públicas se dá na elaboração de planos diretores e planos plurianuais de investimentos (BRASIL, 2011).

Algumas Instituições Internacionais que também possuem destaque na produção de dados são a Organização Mundial da Saúde (OMS); a Organização Pan-Americana da Saúde (OPAS); o Programa das Nações Unidas para o Meio Ambiente (PNUMA), a Organização para a Cooperação e Desenvolvimento Econômico (OCDE); a Organização das Nações Unidas para a Educação, a Ciência e a Cultura (UNESCO) (OPAS, 2018).

Os Sistemas de Informações representam um conjunto de elementos ligados à produção, análise e divulgação de dados, utilizando recursos pessoais (técnicos) e materiais para o processamento dos dados (OMS, 2022).

2.7.1. Sistemas de Informações em Saúde

Na área da saúde, os Sistemas de Informações em Saúde (quadro 02) agregam informações relevantes dos aspectos da saúde individual e coletiva, informando sobre agravos e óbitos (BRASIL, 2011).

Quadro 02- Lista dos principais Sistemas de Informações em Saúde

Sistemas de Informações em Saúde	Sistema de Informações sobre Agravos de Notificação (SINAN)	www.datasus.gov.br
	Sistema de Informações sobre Mortalidade (SIM)	
	Sistema de Informações de Nascidos Vivos (SINASC)	
	Sistema de Informação de Vigilância Alimentar e Nutricional (SINVAN)	
	Sistema de Informações Hospitalares (SIH)	
	Sistema de Informações Ambulatoriais (SAI)	
	Sistema Nacional de Informação Tóxico Farmacológica (SINITOX)	

Fonte: BRASIL, 2011. Organização: LOPES, A. F. A., 2020.

Outro destaque é para a Rede Interagencial de Informação para a Saúde (RIPSA) que contribui para disseminar dados e informações referentes à saúde por meio de um conjunto de indicadores (BRASIL, 2011).

2.7.2. Sistemas de Informações Ambientais

Na área ambiental, o Sistema Nacional de Informação sobre Meio Ambiente é responsável pela produção, coleta e análise das informações ambientais (quadro 03).

Quadro 03- Lista dos principais Sistemas de Informações Ambientais

Sistemas de Informações Ambientais	Instituto Nacional de Pesquisas Espaciais (INPE)	www.mma.gov.br
	Sistema de Informações Ambientais Integrado à Saúde Ambiental (SISAM)	
	Fundação Estadual do Meio Ambiente (FEAM)	
	Sistema de Informação de Vigilância da Qualidade da Água (SISÁGUA)	
	Sistema de Informação de Vigilância da Qualidade do Solo (SISSOLO)	
	Sistema Nacional de Informação sobre Saneamento (SNIS)	

Fonte: BRASIL, 2011. Organização: LOPES, A. F. A., 2020.

Ressalta-se também que o Ministério do Meio Ambiente (MMA), o Instituto Brasileiro do Meio Ambiente e Recursos Naturais Renováveis (IBAMA) e o Programa das Nações Unidas para o Meio Ambiente (PNUMA) desenvolveram o Relatório GEO Brasil 2002 envolvendo temas importantes como biodiversidade, solo, água, atmosfera, ambientes marinhos, desastres (BRASIL, 2011).

Diante do exposto, nota-se o esforço do MMA para a consolidação dos indicadores ambientais como ferramenta de auxílio para ações de preservação e conservação, uso responsável dos recursos e medidas de planejamento.

2.7.3. Sistema Nacional de Informação em Saúde Ambiental

A Análise de Situação em Saúde Ambiental e Saúde do Trabalhador (ASISAST) da (CGVAM/SVS/MS) se destaca por proporcionar um sistema nacional de informação dos aspectos referentes à saúde e ao meio ambiente. Integra ao ASISAST o Painel de Informações em Saúde Ambiental e Saúde do Trabalhador (PISAST), que tem como objetivo disponibilizar para os usuários dados e informações sobre saúde, meio ambiente, aspectos socioeconômicos, saúde ambiental e saúde do trabalhador (BRASIL, 2011).

Dentre as colaborações importantes nessa área tem-se destaque para o documento Vigilância em Saúde Ambiental: dados e indicadores selecionados, uma iniciativa da Secretaria de Vigilância em Saúde, por intermédio do Departamento de Vigilância em Saúde Ambiental e Saúde do Trabalhador. Seu objetivo é apresentar dados e indicadores de interesse para a Vigilância em Saúde Ambiental por meio da metodologia desenvolvida pela Organização Mundial da Saúde (OMS), o modelo Força Motriz - Pressão - Situação - Exposição – Efeito – Ação (FPSEEA) (BRASIL, 2011).

2.8. MODELOS DE ORGANIZAÇÃO E ANÁLISE DOS INDICADORES

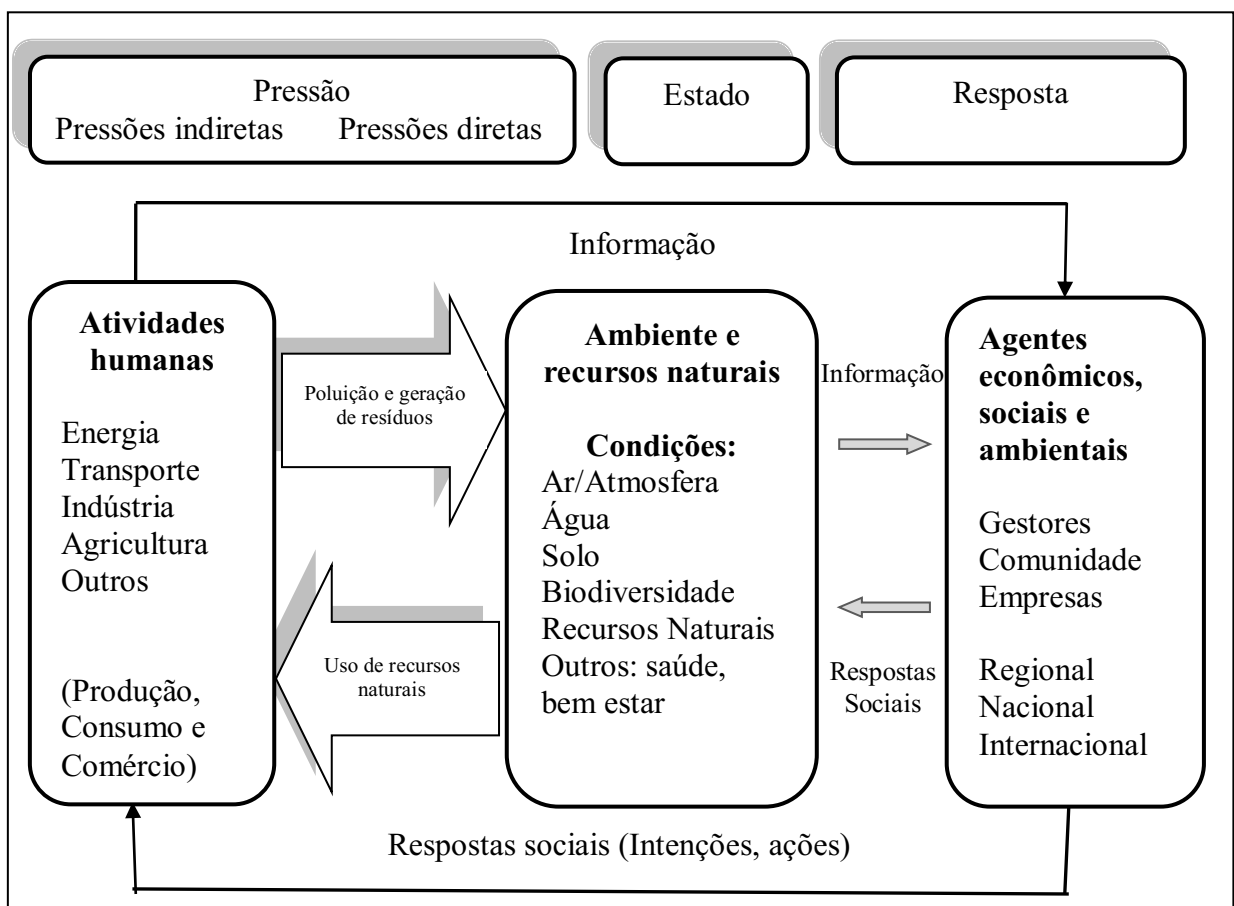
Os modelos de organização dos indicadores podem contribuir para observações e análises, explicando as inter-relações entre os seus diversos níveis. Modelos teórico-conceituais vêm sendo desenvolvidos e aperfeiçoados ao longo do tempo para entender os reflexos das condições ambientais na saúde humana e os indicadores vem sendo agregados a eles para proporcionar análises integradas (BRASIL, 2011).

Nessa seção apresentam-se alguns exemplos de modelos de organização de indicadores na área ambiental que foram elaborados em outros países e que, atualmente, fundamentam-se para a construção de metodologias semelhantes no Brasil.

2.8.1. Modelo Pressão - Estado - Resposta - PER

O primeiro modelo desenvolvido para a análise da situação ambiental foi o Pressão – Estado – Resposta (PER), criado pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE), em 1993 (FERREIRA, 2011). Constituiu-se por indicadores de pressão (P), relacionado às pressões das atividades antrópicas sobre o meio ambiente; indicadores de estado (E) que se refere à situação do ambiente; e os indicadores de resposta (R) que mostra as respostas da sociedade civil ou gestão pública diante das mudanças ambientais (FERREIRA, 2011) (Figura 06).

Figura 06 - Modelo de Organização dos Indicadores: Pressão - Estado - Resposta



Fonte: PNUMA, 2002. Organização: LOPES, A. F. A., 2020.

De acordo com os estudos de Sobral e Freitas (2010), pode-se considerar que o modelo baseia-se no seguinte conceito de causalidade: as atividades antrópicas exercem pressões sobre o ambiente, modificando sua qualidade e a quantidade de recursos naturais; a sociedade, por sua vez, responde a essas mudanças por intermédio de políticas ambientais.

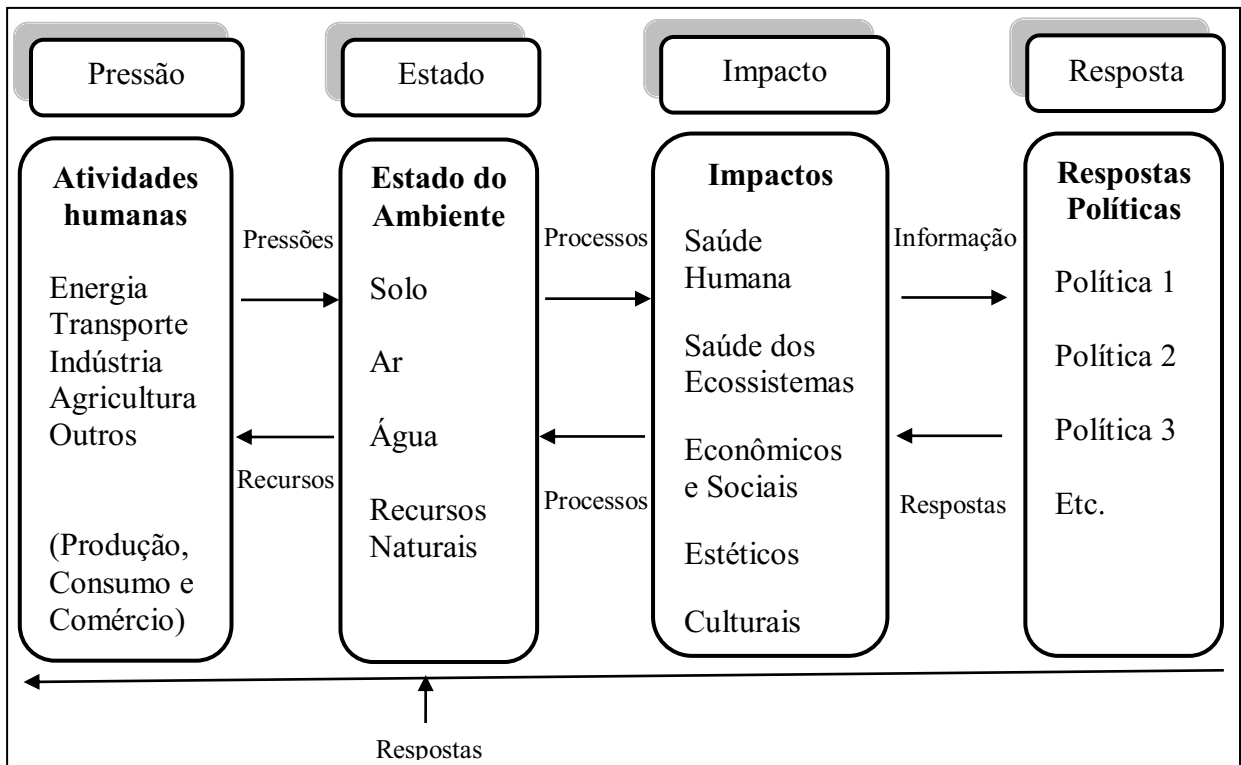
Em geral, essa ferramenta enfatiza a questão ambiental, abrangendo a ação antrópica sobre o ambiente. No entanto, a sua fragilidade está ao considerar apenas as ações humanas, descartando-se as reações naturais aos fenômenos, além do mais, não analisa aspectos sanitários - o que tem tornado o seu uso inviável para essa finalidade.

2.8.2. Modelo Pressão - Estado - Impacto - Resposta - PEIR

Sucessivamente à criação do modelo PER pela OCDE, o Programa das Nações Unidas para o Meio Ambiente criou na década de 1990 uma abordagem conceitual do modelo PEIR, acrescentando o componente “impacto” como desdobramento do “estado”. O objetivo foi de avaliar os efeitos das pressões exercidas sobre as condições ambientais e os possíveis impactos no bem-estar social ou o ecossistema (BRASIL, 2011).

De acordo com Sobral e Freitas (2010), esse modelo apresenta estágios de relações causais em que as atividades humanas exercem pressões (emissões, falta de abastecimento de água) e estado (físico, químico e biológico), causando impactos no ambiente e na saúde. Então, a sociedade responde a essas mudanças por meio de políticas que visem mitigar ou recuperar os danos causados (monitoramento, fixação de metas, indicadores) (Figura 07).

Figura 07 - Modelo de Organização dos Indicadores: Pressão – Estado – Impacto – Resposta



Fonte: BRASIL, 2011. Organização: LOPES, A. F. A., 2020.

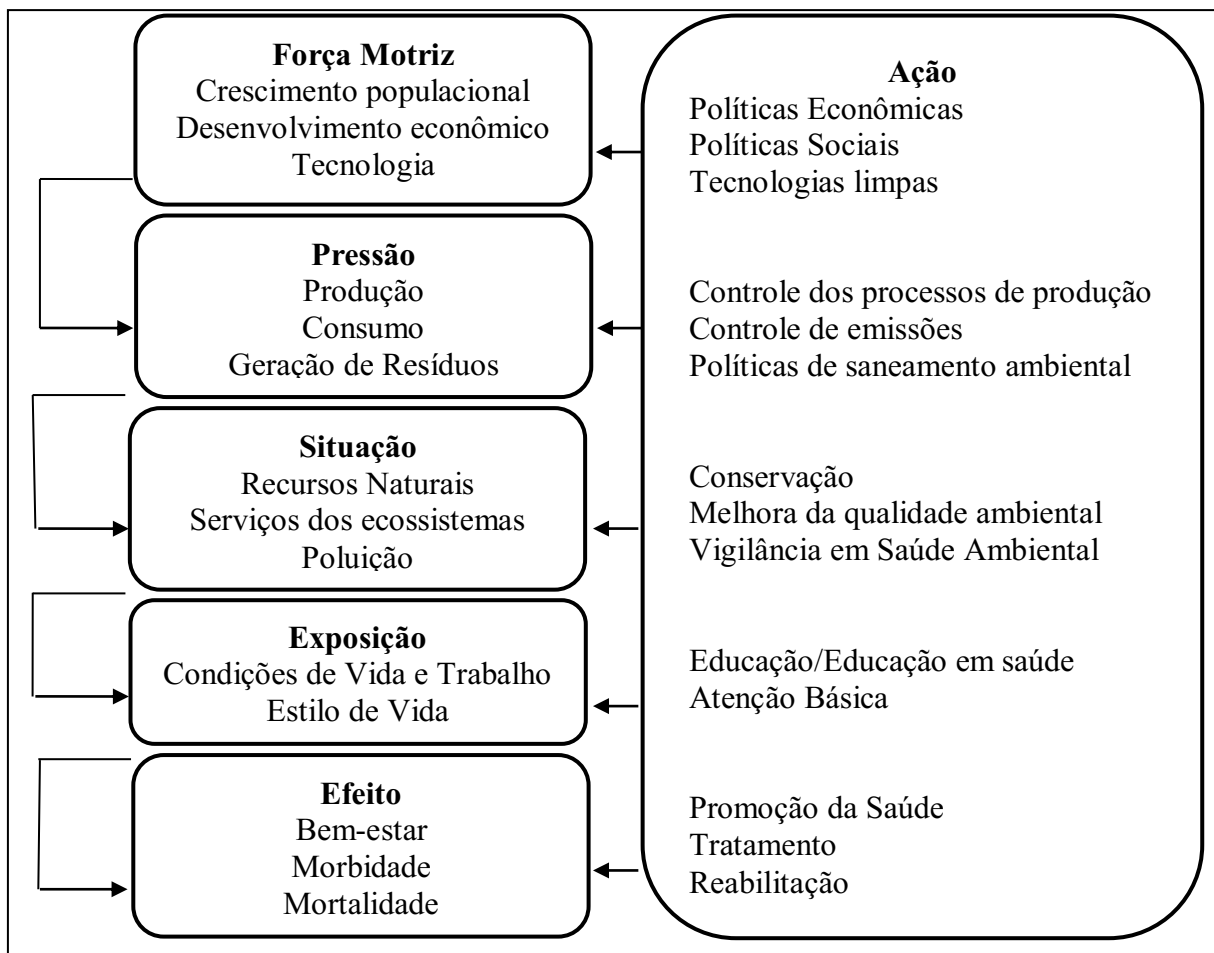
Devido à descrição limitada à exposição, este modelo não pode identificar no caminho, múltiplos pontos de entrada para medidas de ação, além disso, outra crítica recebida é a falta de incorporação dos efeitos à saúde. Para tanto, o próximo modelo foi desenvolvido para atender a complexidade da análise causa-efeito das relações entre saúde e ambiente.

2.8.3. Modelo Força Motriz - Pressão - Situação - Exposição – Efeito – Ação

A abordagem Força Motriz - Pressão - Situação - Exposição – Efeito – Ação (FPSEEA) foi desenvolvida pela Organização Mundial da Saúde, juntamente com o Programa das Nações Unidas para o Meio Ambiente (PNUMA) (SOBRAL e FREITAS, 2010).

A estrutura FPSEEA aponta que as forças motrizes que representam a forma como uma sociedade se organiza política e culturalmente geram pressões que alteram a situação ambiental e conseqüentemente, por meio das formas de exposição causam efeitos adversos na saúde da população. Por fim, é necessária a proposta de ações para a introdução de medidas de prevenção e controle dos riscos (Figura 08).

Figura 08- Modelo de organização de indicadores FPSEEA



Fonte: SOBRAL e FREITAS, 2010. Organização: LOPES, A. F. A., 2020.

As forças motrizes representam a escala macro e que influenciam os vários processos ambientais que vão interferir nas condições de vida e bem-estar da população (como por exemplo, o crescimento da população, o processo de desenvolvimento econômico e tecnológico, a rapidez da urbanização e industrialização) (BRASIL, 2011).

A pressão como nos modelos anteriores, corresponde sobre as ações antrópicas no ambiente como os processos de produção e consumo de bens e serviços, assim como a disposição de resíduos gerados. Como consequências, essas pressões, alteram/modificam/agravam a situação do ambiente (como por exemplo, agravamento da poluição atmosférica, da contaminação da água para consumo humano, das enchentes) resultando em efeitos negativos para a saúde dos indivíduos.

No campo exposição, observa-se como a saúde individual ou coletiva é afetada pelas causas adversas do meio ambiente. Portanto, são analisadas as condições de vida e trabalho das pessoas, as doses de exposição, as vias de entradas de substâncias. É considerado um conceito chave em saúde ambiental, pois estabelece uma inter-relação entre condições ambientais e seus efeitos para a saúde de grupos populacionais em um contexto espacial e temporal. Pode-se citar como exemplo desse contexto indicador como: população residente em domicílios com instalações inadequadas de esgoto, população residente em domicílios sem coleta de lixo, população residente em domicílios sem água tratada (BRASIL, 2011).

Já os efeitos sobre a saúde podem se manifestar quando o indivíduo se submete a uma exposição. Os indicadores de efeitos mais utilizados referem-se à morbimortalidade da população em geral, tais como: internações e óbitos por doenças diarreicas, óbitos por infecção respiratória, intoxicações por agrotóxicos (VEIGA, 2020). Por último, no campo das ações, avaliam-se a formulação e implantação de políticas econômicas e sociais, as tecnologias limpas, promoção da saúde, gestão de risco, vigilância em saúde ambiental, melhorias ambientais e programas educativos.

Portanto, nota-se que além dos indicadores sociais e econômicos que são frequentemente utilizados na avaliação da situação de saúde, esse modelo incorpora os indicadores ambientais referentes à integridade ecológica dos ecossistemas. Seu destaque é reconhecer que alterações na situação ambiental causada pelas forças motrizes e pressões atingem diferentes meios como o solo, a água e o ar resultando em exposições que poderão ter efeito direto ou indireto na saúde da população em curto, médio ou longo prazo.

Entretanto, alguns autores como Arantes e Pereira (2017) consideram que uma das dificuldades de modelos como esse é a organização do grande número de indicadores de forma a permitir uma análise integrada dos diferentes elementos da realidade. Desse modo, ressalta-se que o modelo FPSEEA deve ser utilizado como ferramenta auxiliar para a avaliação da situação de saúde ambiental, pois, embora objetivando uma visão integrada dos indicadores, por si só ela não consegue avaliar toda a complexidade existente da inter-relação entre esses indicadores.

E para complementar esse modelo, essa pesquisa visa empregar métodos estatísticos baseados em análise multivariada de dados para examinar a estrutura de inter-relações dessas múltiplas variáveis e determinar o melhor modelo explicativo das variáveis ambientais que inferem efeitos óbitos na saúde humana.

2.9. CONCEITOS BÁSICOS ESTATÍSTICOS: ANÁLISE MULTIVARIADA DE DADOS

Com a introdução da computação eletrônica, os estudos epidemiológicos se expandiram, e com o uso das análises multivariadas passou a ser possível trabalhar com grandes bancos de dados e controlar as variáveis confundidoras. Apesar de a estatística multivariada ter surgido por volta de 1901, nos dias de hoje essa técnica é bem sucedida devido ao auxílio dos programas computacionais (VICINI, 2005).

A análise multivariada se caracteriza como uma técnica estatística em que se pode realizar a análise simultânea entre mais de duas variáveis. Além disso, “todas as variáveis devem ser aleatórias e inter-relacionadas de tal maneira que seus diferentes efeitos não podem ser significativamente interpretados em separado” (HAIR, 2009, p.23). Pode-se considerar também que o caráter multivariado se caracteriza pelas múltiplas combinações de variáveis e não apenas no número de variáveis e observações.

Para Fávero *et.al.*, (2009), muitas técnicas multivariadas são extensões da análise univariada, ou seja, análise de uma única variável e da análise bivariada para analisar simultaneamente duas variáveis.

A análise fatorial, por exemplo, é uma técnica exclusiva que lida com aspectos multivariados, pois resulta no agrupamento de um conjunto de variáveis. A partir das técnicas multivariadas é possível analisar a relação que existe entre as variáveis e obter um entendimento adequado para a tomada de decisões.

Para tanto, primeiramente, é importante uma compreensão conceitual do constructo da técnica multivariada, assim, os próximos subitens abordarão aspectos sobre a variável estatística, os tipos de escalas de medida e questões estatísticas de níveis de validade/confiabilidade e testes de significância - conceitos estes fundamentais para uma aplicação satisfatória da análise multivariada.

2.9.1. A variável estatística

O constructo da análise multivariada é a variável estatística, em que há uma combinação linear de variáveis (especificadas pelo pesquisador) com seus respectivos pesos (determinados pela técnica estatística). De acordo com Hair (2009), uma variável estatística de n variáveis (X_1 até X_n) pode ser definida matematicamente da seguinte maneira:

$$\text{Valor da variável estatística} = w_1X_1 + w_2X_2 + w_3X_3 + \dots + w_nX_n$$

Onde:

x_n = Variável observada

w_n = peso determinado pelo método estatístico

A partir disso, obtêm-se um único valor da combinação de todas as variáveis. Para Mingoti (2005) na análise multivariada de dados, as combinações lineares são importantes uma vez que sintetizam as informações das n variáveis originais.

2.9.2. Tipos de escalas de medida

Os dados podem ser classificados como não-métricos (qualitativos) e métricos (quantitativos) e cabe ao pesquisador definir o tipo de medida para cada variável.

2.9.2.1. Escalas de medidas não-métricas

Referem-se a determinado tipo ou natureza e indicam a presença ou ausência de uma característica ou propriedade. As escalas de medidas não-métricas são realizadas com uma escala **nominal** ou **ordinal**. As escalas nominais são conhecidas também como escalas categóricas e designam números para identificar indivíduos ou objetos (HAIR, 2009). No

entanto, esses números não apresentam significado quantitativo, por exemplo, no caso do sexo feminino e masculino, pode-se definir o número 1 para mulheres e o número 2 para homens. Portanto, esses números servem para representar categorias ou classes.

Nas escalas ordinais, as variáveis podem ser ordenadas, permitindo a comparação entre indivíduos ou objetos com outro em uma relação “maior que” ou “menor que” (HAIR, 2009). Por exemplo, pode-se ordenar o nível de satisfação em relação a novos diversos produtos em um número 3 para muito satisfeito, 2 para satisfeito e 1 para pouco satisfeito. Nesse caso, os números indicam apenas posições relativas em uma série ordenada e não a quantia de valores (HAIR, 2009). Os cálculos que podem ser utilizados para essas variáveis ordinais referem-se apenas à informação da ordem.

2.9.2.2. Escalas de medidas métricas

As escalas de medida métricas referem-se à quantidade de determinado atributo, ou seja, envolvem quantia ou magnitude, ao contrário das medidas não-métricas. E as duas escalas de medidas métricas são as **intervalares** e de **razão**, ambas apresentam alto nível de precisão de medida, podendo ser realizada quase todas as operações matemáticas (HAIR, 2009).

A única diferença entre as duas é que as escalas intervalares possuem um ponto zero arbitrário e as de razão têm um ponto zero absoluto. As escalas intervalares mais conhecidas são as de temperatura Fahrenheit e Celsius, portanto, nesse tipo de escala, podem-se utilizar valores negativos (HAIR, 2009). Por sua vez, as escalas de razão apresentam um ponto zero absoluto, como exemplo tem-se as balanças de banheiro usadas para medir peso. Nesse caso, pode-se relacionar em termos de múltiplos, por exemplo, 100 libras é duas vezes o peso de 50 libras (HAIR, 2009). Portanto, nesse tipo de escala é permitido o uso de todos os tipos de cálculos como múltiplos, submúltiplos, soma, diferença, “maior que” ou “menor que”, moda, percentual e mediana - o que permite uma capacidade de análise ampla (HAIR, 2009).

Nesse sentido, é importante definir quais escalas de medidas serão utilizadas em cada variável para evitar usos incorretos entre as medidas, tais como os erros de medida e também para que se busque a técnica estatística mais apropriada para o estudo. Os erros de medidas se caracterizam como o grau em que as medidas observadas se distanciam dos valores reais e ocorrem entre outros aspectos devido a erros na entrada de dados, imprecisões das medidas e informações imprecisas dos respondentes (HAIR, 2009).

2.9.3. Validade e Confiabilidade

Tendo em vista a finalidade de reduzir o erro de medida, devem considerar dois importantes conceitos: a validade e a confiabilidade. Ambas são estreitamente relacionadas e desempenham papel complementar. “Validade é o grau em quem uma medida representa precisamente aquilo que se espera” (HAIR, 2009, p. 26). Ou seja, está relacionada com o que deve ser medido (grau de certeza daquilo que se quer medir).

Um estudo válido é aquele em que realmente se conseguiu medir aquilo que foi proposto medir. A validade pode ser classificada em interna e externa. A validade interna se refere a uma medida válida para uma amostra em particular, ou seja, os resultados observados representam uma verdade para a população que está sendo analisada. Por sua vez, a validade externa permite generalizar os resultados de uma amostra em particular para a população como um todo (PATINO e FERREIRA, 2018).

A confiabilidade por sua vez, representa o grau em que a variável analisada mede o valor verdadeiro e assim estando “livre de erro” (HAIR, 2009). As medidas mais confiáveis apresentam maior consistência do que as não confiáveis. Portanto, a confiabilidade está relacionada com a exatidão do procedimento de mensuração (modo de como é medido) e com a precisão (reprodutibilidade do estudo).

2.9.4. Testes de Significância

Para Mingoti (2005) um teste de significância testa uma hipótese específica, usando dados amostrais para decidir sobre sua validade. Além da estatística é preciso uma hipótese nula (H_0) a ser testada e uma hipótese alternativa (H_1) para qual se procura a evidência. Para tanto, um teste de significância avaliará a força da evidência contra H_0 .

O nível de significância α conhecido como erro tipo I é a probabilidade de rejeitar a hipótese nula quando ela é verdadeira, ou seja, o teste exibe significância estatística quando essa não está presente, chamado de “falso positivo” (HAIR, 2009). O erro tipo II ou β é “a probabilidade de não rejeitar a hipótese nula quando na realidade essa é falsa” (HAIR, 2009, p. 27), ou seja, o teste não exibe significância estatística quando na verdade essa está presente, chamado de “falso negativo”.

Já o poder de inferência estatística é a “probabilidade de rejeitar corretamente a hipótese nula quando esta deve ser rejeitada” (HAIR, 2009, p. 27). Assim, “poder é a probabilidade de

a significância estatística ser indicada se estiver presente” (HAIR, 2009, p. 27). E apesar de α estabelecer o nível de significância estatística aceitável, é o nível de poder que garante em encontrar as diferenças ou efeitos se eles realmente existirem. E não dá para fixar α e β em níveis aceitáveis porque ambos os erros são inversamente relacionados, ou seja, à medida que o erro tipo I se aproxima de zero, a probabilidade de um erro tipo II aumenta. E reduzir erros do tipo I, diminui o poder estatístico (MANLY, 2008). Assim sendo, o pesquisador deve considerar que o melhor resultado possível decorre de um equilíbrio entre o nível α e poder resultante (HAIR, 2009).

No entanto, nem sempre níveis elevados de poder são alcançados, uma vez que dependem do alfa (α), do tamanho do efeito e do tamanho da amostra. Como já supracitado, quando α se torna mais restritivo, o poder também reduz, uma vez que o erro tipo II aumenta. “Portanto, quando o pesquisador reduz a chance de incorretamente dizer que um efeito é significativo quando não o é, a probabilidade de corretamente encontrar um efeito também diminui” (HAIR, 2009, p. 27). Ou seja, quando se reduz o erro tipo I ou α , aumenta-se o erro tipo II ou β , assim sendo, o poder diminui (capacidade de encontrar efeito).

Em testes de hipóteses estatísticas, afirma-se que o resultado é estatisticamente significativo quando o p-valor é menor do que o nível de significância α definido no estudo. Esses níveis são fixados por diretrizes convencionais no valor de 0,05 ou 0,01, dependendo da área de estudo (BORROR, 2009). Desse modo, o p-valor pode ser considerado como a expressão numérica da certeza, ou seja, quanto menor ele for maior será a certeza de que a hipótese nula seja falsa e assim rejeitá-la.

Consecutivamente, o tamanho de efeito se refere a uma estatística descritiva que complementa o teste de significância e caracteriza na verdadeira magnitude do efeito de interesse (LINDENAU e GUIMARÃES, 2012). Por exemplo, de acordo com Hair (2009), uma diferença de médias entre dois grupos ou a correlação entre variáveis na população é denominado tamanho do efeito e são definidos em termos padronizados para facilitar a comparação.

“As diferenças de média são dadas em termos de desvios-padrão, de modo que um tamanho de efeito de 0,5 indica que a diferença de média é metade de um desvio-padrão. Para correlações, o tamanho do efeito é baseado na real correlação entre as variáveis” (HAIR, 2009, p. 27).

Com relação ao tamanho da amostra, em qualquer nível alfa, as amostras maiores produzem maior poder estatístico. No entanto, é necessário ter cautela no caso de “poder excessivo”, pois, ao aumentar o tamanho amostral, os efeitos menores serão percebidos como estatisticamente significantes - tornando o teste exageradamente sensível. Por outro lado, amostras reduzidas tornam o teste insensível (VIEIRA, 2008). Assim sendo, o tamanho da amostra poderá impactar no teste estatístico por isso é importante definir alguns aspectos principais como a margem do erro, o poder desejado para o teste, tempo, recursos financeiros, assim como obtenção dos dados.

Diante disso, as relações entre alfa, tamanho da amostra, tamanho do efeito e poder são consideradas complexas. Para tanto, Cohen *apud* Hair (2009) apresenta uma orientação para níveis aceitáveis de poder e sugere que os estudos devem ser planejados para atingir níveis α de pelo menos 0,05 com níveis de poder de 80% - considerando simultaneamente os fatores alfa, tamanho da amostra e tamanho do efeito. Primeiramente, estima-se o tamanho esperado do efeito para posteriormente selecionar o tamanho da amostra e alfa para que se alcance o nível de poder desejado. Ressalta-se que essa é uma sugestão e não uma regra, pois existem casos específicos em que não é possível aumentar o tamanho da amostra (como no caso de doenças raras), assim, é necessário fazer ajustes no valor do alpha ($> 0,05$) para que se possa aumentar o poder e encontrar o efeito.

2.10. CLASSIFICAÇÃO E TIPOS DE TÉCNICAS MULTIVARIADAS

As técnicas multivariadas permitem analisar um vasto conjunto de variáveis e podem ser classificadas em técnicas de dependência e interdependência (FERREIRA, 2008). Uma técnica de dependência é aquela em que uma variável ou conjunto de variáveis (dependentes) são previstas ou explicadas por outras variáveis (as independentes). Ou seja, as variáveis são dependentes do comportamento de outras. Têm-se como técnicas de dependência a regressão múltipla, análise conjunta, correlação canônica, análise multivariada de variância, análise discriminante e modelagem de equações estruturais (BAKKE, LEITE e SILVA, 2008).

Por outro lado, a técnica de interdependência é considerada aquela em que nenhuma variável ou grupo de variáveis são classificadas como dependentes ou independentes, sendo que o procedimento envolve a análise de todas em conjunto (VIEIRA, 2008). A análise fatorial, a análise de Agrupamentos, o escalonamento multidimensional e a análise de correspondência são exemplos dessa técnica.

Portanto, as variáveis dependentes (y) são consideradas resposta e são valores a serem explicados por outras variáveis (as independentes). Assim, as variáveis independentes (x) são aquelas que influenciam, afetam, prevêm ou explicam outras variáveis, no caso as dependentes (TRIOLA, 2013).

Dentre as técnicas mais estabelecidas de acordo com Hair (2009), se destacam:

- 1- Análise Fatorial (Análise de Componentes Principais e Análise dos Fatores Comuns);
- 2- Regressão Múltipla e Correlação Múltipla;
- 3- Análise Discriminante Múltipla e Regressão Logística;
- 4- Análise de Correlação Canônica;
- 5- Análise Multivariada de Variância e Covariância;
- 6- Análise Conjunta;
- 7- Análise de Agrupamentos;
- 8- Mapeamento Perceptual, também conhecido como Escalonamento Multidimensional;
- 9- Análise de Correspondência;
- 10- Modelagem de Equações Estruturais e Análise Fatorial Confirmatória;

Por conseguinte, serão apresentadas brevemente as características da Análise Fatorial e da Regressão Linear Múltipla, assim como os objetivos para sua aplicação, técnicas estas utilizadas nessa pesquisa. Ressalta-se que as etapas para a sua realização são apresentadas na seção “Resultados” (artigo 2 e artigo 3). Para as demais técnicas de análise supracitadas recomenda-se consultar a obra de Hair (2009) para uma análise minuciosa de cada uma, tendo em vista que a escolha da análise apropriada para cada tipo de estudo, bem como seu uso adequado é fundamental para a geração de resultados corretos e confiáveis.

2.10.1. ANÁLISE FATORIAL (ANÁLISE DE COMPONENTES PRINCIPAIS E ANÁLISE DOS FATORES COMUNS)

Essa técnica estatística multivariada de interdependência é utilizada com a finalidade de simplificar os dados, ou seja, reduzir o número de variáveis iniciais com a menor perda de informação. A Análise Fatorial (AF) inclui a Análise de Componentes Principais (ACP) e Análise dos Fatores Comuns (AFC) (MINGOTI, 2005).

A Análise dos Fatores Comuns (AFC) tem como propósito principal definir a estrutura latente, ou seja, ela permite analisar como o grande número de variáveis iniciais se inter-relacionam. Assim, pode-se definir o conjunto de variáveis fortemente relacionadas que podem ser agrupadas em fatores, permitindo o resumo de dados. Para Mingoti (2005), os fatores condensam as várias variáveis em um conjunto menor e podendo ser uma variável que não foi percebida no estudo. Por sua vez, a Análise de Componentes Principais (ACP) é adequada para encontrar a variável mais representativa, típica ou marcadora do grupo, propiciando a redução dos dados.

O agrupamento das variáveis é realizado de acordo com suas correlações, assim, são gerados grupos com variáveis altamente correlacionadas entre si e também com baixas correlações com um grupo que seja diferente. Cada grupo dessas variáveis apresenta um constructo ou fator, responsável pelas correlações observadas (BAKKE, LEITE e SILVA, 2008). Assim, essa técnica sumariza o conjunto de variáveis e aponta os fatores latentes em um número menor do que as variáveis iniciais (originais) inseridas no estudo.

Vale lembrar que essa técnica considera simultaneamente todas as variáveis, empregando o conceito de variável estatística como supracitado, ou seja, a composição linear de variáveis marcadoras que permitem maximizar a explicação do conjunto original inicial.

A Análise Fatorial também permite a criação de uma nova medida composta para representar cada grupo de variáveis a partir das variáveis mais representativas, nesse caso, o pesquisador pretende realizar apenas a redução do número de variáveis (HAIR, 2009).

A AF pode atingir o que se propõe a partir de uma perspectiva exploratória e/ou confirmatória. A Análise Fatorial Exploratória se caracteriza pela busca da estrutura do conjunto de variáveis, ou seja, com base na teoria agrupam-se aquelas que são inter-relacionadas. Consecutivamente, é possível realizar uma redução dos dados, encontrando a(s) variável (is) marcadora(s) de cada fator. Em síntese, realiza-se uma sugestão do conjunto de variáveis e seus fatores a partir de um embasamento teórico (HONGYU, 2018).

Por sua vez, a Análise Fatorial Confirmatória permite testar hipóteses, a partir de idéias preconcebidas por meio de uma forte base teórica para definir os fatores e o conjunto de variáveis que os representam (HONGYU, 2018).

2.10.2. REGRESSÃO LINEAR MÚLTIPLA

A Regressão Linear Múltipla (RLM) se caracteriza como uma técnica estatística multivariada de dependência utilizada para analisar a relação de uma única variável dependente (y) classificada como métrica e várias variáveis independentes ou preditoras (x) que também devem ser classificadas como métricas ou adequadamente transformadas (HAIR, 2009).

Os principais objetivos da RLM são prever ou explicar a variação da variável dependente analisada a partir do conjunto de valores conhecidos das variáveis independentes (FURLETTI et al., 2022). Os pesos apontam a contribuição ou influência de cada variável independente para a previsão geral em (y). E o conjunto dessas variáveis independentes forma a variável estatística de regressão (combinação linear das múltiplas variáveis x que melhor prevê a variável y) (FURLETTI et al., 2022). Portanto, a equação ou modelo de regressão pode ser apresentado da seguinte maneira (HAIR, 2009):

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_n x_n$$

Onde:

y = variável dependente;

β_0 = intercepto

β_n = peso atribuído a cada variável independente;

x_n = variável independente;

Os coeficientes betas ou coeficientes de regressão estimados representam a relação (positiva ou negativa) e a força da relação entre as variáveis explicativas e a variável dependente na variável estatística de regressão. Se o sinal do beta for positivo indica a relação positiva e se este for negativo representa a relação inversa ou negativa entre x e y . E o valor do coeficiente denota a variação no valor dependente cada vez que a variável independente varia em uma unidade (HAIR, 2009). O intercepto por sua vez, é o valor da variável dependente quando a variável estatística for igual a zero ou nula (HAIR, 2009).

Assim sendo, além de avaliar os coeficientes estimados, deve-se também verificar sua significância estatística. No artigo 3 (Resultados) são apresentadas as suposições importantes para a realização da RLM. Para tanto, é fundamental definir os objetivos da pesquisa para escolha da técnica estatística apropriada, aspectos estes apresentados no próximo subitem.

3. OBJETIVOS

3.1. OBJETIVO GERAL

Criar um modelo matemático baseado em múltiplas variáveis ambientais que seja capaz de estimar mortalidade em saúde pública.

3.2. OBJETIVOS ESPECÍFICOS

- ✓ Explorar os conceitos sobre o tema saúde/meio ambiente, os modelos de organização de indicadores e os métodos estatísticos de análise de múltiplas variáveis;
- ✓ Selecionar múltiplas variáveis de saúde ambiental e organizá-las por meio do modelo Força Motriz, Pressão, Situação, Exposição e Efeito (FPSEE);
- ✓ Determinar a estrutura latente de relações de interdependência entre todas as variáveis envolvidas no estudo através da Análise Fatorial Exploratória (AFE);
- ✓ Identificar as variáveis marcadoras em cada fator obtido pela AFE que sejam, em conjunto, capazes de reduzir todas as variáveis originais com mínima perda de informação;
- ✓ Determinar por meio da Regressão Linear Múltipla *Stepwise* o melhor modelo de combinação linear das variáveis marcadoras capaz de explicar o comportamento de óbitos em saúde pública a fim de subsidiar tomadas de decisão coletiva, integrada e territorializada em saúde;

24 RESUMO

25 Os Indicadores de Saúde Ambiental (ISA) são considerados medidas-síntese que possibilitam
26 a avaliação da saúde por meio de indicadores do ambiente. No entanto, seu levantamento,
27 seleção e organização é uma tarefa complexa devido ao seu arco de abrangência e dimensões
28 que envolvem desde aspectos ambientais quanto sociais, econômicos e demográficos. Diante
29 disso, essa pesquisa visa realizar um levantamento, seleção e organização de múltiplas
30 variáveis de saúde ambiental a fim de contribuir para o processo de vigilância em saúde e
31 realizar uma análise estatística exploratória e descritiva deste banco de dados. Primeiramente,
32 realizou-se um levantamento e seleção dos indicadores tendo como principal fundamento a
33 teoria científica. Em seguida, a organização dos indicadores se pautou no modelo Força
34 Motriz-Pressão-Situação-Exposição-Efeito (FPSEE) recomendado pela Organização Mundial
35 da Saúde. Por fim, realizou-se a análise estatística exploratória e descritiva do banco de dados
36 por meio do software IBM-SPSS *Statistics*, versão 22.0. A amostra foi composta por 853
37 observações que se referem aos municípios de Minas Gerais, região Sudeste, Brasil e os dados
38 utilizados foram obtidos em sistemas de informações virtuais de acesso público para o ano de
39 2017. De acordo com a teoria científica subjacente e os critérios para o desenvolvimento dos
40 ISA, foram selecionadas 130 variáveis para análise e agrupadas, a princípio, em 14 fatores:
41 (1) demográfico; (2) sócio-econômico; (3) assistência à saúde; (4) pressão e estado do meio
42 ambiente; (5) água e saneamento básico; (6) qualidade do ar; (7) qualidade do solo,
43 contaminantes ambientais e substâncias químicas; (8) desastres naturais; (9) habitação; (10)
44 vulnerabilidade social; (11) violência; (12) ambiente de trabalho; (13) doenças e agravos de
45 notificação e, por fim, (14) efeitos à saúde. A partir do modelo FPSEE identificou-se que
46 19,23% das variáveis foram classificadas como Força Motriz; 6,92% como Pressão; 14,61%
47 como Situação; 21,54% como Exposição e 37,70% como Efeito à saúde. A análise estatística
48 exploratória e descritiva permitiu explorar e descrever o comportamento das variáveis
49 selecionadas. E pode-se constatar a existência de 54,73% de correlações lineares entre as
50 variáveis a nível alfa de 5%. A seleção de múltiplas variáveis de saúde ambiental, assim como
51 a exploração e desenvolvimento de estudos teóricos ou metodológicos para análises de suas
52 inter-relações é fundamental para subsidiar e planejar ações de vigilância em saúde que
53 envolva fatores de riscos ambientais associados à efeitos danosos à saúde pública.

54 **Palavras-chave:** Indicadores de Saúde Ambiental. Análise Exploratória. Estatística
55 Descritiva. Vigilância em Saúde.

56

57 **ABSTRACT**

58 Environmental Health Indicators (ISA) are considered synthesis measures that allow the
59 assessment of health through environmental indicators. However, its survey, selection and
60 organization is a complex task due to its scope and dimensions that involve environmental,
61 social, economic and demographic aspects. That said, this research aims to carry out a survey,
62 selection and organization of multiple environmental health variables in order to contribute to
63 the health surveillance process and carry out an exploratory and descriptive statistical analysis
64 of this database. Firstly, a survey and selection of indicators was carried out, having as main
65 foundation the scientific theory. Then, the organization of indicators was based on the Force-
66 Pressure-Situation-Exposure-Effect (FPSEE) model recommended by the World Health
67 Organization. Finally, an exploratory and descriptive statistical analysis of the database was
68 carried out using the IBM-SPSS Statistics software, version 22.0. The sample consisted of 853
69 observations referring to the municipalities of Minas Gerais, Southeast region, Brazil and the
70 data used were obtained from publicly accessible virtual information systems for the year
71 2017. According to the underlying scientific theory and criteria for the development of ISAs,
72 130 variables were selected for analysis and grouped, at first, into 14 factors: (1) demographic;
73 (2) socio-economic; (3) health care; (4) pressure and state of the environment; (5) water and
74 sanitation; (6) air quality; (7) soil quality, environmental contaminants and chemicals; (8)
75 natural disasters; (9) housing; (10) social vulnerability; (11) violence; (12) work environment;
76 (13) notifiable diseases and conditions and, finally, (14) health effects. From the FPSEE
77 model, it was identified that 19.23% of the variables were classified as Driving Force; 6.92%
78 as Pressure; 14.61% as Status; 21.54% as Exposure and 37.70% as Health Effect. Exploratory
79 and descriptive statistical analysis made it possible to explore and describe the behavior of the
80 selected variables. The existence of 54.73% of linear correlations between the variables at an
81 alpha level of 5% can be seen. The selection of multiple environmental health variables, as
82 well as the exploration and development of theoretical or methodological studies to analyze
83 their interrelationships, is essential to support and plan health surveillance actions that involve
84 environmental risk factors associated with harmful effects on health public.

85 **Keywords:** Environmental Health Indicators. Exploratory Analysis. Descriptive statistics.
86 Health Surveillance.

87

88 **Informações de Financiamento:**

89 Esta pesquisa foi financiada pela Coordenação de Aperfeiçoamento de Pessoal de Nível
90 Superior (CAPES) [Bolsa n° 88882.461704, 2019];

91 **1. Introdução**

92 A urbanização intensa e mal planejada, o crescimento populacional e o desenvolvimento
93 econômico insustentável têm corroborado negativamente com o meio ambiente e vem
94 impactando diretamente na saúde humana.

95 A relação existente entre o meio ambiente e a saúde pública evidencia que os fatores de
96 riscos ambientais vêm ameaçando as condições de vida das populações urbanas brasileiras
97 como: poluição atmosférica, alterações climáticas, resíduos, contaminação hídrica,
98 saneamento básico inadequado, habitações insalubres, áreas de risco e vetores transmissores
99 de doenças (VEIGA, 2020) e (SALDIVA, 2019).

100 A pesquisa de Abreu *et al.* (2020) comprova a relação entre mudanças nos índices
101 pluviométricos e de umidade com riscos de propagação de doenças infecciosas e transmitidas
102 por vetores. No que tange ao uso do solo, a urbanização, o desflorestamento, as queimadas, a
103 expansão da agricultura e do uso de agrotóxicos aumentam a incidência de câncer, malária e
104 doenças respiratórias (ABREU *et al.*, 2020) e (TUCKER, 2017).

105 Eventos climáticos estão associados não apenas à secas severas e altas temperaturas, mas
106 também ao aumento de doenças como dengue, febre amarela, chikungunya, zika vírus,
107 hantavirose, leptospirose, leishmaniose e doença de chagas (NAVA *et al.*, 2017).

108 É diante destes problemas ambientais que os estudos na área da saúde ambiental vêm se
109 desenvolvendo na tentativa de minimizar os riscos e orientar medidas pertinentes. Para tanto,
110 é imprescindível a existência de instrumentos capazes de auxiliar os pesquisadores e
111 governantes na avaliação de cada cenário. Estes instrumentos, compostos por indicadores
112 inter-relacionados podem proporcionar melhorias na qualidade de vida da população na
113 medida em que são monitorados.

114 Diante disso, essa pesquisa tem como objetivo realizar um levantamento, seleção e
115 organização de múltiplas variáveis/Indicadores de Saúde Ambiental (ISA) e realizar uma
116 análise estatística exploratória e descritiva deste banco de dados. A relevância deste trabalho
117 fica evidente diante das condições ambientais insatisfatórias que se acentuam em âmbito
118 global. Ao identificar as principais variáveis ambientais relacionadas à efeitos adversos à
119 saúde humana, é possível contribuir com o processo de vigilância em saúde pública.

120 A temática é abrangente, envolvendo múltiplas variáveis e dimensões, o que requer
121 diferentes aportes teórico-metodológicos para análise. Para tanto, essa pesquisa teve como
122 fundamentos básicos a teoria científica subjacente, o Guia Básico para a Construção de
123 Indicadores em Saúde Ambiental (2011) e o documento “*Environmental Health Indicators:*

124 *Development of a Methodology for the WHO European Region*” desenvolvido pela
125 Organização Mundial da Saúde – OMS (OMS, 2001).

126 Destarte, o planejamento da matriz das variáveis foi elaborado de acordo com o modelo de
127 análise mais recente desenvolvido pela OMS, o modelo Força Motriz-Pressão-Situação-
128 Exposição-Efeito-Ação (FPSEEA) (ARANTES e PEREIRA, 2017). Esta escolha se deve à
129 sua adequação para o estabelecimento de uma base racional capaz de expor de forma
130 estruturada, a matriz de indicadores integrantes da cadeia relacional entre os determinantes
131 ambientais e efeitos à saúde, bem como à sua flexibilidade de análise. Esta análise estatística
132 exploratória e descritiva de dados teve como referência os municípios de Minas Gerais (MG),
133 fonte de dados cuja variância é suficiente para compreender os aspectos que permitem atingir
134 os objetivos propostos.

135

136 **2. Materiais e métodos**

137 **2.1. Delineamento do Estudo**

138 Esse estudo se caracteriza como observacional, exploratório e descritivo, com
139 abordagem quali-quantitativa, baseado, em primeiro momento, na construção de uma lista de
140 indicadores/variáveis de saúde ambiental e, em seguida, na análise estatística de medidas de
141 centro e dispersão dos valores das variáveis selecionadas (LAKATOS e MARCONI, 2022).

142

143 **2.2. Número de observações (municípios)**

144 O número de observações foi de 853 municípios do estado de Minas Gerais, região
145 Sudeste, Brasil. A justificativa desta escolha se deve a três aspectos principais: primeiro, pelo
146 fato da região sudeste apresentar o maior Índice de Desenvolvimento Humano registrado com
147 valor 0,766 (BRASIL, 2016), segundo, o estado de Minas Gerais proporciona o maior
148 número de observações para a análise estatística (853 municípios) (IBGE, 2017). E, por fim,
149 pela marcante heterogeneidade presente nesses municípios. Como destaca Queiroz (2001),
150 este estado é uma das regiões mais heterogêneas do país, com algumas regiões desenvolvidas
151 e outras que oferecem condições de vida insatisfatórias à sua população, sendo o retrato-
152 síntese do Brasil.

153

154 **2.3. Levantamento, Seleção e Organização dos Indicadores**

155 Primeiramente, foi realizada uma pesquisa bibliográfica para embasamento teórico e
156 levantamento de variáveis ambientais capazes de expressar a relação de dependência ou
157 interdependência com a saúde pública.

158 O levantamento bibliográfico sobre o assunto foi realizado nas seguintes bases de
 159 dados: *Medical Literature Analysis and Retrieval System Online* - MEDLINE/PUBMED,
 160 *Literatura Latino-Americana em Saúde da América Latina e Caribe* – LILACS e *Scientific*
 161 *Electronic Library Online* – SCIELO com as seguintes terminologias cadastradas nos
 162 Descritores em Ciências da Saúde (DeCS): Meio Ambiente e Saúde Pública; Indicadores
 163 Básicos de Saúde; Indicadores Ambientais.

164 A partir dessa revisão de literatura, foi observada a abrangência das variáveis e
 165 considerada a necessidade de modelos teórico-metodológicos que as analisassem
 166 simultaneamente. Para isso, foram usados os fundamentos conceituais de análise pautada na
 167 metodologia proposta pelo Guia Básico para a Construção de ISA (BRASIL, 2011).

168 Nessa etapa, também foram considerados os eixos de saúde ambiental propostos pela
 169 OMS apresentados no documento: “*Environmental Health Indicators: Development of a*
 170 *Methodology for the WHO European Region*” (OMS, 2001). Este fornece um sistema de
 171 variáveis que engloba questões ambientais de importância para a saúde pública, não apenas a
 172 nível da região européia, mas também dentro dos outros países, tais como: qualidade do ar,
 173 habitação, resíduos, solos contaminados, água para consumo, segurança alimentar,
 174 substâncias químicas e ambiente de trabalho. Ademais, foram considerados os eixos da
 175 Vigilância Ambiental em Saúde, de competência da Coordenação Geral de Vigilância em
 176 Saúde Ambiental (CGVAM) criada pelo Ministério da Saúde: água, ar, solo, desastres
 177 naturais e contaminantes.

178 Desse modo, a organização das variáveis pautou-se no modelo Força Motriz-Pressão-
 179 Situação-Exposição-Efeito definida pela OMS. Essa estrutura determina que as forças
 180 condutoras (motrizes) geram pressões capazes de alterar a situação ambiental e, induz a
 181 exposição das pessoas às condições adversas à saúde (efeitos). Vale salientar, que houve
 182 adaptações e inclusões de novas variáveis com base na teoria subjacente e em conformidade
 183 com a disponibilidade de dados e informações.

184

185 **2.4. Coleta de dados e informações**

186 A coleta dos dados foi realizada em sistemas de informações virtuais, sendo: (1)
 187 Departamento de Informática do Sistema Único de Saúde (DATASUS), (2) Instituto
 188 Brasileiro de Geografia e Estatística (IBGE), (3) Fundação João Pinheiro (FJP), (4) Sistema
 189 de Informações Ambientais Integrados a Saúde (SISAM), (5) Sistema de Informação de
 190 Vigilância da Qualidade da Água (SISAGUA), (6) Sistema Nacional de Informações sobre
 191 Saneamento (SNIS) e (7) Instituto Nacional de Pesquisas Espaciais (INPE).

192 O período de abrangência se refere ao ano 2017 devido à disponibilidade dos dados
 193 neste período em comum à todos os sistemas consultados e também devido à atualidade das
 194 informações (período inferior a cinco anos). Ademais, uma vez que as relações aqui
 195 exploradas devem ser válidas em qualquer época no tempo, mesmo que mudem as
 196 intensidades destas relações, não é relevante o período analisado, desde que se mantenha o
 197 pareamento das variáveis. Os dados foram organizados em planilha do Excel para posterior
 198 análise estatística, e todas as análises foram processadas no software IBM-SPSS *Statistics*
 199 versão 22.0.

200 201 **2.5. Análise Estatística Exploratória e Descritiva**

202 Após a seleção, organização e coleta dos dados, foi processada a análise estatística
 203 exploratória e descritiva a fim de analisar e compreender as características das variáveis
 204 selecionadas. Esta etapa considerou a análise das seguintes características conforme Triola
 205 (2013): (1) Medida de tendência central: representa um valor médio ou representativo; (2)
 206 Forma de espalhamento: medidas de quanto os valores dos dados variam entre eles; (3) Forma
 207 de distribuição: natureza ou forma de distribuição dos dados; (4) *Outliers*: presença de
 208 valores discrepantes em comparação à maioria dos outros valores amostrais; (5) Efeito do
 209 tempo: avaliação de possibilidade de alteração com o tempo.

210 211 **2.6. Aspectos Éticos**

212 Essa pesquisa não foi submetida ao comitê de ética conforme orienta a resolução
 213 466/2012 do Conselho Nacional de Saúde, pois o banco de dados não apresentou variáveis de
 214 identificação pessoal e constitui informação de domínio público. Todos os dados utilizados
 215 estavam disponíveis em sistemas de informações virtuais de acesso público e irrestrito.

216 217 **3. Resultados e Discussão**

218 **3.1. Lista de indicadores/variáveis de Saúde Ambiental, fatores e classificação FPSEE**

219 A princípio, selecionou-se 130 variáveis para análise, aderindo-se a critérios desejáveis,
 220 como: relevância social, validade, confiabilidade e obtenção de dados. Estas foram
 221 delimitadas, a princípio em 14 fatores supostos, sendo: 1) demográfico; 2) sócio-econômico;
 222 3) assistência à saúde; 4) pressão e estado do meio ambiente; 5) água e saneamento básico; 6)
 223 qualidade do ar; 7) qualidade do solo, contaminantes ambientais e substâncias químicas; 8)
 224 desastres naturais; 9) habitação; 10) vulnerabilidade social; 11) violência; 12) ambiente de
 225 trabalho; 13) doenças e agravos de notificação e 14) efeitos à saúde (Quadro 01).

Quadro 01: Descrição das variáveis de saúde ambiental, fator e classificação no modelo FPSEE.

FATOR	VAR.	DESCRIÇÃO	FPSEE
1- Demográfico	X1	População total do município;	Força Motriz
	X2	População urbana do município;	Força Motriz
	X3	Densidade populacional (hab/km ²);	Força Motriz
	X4	Taxa de urbanização;	Força Motriz
	X5	Esperança de vida ao nascer;	Força Motriz
	X6	Índice de envelhecimento;	Força Motriz
	X7	Mortalidade proporcional da população idosa (%);	Força Motriz
	X8	Mortalidade até 5 anos de idade (%);	Força Motriz
	X9	Taxa de mortalidade infantil;	Força Motriz
	X10	Taxa bruta de mortalidade;	Força Motriz
2- Sócio-Econômico	X11	Taxa de escolarização líquida (ensino médio) (%);	Força Motriz
	X12	Número de famílias com renda per capita até 1/2 salário mínimo;	Força Motriz
	X13	Percentual de pobres;	Força Motriz
	X14	Percentual da pop. pobre no Cadastro Único em relação a pop. total;	Força Motriz
	X15	Percentual de pessoas em idade produtiva e sem ocupação do Cadastro Único;	Força Motriz
	X16	Taxa de emprego no setor formal (%);	Força Motriz
	X17	Razão de dependência;	Força Motriz
3- Assistência à Saúde	X18	Gasto per capita com atividades de saúde;	Força Motriz
	X19	Número geral de estabelecimentos de saúde por nível de atenção;	Força Motriz
	X20	Número de profissionais da saúde;	Força Motriz
	X21	Coberturas vacinais;	Força Motriz
	X22	Proporção da população atendida pela Estratégia de Saúde da Família;	Força Motriz
	X23	Percentual de pessoas cobertas por planos de saúde suplementar;	Força Motriz
	X24	Percentual de nascidos vivos com pelo menos 7 consultas de pré-natal;	Força Motriz
	X25	Proporção de nascidos vivos com baixo peso (%);	Força Motriz
	X26	Percentual de óbitos por causas mal definidas;	Efeito
X27	Percentual de óbitos por causas mal definidas sem assistência médica;	Efeito	
4- Pressão Ambiental	X28	Número de Unidades Locais - Indústria, Const, Comércio e Serviços;	Pressão
	X29	Cobertura por agropecuária (%);	Pressão
	X30	Cobertura por plantação de cana-de-açúcar (%);	Pressão
	X31	Densidade de veículos (unidade/km ²);	Pressão
	X32	Densidade de rebanho bovino (unidade/km ²);	Pressão
	X33	Cobertura por infraestrutura urbana (%);	Pressão
	X34	Cobertura vegetal natural (%);	Pressão
	X35	Cobertura por floresta plantada (%);	Pressão
	X36	Participação nos focos de calor do estado (%);	Pressão
5- Água para consumo humano e Saneamento Básico	X37	Percentual de água tratada;	Situação
	X38	Percentual de amostras fora dos padrões microbiológicos;	Situação
	X39	Percentual de amostras fora dos padrões organolépticos;	Situação
	X40	Índice de Qualidade da Água – Média Anual;	Situação
	X41	Qualidade da água: Classe de contaminação por tóxicos;	Situação
	X42	Percentual de esgoto tratado;	Situação
	X43	Taxa de mortalidade por cólera;	Efeito
	X44	Taxa de mortalidade por diarreia e gastroenterite;	Efeito
	X45	Taxa de mortalidade por outras doenças infecciosas intestinais;	Efeito
6- Qualidade do Ar	X46	Concentração de CO (ppb);	Situação
	X47	Concentração de O ₃ (ppb);	Situação
	X48	Concentração de NO ₂ (ppb);	Situação
	X49	Concentração de SO ₂ (µg/m ³);	Situação

	X50	Concentração de PM _{2.5} (µg/m ³);	Situação
	X51	Concentração de PM ₁₀ (µg/m ³);	Situação
	X52	Temperatura do ar (°C);	Situação
	X53	Umidade relativa do ar (%);	Situação
	X54	Precipitação acumulada (mm);	Situação
	X55	Taxa de mortalidade por câncer de pulmão (por 100 mil habitantes);	Efeito
	X56	Número de óbitos por influenza (gripe);	Efeito
	X57	Número de óbitos por pneumonia;	Efeito
	X58	Número de óbitos por Doenças Crônicas das Vias Aéreas Inferiores (bronquite, enfisema, asma, estado de mal asmático, bronquectasia);	Efeito
	X59	Número de óbitos por outras Doenças do Aparelho Respiratório (nasofaringite, sinusite, faringite, amigdalite aguda, laringite);	Efeito
	X60	Número de óbitos por neoplasias da traquéia, brônquios e pulmões;	Efeito
	X61	Taxa de mortalidade por doenças respiratórias em crianças < 1 ano;	Efeito
	X62	Número de óbitos por Doenças Isquêmicas do Coração;	Efeito
	X63	Número de óbitos por Infarto Agudo do Miocárdio;	Efeito
7- Qualidade do solo, Contaminantes Ambientais e Substâncias Químicas	X64	Os resíduos possuem destino final em lixões a céu aberto?	Situação
	X65	O município apresentou registro de áreas contaminadas?	Situação
	X66	O município apresentou registro de acidentes ambientais?	Situação
	X67	Número de áreas cadastradas como contaminadas e reabilitadas para municípios com 4 ou mais áreas;	Situação
	X68	Número de óbitos por Intoxicação Exógena;	Efeito
	X69	Número de óbitos de intoxicação exógena por agrotóxico;	Efeito
	X70	Número de óbitos de intoxicação exógena por produto químico;	Efeito
	X71	Número de óbitos de intoxicação exógena por alimento e bebida;	Efeito
X72	Número de óbitos por exposição a substâncias nocivas;	Efeito	
8- Desastres Naturais e Fatores da Natureza	X73	Índice de vulnerabilidade a mudanças climáticas do território; (grau de exposição, sensibilidade e capacidade de adaptação);	Exposição
	X74	Parcela de Domicílios em Situação de Risco de Inundação;	Exposição
	X75	Número de óbitos por exposição à corrente elétrica, radiação, temperatura e pressão extrema ambiental;	Efeito
	X76	Número de óbitos por exposição às forças da natureza (calor excessivo, frio excessivo, raio, desabamento de terra, inundação);	Efeito
	X77	Índice de óbitos em decorrência de eventos hidrológicos;	Efeito
	X78	Número de óbitos por Doenças Hipertensivas (DH);	Efeito
	X79	Número de óbitos por neoplasia maligna da pele;	Efeito
9- Habitação	X80	Percentual da população urbana em domicílios com abastecimento de água;	Exposição
	X81	Percentual da população urbana em domicílios com esgotamento sanitário;	Exposição
	X82	Percentual da população urbana atendida por serviços de coleta de resíduos;	Exposição
	X83	Existência de coleta seletiva;	Exposição
	X84	Percentual da população em domicílios com banheiro e água encanada;	Exposição
	X85	Déficit habitacional;	Exposição
	X86	Número de domicílios precários;	Exposição
	X87	Número de domicílios em situação de coabitação familiar;	Exposição
	X88	Número de domicílios c/ pelo menos um tipo de serv. básico inadequado;	Exposição
	X89	Número de domicílios alugados com adensamento excessivo;	Exposição
	X90	Número de domicílios urbanos próprios com adensamento excessivo;	Exposição

10- Vulnerabilidade Social	X91	Percentual de pessoas em situação de vulnerabilidade pelas condições de saneamento básico;	Exposição
	X92	Percentual de pessoas inscritas no Cadastro Único sem abastecimento de água, esgotamento sanitário e coleta de lixo adequados;	Exposição
	X93	Percentual de pessoas inscritas no Cadastro Único sem abastecimento de água;	Exposição
	X94	Percentual de pessoas inscritas no Cadastro Único sem esgotamento sanitário;	Exposição
	X95	Percentual de pessoas inscritas no Cadastro Único sem coleta de lixo adequada;	Exposição
	X96	Percentual de pessoas em domicílios com abastecimento de água e esgotamento sanitário inadequados;	Exposição
	X97	Percentual de pessoas em domicílios sem energia elétrica;	Exposição
	X98	Percentual de pessoas em domicílios com paredes que não sejam de alvenaria;	Exposição
	X99	Número de domicílios urbanos sem banheiro;	Exposição
	X100	Percentual de pessoas em domicílios vulneráveis à pobreza e que gastam mais de uma hora até o trabalho;	Exposição
	X101	Percentual de meninas de 10 a 14 anos que tiveram filhos;	Exposição
	X102	Percentual de adolescentes de 15 a 17 anos que tiveram filhos;	Exposição
11- Violência	X103	Taxa de crimes violentos (por 100 mil habitantes);	Exposição
	X104	Taxa de crimes violentos contra a pessoa (por 100 mil habitantes);	Exposição
	X105	Taxa de ocorrências de homicídios dolosos (por 100 mil habitantes);	Exposição
	X106	Taxa de mortalidade por homicídio (por 100 mil habitantes);	Efeito
	X107	Taxa de mortalidade por agressão (por 100 mil habitantes);	Efeito
	X108	Taxa de mortalidade por suicídio;	Efeito
12- Ambiente de Trabalho	X109	Número de óbitos por doenças relacionadas ao trabalho;	Efeito
	X110	Número de óbitos de pneumoconiose (de mineiros de carvão, devido amianto ou outras fibras minerais, à poeira de sílica, outras poeiras);	Efeito
	X111	Número de óbitos de intoxicações exógenas por exposição ao trabalho;	Efeito
	X112	Número de óbitos por acidente de trabalho grave;	Efeito
	X113	Número de óbitos de câncer relacionado ao trabalho;	Efeito
13- Doenças e Agravos de Notificação (Vetores, Hospedeiros, outros)	X114	Número de óbitos notificados de dengue;	Efeito
	X115	Número de óbitos notificados de febre amarela;	Efeito
	X116	Número de óbitos notificados de Leishmaniose;	Efeito
	X117	Número de óbitos notificados de Esquistossomose;	Efeito
	X118	Número de óbitos notificados por doença de Chagas;	Efeito
	X119	Número de óbitos notificados de Leptospirose;	Efeito
	X120	Número de óbitos notificados de hepatites;	Efeito
14- Efeitos à Saúde	X121	Número de óbitos por algumas doenças infecciosas e parasitárias (cólera, tuberculose, difteria, malária, toxoplasmose);	Efeito
	X122	Número de óbitos por neoplasias;	Efeito
	X123	Número de óbitos por doenças endócrinas nutricionais e metabólicas;	Efeito
	X124	Número de óbitos por transtornos mentais e comportamentais;	Efeito
	X125	Número de óbitos por doenças do sistema nervoso;	Efeito
	X126	Número de óbitos por doenças do aparelho circulatório;	Efeito
	X127	Número de óbitos por doenças do aparelho respiratório;	Efeito
	X128	Número de óbitos por doenças da pele e do tecido subcutâneo;	Efeito
	X129	Número de óbitos por malformações congênitas, deformidades e anomalias cromossômicas;	Efeito
	X130	Número de óbitos por causas externas;	Efeito

227 O modelo Força Motriz-Pressão-Situação-Exposição-Efeito-Ação (FPSEEA) foi
228 desenvolvido pela OMS, juntamente com o Programa das Nações Unidas para o Meio
229 Ambiente (PNUMA) e a Agência de Proteção Ambiental dos Estados Unidos (USEPA),
230 incluindo conceitos de forças motrizes, exposição ambiental e efeitos na saúde (ARANTES e
231 PEREIRA, 2017).

232 As forças motrizes representam a escala macro e que influenciam os vários processos
233 ambientais que vão interferir nas condições de vida e bem-estar da população. A pressão
234 corresponde sobre as ações antrópicas no ambiente como os processos de produção e
235 consumo de bens e serviços. Como consequências, essas pressões,
236 alteram/modificam/agravam a situação do ambiente, resultando em exposições e efeitos
237 negativos para a saúde dos indivíduos.

238 O destaque deste modelo é reconhecer que alterações na situação ambiental causada pelas
239 forças condutoras (motrizes) e pressões atingem diferentes meios como o solo, a água e o ar,
240 resultando em exposições humanas que poderão ter efeito direto ou indireto na saúde em
241 curto, médio ou longo prazo (ARANTES e PEREIRA, 2017).

242 Ressalta-se que a última classificação (ação) que visa determinar medidas efetivas será
243 estabelecida em trabalhos futuros após identificar as variáveis críticas à saúde. A partir da
244 identificação dos indicadores que mais afetam a saúde humana, podem-se definir ações no
245 campo da formulação e implantação de políticas públicas em saúde.

246 A partir dos resultados demonstrados no quadro 01, observa-se que os fatores
247 demográfico, socioeconômico e assistência à saúde representam a escala macro, portanto, as
248 variáveis estão organizadas hegemonicamente em Força Motriz. Em contrapartida, os
249 constructos pressão do meio ambiente, água para consumo humano, qualidade do ar e do solo
250 representam as variáveis de pressão e situação ambiental, incluindo também os efeitos na
251 saúde. Por sua vez, os eixos desastres naturais, habitação, vulnerabilidade social, violência e
252 ambiente de trabalho agrupam indicadores que representam a exposição das pessoas a essas
253 adversidades do ambiente. Por fim, a última dimensão reúne as variáveis de efeitos à saúde,
254 no caso desse estudo, o desfecho considerado foi o número de óbitos. Considerando todos os
255 constructos analisados constata-se que 19,23% das variáveis foram classificadas como Força
256 Motriz; 6,92% como pressão; 14,61% como situação; 21,54% como exposição e 37,70%
257 como efeito.

258 É importante salientar a relevância do modelo FPSEEA ser utilizado como ferramenta
259 auxiliar com outras análises, como estatísticas multivariadas para avaliar toda a complexidade
260 existente nas inter-relações entre esses indicadores.

261 **3.2. Análise Estatística Exploratória e Descritiva**

262 Primeiramente, ao explorar o banco de dados, observou-se o tipo e a natureza das
263 variáveis, sendo em sua maioria métricas (Tabela 01).

264

265 Tabela 01: Número de fatores, n amostral, tipo e natureza das variáveis selecionadas

FATOR	N AMOSTRAL	TIPO E NATUREZA (var.)
1. Demográfico	n=10	Métricas de razão;
2. Sócio-econômico	n=07	Métricas de razão;
3. Assistência à saúde	n=10	Métricas de razão;
4. Pressão e estado do meio ambiente	n=09	Métricas de razão;
5. Água para consumo humano e saneamento básico	n=09	Métricas de razão e duas não-métricas ordinais;
6. Qualidade do ar	n=18	Métricas de razão;
7. Qualidade solo, cont. ambientais e subst. químicas	n=09	Métricas de razão e três binárias;
8. Desastres naturais e fatores da natureza	n=07	Métricas de razão;
9. Habitação	n=11	Métricas de razão e uma binária;
10. Vulnerabilidade social	n=12	Métricas de razão;
11. Violência;	n=06	Métricas de razão;
12. Ambiente de trabalho	n=05	Métricas de razão;
13. Doenças e agravos de notificação	n=07	Métricas de razão;
14. Efeitos à saúde	n=10	Métricas de razão;

266 Fonte: LOPES, A. F. A. 2022.

267

268 Portanto, foi identificado que, do total amostral n=130 variáveis, cento e vinte e três foram
269 métricas e de escala de razão, apenas duas foram classificadas como não métricas ordinais
270 (VAR 40, VAR 41) e quatro foram convertidas em métricas (binárias) (VAR 64, VAR 65,
271 VAR 66 e VAR 83).

272 As variáveis não métricas são qualitativas e apresentam escala nominal e ordinal (HAIR,
273 2009). As escalas nominais são conhecidas como categóricas e designam números para
274 identificar indivíduos ou objetos (HAIR, 2009). Nas escalas ordinais, as variáveis podem ser
275 ordenadas, permitindo a comparação entre indivíduos ou objetos (HAIR, 2009). Por sua vez,
276 as variáveis métricas são quantitativas e apresentam duas escalas de medida, intervalares e de
277 razão. A única diferença entre as duas é que as escalas intervalares possuem um ponto zero
278 arbitrário (valores negativos) e as de razão têm um ponto zero absoluto (ausência de valores
279 negativos) (HAIR, 2009).

280 Ressalta-se que nessa etapa de análise das variáveis ponderaram-se os requisitos
281 necessários para que entrassem em um modelo de análise fatorial posterior, considerando o

282 tipo de variável e sua natureza (métricas e de razão). Além disso, considerou-se minimamente
 283 05 variáveis por fator e um número de 853 observações, atendendo assim os requisitos
 284 mínimos de observações por variável (5x1) como determina Hair (2009).

285 A análise estatística descritiva foi realizada no software IBM-SPSS *Statistics* versão 22.0,
 286 com a descrição de cada uma das variáveis correspondentes aos seus respectivos constructos.
 287 No entanto, devido ao arco de abrangência dessas múltiplas variáveis selecionadas, optamos,
 288 neste estudo, por apresentar a descrição das tabelas das cinco principais variáveis marcadoras
 289 identificadas após o processamento por Análise Fatorial Exploratória (AFE), sendo: número
 290 de óbitos por neoplasia (VAR 122) (Tabela 02); Percentual de pessoas inscritas no Cadastro
 291 Único sem abastecimento de água adequado (VAR 93); Concentração de NO₂ (VAR 48);
 292 Taxa de mortalidade por homicídio (VAR 106) e cobertura vegetal natural (VAR 34). Estas
 293 sintetizam as informações das 130 variáveis originais, ademais foram reduzidas para cinco
 294 constructos que apresentaram mais da metade do percentual de variância explicada (59,78%).

295 A estatística descritiva pode ser dividida em medidas de tendência central, forma de
 296 espalhamento, forma de distribuição, *outliers* (valores atípicos) e efeito do tempo (TRIOLA,
 297 2013). Para a tendência central foram calculadas a média e a mediana; para a forma de
 298 espalhamento, observou-se a variância dos dados e amplitude interquartil; e para a forma de
 299 distribuição foram observados os valores de assimetria e curtose.

300

301 Tabela 02: Estatística Descritiva (variável número de óbitos por neoplasia)

DESCRITIVOS		Estatística	Erro	
var_122	Média	27.3955	4.53046	
	95% Intervalo de Confiança para Média	Limite inferior	18.5033	
		Limite superior	36.2878	
	5% da média aparada	14.5150		
	Mediana	9.0000		
	Variância	17384.733		
	Erro Desvio	131.85118		
	Mínimo	1.00		
	Máximo	3482.00		
	Intervalo	3481.00		
	Amplitude interquartil	15.00		
	Assimetria	21.958	0.084	
Curtose	561.086	0.168		

302 Fonte: LOPES, A. F. A. 2022.

303

304 A variável 122 ‘número de óbitos por neoplasia’ refere-se ao constructo efeitos à saúde e
 305 os dados apresentaram média de 27.3955 com 95% de certeza do Intervalo de Confiança (IC)

306 apresentar o verdadeiro valor no intervalo IC=[18.5033-36.2878], sendo o valor central
 307 (mediana= 9.0000). A variância é considerada uma boa medida de variabilidade, pois
 308 considera todo o conjunto de dados, sendo a medida que expressa um desvio quadrático
 309 médio dos dados em relação à média (TRIOLA, 2013).

310 O valor máximo do conjunto de dados foi igual a 3482.00 e o valor mínimo foi 1.000. A
 311 amplitude interquartil é a diferença entre o terceiro e o primeiro quartil e abrange 50% dos
 312 dados, sendo útil para detectar valores discrepantes (RODRIGUES, *et.al.*, 2017). Para essa
 313 variável a amplitude interquartil da variável 122 equivale a 15.00.

314 A assimetria e curtose indicam a forma da distribuição dos dados, que também podem ser
 315 verificados por meio dos testes de normalidade de *Kolmogorov-Smirnova* e *Shapiro-Wilk*.
 316 Uma distribuição é considerada como simétrica quando os valores da média, mediana e moda
 317 são iguais; assimétrica negativa quando se $média \leq mediana \leq moda$ e assimétrica positiva se
 318 $moda \leq mediana \leq média$ (RODRIGUES, *et.al.*, 2017). Os valores de assimetria e curtose
 319 próximos à zero indicam a normalidade dos dados (MEDRI, 2011). Nesse caso, como a média
 320 apresentou valores acima da mediana, têm-se uma distribuição assimétrica positiva ou à
 321 direita. A próxima tabela apresenta a descrição da variável 93 (tabela 03).

322

323 Tabela 03: Estatística Descritiva (variável Cadastro Único sem abastecimento de água
 324 adequado)

DESCRITIVOS		Estatística	Erro	
var_93	Média	26.0098	0.59808	
	95% Intervalo de Confiança para Média	Limite inferior	24.8359	
		Limite superior	27.1837	
	5% da média aparada	25.0544		
	Mediana	21.7950		
	Variância	304.041		
	Erro Desvio	17.43677		
	Mínimo	0.19		
	Máximo	76.10		
	Intervalo	75.91		
	Amplitude interquartil	25.44		
	Assimetria	0.723	0.084	
Curtose	-0.273	0.168		

325 Fonte: LOPES, A. F. A. 2022.

326

327 A variável 93 ‘Cadastro Único sem abastecimento de água adequado’ refere-se ao fator
 328 vulnerabilidade social e os dados apresentaram média de 26.0098 com 95% de certeza do
 329 Intervalo de Confiança (IC) conter o verdadeiro valor IC= [24.8359-27.1837] e mediana igual
 330 a 21.7950. O valor máximo do conjunto de dados foi igual a 76.10 e o valor mínimo foi 0.19.
 331 Para essa variável, nota-se que os dados se aproximam de uma distribuição normal com

332 valores de curtose e assimetria próximos a zero. Ademais, o valor da mediana se aproxima ao
 333 valor da média. A seguir, têm-se a descrição da variável 48 (tabela 04).

334

335 Tabela 04: Estatística Descritiva (variável concentração de Dióxido de Nitrogênio - NO₂)

DESCRITIVOS		Estatística	Erro	
var_48	Média	1.8566	0.03115	
	95% Intervalo de Confiança para Média	Limite inferior	1.7954	
		Limite superior	1.9177	
	5% da média aparada	1.8065		
	Mediana	1.7600		
	Variância	0.827		
	Erro Desvio	0.90924		
	Mínimo	0.37		
	Máximo	4.85		
	Intervalo	4.48		
	Amplitude interquartil	1.25		
	Assimetria	0.690	0.084	
	Curtose	0.296	0.167	

336 Fonte: LOPES, A. F. A. 2022.

337

338 A variável 48 ‘concentração de Dióxido de Nitrogênio (NO₂)’ pertence ao fator qualidade
 339 do ar e os dados apresentaram média de 1.8566 com 95% de certeza do Intervalo de
 340 Confiança (IC) conter o verdadeiro valor IC=[1.7954-1.9177] e mediana 1.7600. O valor
 341 máximo do conjunto de dados foi igual a 4,85 e o valor mínimo foi 0,37. Como os valores de
 342 assimetria e curtose estão próximos de zero e o valor da mediana também se aproxima ao
 343 valor da média, podemos considerar que há uma distribuição normal dos dados para essa
 344 variável. A seguir, têm-se a descrição da variável 106 (tabela 05).

345

346 Tabela 05: Estatística Descritiva (variável Taxa de mortalidade por homicídio)

DESCRITIVOS		Estatística	Erro	
var_106	Média	16.7649	0.68027	
	95% Intervalo de Confiança para Média	Limite inferior	15.4297	
		Limite superior	18.1001	
	5% da média aparada	14.4059		
	Mediana	12.3450		
	Variância	394.275		
	Erro Desvio	19.85637		
	Mínimo	0.00		
	Máximo	120.90		
	Intervalo	120.90		
	Amplitude interquartil	25.32		
	Assimetria	1.823	0.084	
	Curtose	4.505	0.167	

347 Fonte: LOPES, A. F. A. 2022.

348

349 A variável 106 ‘Taxa de mortalidade por homicídio’ pertence ao fator violência e os dados
 350 apresentaram média de 16.7649 com 95% de confiança de o intervalo conter o verdadeiro
 351 valor IC= [15.4297-18.1001] e mediana equivalente a 12.3450. O valor máximo do conjunto
 352 de dados foi igual a 120.90 e o valor mínimo foi 0.00. Nesse caso, nota-se que os dados
 353 apresentam uma pequena assimetria à direita com valor da média (16.7649) um pouco acima
 354 ao valor da mediana (12.3450). Por fim, apresentamos a tabela descritiva da variável 34
 355 (tabela 06).

356

357 Tabela 06: Estatística Descritiva (variável cobertura vegetal natural)

DESCRITIVOS			Estatística	Erro
var_34	Média		27.3735	0.60489
	95% Intervalo de Confiança para Média	Limite inferior	26.1863	
		Limite superior	28.5608	
	5% da média aparada		26.2283	
	Mediana		23.2700	
	Variância		311.379	
	Erro Desvio		17.64592	
	Mínimo		0.29	
	Máximo		95.59	
	Intervalo		95.30	
	Amplitude interquartil		24.30	
	Assimetria		0.938	0.084
	Curtose		0.487	0.167

358

Fonte: LOPES, A. F. A. 2022.

359

360 A variável 34 ‘cobertura vegetal natural’ pertence ao fator Pressão Ambiental e os dados
 361 apresentaram média de 27.3735 com 95% de confiança de o intervalo conter este valor
 362 IC=[26.1863-28.5608] e mediana 23.2700. Podemos observar que os valores de assimetria e
 363 curtose estão próximos de zero indicando uma normalidade da distribuição dos dados.

364

Em geral, a diversidade da variância do conjunto de dados observada a priori (130
 365 variáveis e 853 observações), pode ser explicada pelo fato dos resultados das variáveis
 366 referentes às observações se diferirem entre si. Enquanto algumas variáveis apresentam
 367 valores elevados em determinados municípios, outras já apresentam valores muito baixos, e
 368 isso pode ser justificado pelas diferenças demográficas e socioeconômicas entre eles.

369

Outro aspecto se refere ao número de casos omissos identificados, notavelmente, em 04
 370 fatores por meio da estatística descritiva (resumo de processamento do caso) (tabela 07)
 371 sendo: qualidade do solo, contaminantes ambientais e substâncias químicas (VAR 67 a VAR
 372 72 com média de 96,63% de casos omissos); desastres naturais e fatores da natureza (VAR

373 75, 76 e VAR 79 com média de 94,3%); fator ambiente de trabalho (VAR 109 a VAR 113
 374 com média de 92,92%); e doenças e agravos de notificação (VAR 114 a VAR 120 com média
 375 de 91,38%).

376

377 Tabela 07: Resumo de processamento do caso, fator (qualidade do solo, contaminantes
 378 ambientais e substâncias químicas: var 64 – var 72, consultar quadro 01).

	Casos					
	Válido		Omisso		Total	
	N	Porcentagem	N	Porcentagem	N	Porcentagem
var_64	851	99.9%	1	0.1%	852	100.0%
var_65	852	100.0%	0	0.0%	852	100.0%
var_66	852	100.0%	0	0.0%	852	100.0%
var_67	28	3.3%	824	96.7%	852	100.0%
var_68	91	10.7%	761	89.3%	852	100.0%
var_69	16	1.9%	836	98.1%	852	100.0%
var_70	7	0.8%	845	99.2%	852	100.0%
var_71	3	0.4%	849	99.6%	852	100.0%
var_72	26	3.1%	826	96.9%	852	100.0%

379 Fonte: LOPES, A. F. A. 2022.

380

381 Ressalta-se que essa ausência de dados nestes constructos se deve à indisponibilidade nos
 382 sistemas de informações e/ou devido à própria inexistência de casos que não foram lançados
 383 nos bancos virtuais.

384 Com relação à forma de distribuição, a distribuição não homogênea dos dados não
 385 influencia os resultados quando existe linearidade e homocedasticidade e isso pode ser
 386 comprovado pelo número substancial de correlações lineares significantes entre as variáveis a
 387 nível alfa de 5% (54,73%) encontradas pela matriz de correlações de *Spearman*.

388 A presença de *outliers* também pode alterar a forma de distribuição dos dados. Estes
 389 foram identificados como sendo todos os pontos maiores que 1,5 o desvio interquartil
 390 acima do 3º quartil ou abaixo do 1º quartil.

391 No que diz respeito ao efeito do tempo, destaca-se que os dados das variáveis podem ser
 392 coletados no ano de interesse definido pelo pesquisador e, apesar de sofrerem efeito do tempo
 393 e as intensidades das relações e inter-relações entre as variáveis oscilarem, a estrutura do
 394 constructo deve ser a mesma independente do tempo e ser válida para qualquer época.

395 4. Conclusão

396 A temática saúde e meio ambiente apresenta um amplo campo de estudo, eminentemente
397 interdisciplinar que envolve a união de profissionais de diversas formações acadêmicas e
398 técnicas, tanto das áreas biomédicas quanto das ciências da natureza e das ciências exatas. A
399 união dessas áreas para propor ISA assim como sua análise, organização e mensuração é
400 fundamental para a vigilância em saúde ambiental e a implementação eficaz de ações em
401 saúde.

402 A partir do embasamento teórico realizado nesse estudo, foi possível propor uma lista de
403 indicadores de saúde ambiental a fim de contribuir para seu monitoramento, sendo possível a
404 sua organização no modelo FPSEE recomendado pela OMS e o agrupamento em fatores ou
405 constructos. Ademais, foi possível realizar o levantamento dos dados nos sistemas de
406 informações, apesar de algumas limitações encontradas como desvinculação entre os bancos
407 de dados e ausência de informações.

408 A análise estatística permitiu explorar e descrever o comportamento das variáveis
409 selecionadas (*a priori*). Foi possível analisar a sua natureza, assim como os requisitos para
410 escolha daquelas que poderiam entrar em um modelo de Análise Fatorial Exploratória (AFE)
411 a fim de identificar a relação e o constructo latente entre elas.

412 A partir da análise estatística descritiva pode-se concluir que do ponto de vista estatístico,
413 a variância dos dados e a distribuição não-homogênea não altera os resultados quando se tem
414 um número substancial de correlações lineares entre as variáveis analisadas.

415 Ressalta-se que a exploração do banco de dados é fundamental para identificar a presença
416 de *outliers*, variáveis com correlações insignificantes e identificação de casos omissos
417 elevados, a fim de garantir a certeza dos resultados encontrados.

418 Em suma, a seleção das variáveis propostas nesse estudo resultou em um número
419 substancial de correlações entre elas, no entanto, essa seleção pode ser modificada (com
420 inclusão ou exclusão) considerando a teoria científica, as dimensões de cada pesquisa, as
421 especificidades de cada unidade de análise e a disponibilidade de dados e informações.
422 Portanto, a seleção de múltiplas variáveis de saúde ambiental, assim como a exploração e
423 desenvolvimento de aportes teórico-metodológicos para análises de suas inter-relações é
424 fundamental para subsidiar e planejar ações de vigilância em saúde que envolva fatores de
425 riscos ambientais que provocam efeitos deletérios na saúde pública.

426 **5. Agradecimentos**

427 À Universidade Federal de Uberlândia por meio do Programa de Pós-Graduação em Ciências da
428 Saúde PPCSA/UFU e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.
429

430 **6. Referências**

- 431 ABREU, A. M. et al. A interface entre saúde, mudanças climáticas e uso do solo no Brasil: uma análise da
432 evolução da produção científica internacional entre 1990 e 2019. **Saúde e Sociedade**, v. 29, n. 2, 2020. DOI:
433 <https://doi.org/10.1590/S0104-12902020180866>. Disponível em: <https://www.scielo.br/j/sausoc/>. Acesso em: 18
434 de abril de 2022.
- 435 ARANTES, K. M.; PEREIRA, B. B. Levantamento, análise e seleção de indicadores ambientais e
436 socioeconômicos como subsídio para o fortalecimento das estratégias de controle da dengue no município de
437 Uberlândia- MG. **JHBS- Journal of Health and Biological Sciences**, v.5, n. 1, 2017. DOI:
438 <http://dx.doi.org/10.12662/2317-3076jhbs.v5i1.1104.p86-94.2017>. Disponível em:
439 <<https://periodicos.unichristus.edu.br/jhbs/article>>. Acesso em 16 de maio de 2022.
- 440 BRASIL. **Desenvolvimento humano nas macrorregiões brasileiras**: 2016. – Brasília: PNUD : IPEA : FJP,
441 2016. 55 p.: il., gráfs, mapas color. ISBN: 978-85-88201-31-6.
- 442 _____. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância em Saúde
443 Ambiental e Saúde do Trabalhador. **Saúde ambiental**: guia básico para construção de indicadores. Brasília:
444 Ministério da Saúde, 2011.
- 445 HAIR, J. F. **Análise multivariada de dados** [recurso eletrônico] / tradução Adonai Schlup Sant’Anna. – 6. ed. –
446 Dados eletrônicos. – Porto Alegre: Bookman, 2009.
- 447 INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA (IBGE). **Panorama Brasil**, v4.6.25.1, 2017.
448 Disponível em: <https://cidades.ibge.gov.br/brasil/panorama>. Acesso em: 02 de maio de 2022.
- 449 LAKATOS, M. E. MARCONI, M. A. **Metodologia Científica**. 8 ed. 392p. São Paulo: Atlas, 2022.
- 450 MEDRI, W. **Análise Exploratória de dados**. Curso de Especialização Lato Sensu . Centro de Ciências Exatas –
451 CCE, Departamento de Estatística. Universidade Estadual de Londrina, Londrina/PR, 2011.
- 452 NAVA, A. et al. *The impact of global environmental changes on infectious disease emergence with a focus on*
453 *risks for Brazil*. **ILAR Journal**, Londres, p. 1-8, 2017. DOI: 10.1093/ilar/ilx034. Disponível em:
454 <https://pubmed.ncbi.gov/>. Acesso em: 05 de abril de 2022.
- 455 ORGANIZAÇÃO MUNDIAL DE SAÚDE (OMS). **Environmental Health Indicators: Development of a**
456 **Methodology for the WHO European Region**. Dinamarca, 2001.
- 457 QUEIROZ, B.L. **Diferenciais de salários nas regiões mineiras**. Belo Horizonte, 2001. 191f. Dissertação
458 (mestrado em demografia) – Faculdade de Ciências Econômicas, Universidade Federal de Minas Gerais, Belo
459 Horizonte, 2001.
- 460 RODRIGUES *et.al*. *Importance of using basic statistics adequately in clinical research Study performed at*
461 *Universidade Federal de Alagoas*. **Revista Brasileira de Anestesiologia** [online]. 2017, v. 67, n. 6 , pp. 619-
462 625. DOI: <https://doi.org/10.1016/j.bjane.2017.01.011>. Disponível em:
463 <<https://www.scielo.br/j/rba/a/N5PgBCrzhDkfRbX8QXsctHx/?lang=pt#>>. Acessoem: 02 de maio de 2022.
- 464 SALDIVA, P; DAVIDOVICH, L. **Air Pollution and Health: A Science Policy Initiative of the National**
465 **Academie, 2019**. *Annals of Global Health*, 85(1), p.143. DOI: <http://doi.org/10.5334/aogh.2670>. Disponível em:
466 <https://annalsofglobalhealth.org/articles/>. Acesso em: 05 de abril de 2022.
- 467 TRIOLA, M.F. **Introdução à Estatística**: atualização da tecnologia / tradução e revisão técnica Ana Maria L. de
468 Farias, Vera Regina L.F. e Flores. Rio de Janeiro: LTC, 2013.
- 469 TUCKER, J. M. et al. *Does deforestation promote or inhibit malaria transmission in the Amazon? A systematic*
470 *literature review and critical appraisal of current evidence*. **Philosophical Transactions of the Royal Society of**
471 **London: Biological Sciences**, Londres, v. 372, n. 1722, p. 20160125, 2017. DOI: 10.1098/rstb.2016.0125.
472 Disponível em: <https://pubmed.ncbi.gov/>. Acesso em: 01 jun. 2022.
- 473 VEIGA, J. E. **Saúde e sustentabilidade**. *Estudos Avançados*, 2020, v. 34, n. 99, pp. 303-310, ISSN 1806-9592.
474 DOI: <https://doi.org/10.1590/s0103-4014.2020.3499.018>. Disponível em:
475 <<https://www.scielo.br/j/ea/a/xvdzDcv5mdrcZ7VMCzL6gf/?lang=pt>>. Acesso em: 16 de maio de 2022.

Artigo 2. Artigo intitulado: “Meio Ambiente e Saúde Humana: O uso da Análise Fatorial Exploratória – AFE como estatística de análise de múltiplas variáveis”, formatado para submissão à revista *Environmental Research* (Impact Factor=8.431).

476 **Meio Ambiente e Saúde Humana: O uso da Análise Fatorial Exploratória – AFE como**
477 **estatística de análise de múltiplas variáveis²**
478

479

480

481

482

Anáisa Filmiano Andrade Lopes

483

Universidade Federal de Uberlândia, Programa de Pós-Graduação em Ciências da Saúde,
484 Faculdade de Medicina, Av. Pará - 1358 - Bloco 2U - Bairro Umuarama, Uberlândia, Minas

485

Gerais, Brasil

486

ana_isaandrade@ufu.br

487

488

Yara Cristina de Paiva Maia

489

Universidade Federal de Uberlândia, Programa de Pós-Graduação em Ciências da Saúde,
490 Faculdade de Medicina, Av. Pará - 1720 - Bloco 2U - Bairro Umuarama, Uberlândia, Minas

491

Gerais, Brasil

492

yara.maia@ufu.br

493

494

Morun Bernardino Neto

495

Universidade de São Paulo, Departamento de Ciências Básicas e Ambientais, Escola de
496 Engenharia de Lorena, São Paulo, Brasil

497

morun@usp.br

498

² Artigo formatado segundo as normas da revista *Environmental Research* (ISSN 0013-9351).

499 **RESUMO**

500 A análise de forma integrada das principais variáveis de saúde ambiental que causam efeitos
501 adversos na saúde da população é uma etapa fundamental para o desenvolvimento de políticas
502 públicas em saúde. Nesse sentido, essa pesquisa tem como objetivo analisar simultaneamente
503 múltiplas variáveis ambientais e identificar sua estrutura latente com a finalidade de
504 determinar aquelas que são mais eficientes na identificação de efeitos críticos sobre a saúde
505 pública. A teoria científica subjacente e a Análise Fatorial Exploratória (AFE) foram usadas
506 para atingir o objetivo proposto. As suposições teóricas e estatísticas para AFE foram
507 verificadas e os cálculos foram feitos no software IBM-SPSS *Statistics*, versão 22.0. A
508 amostra é composta por 853 observações que se referem aos municípios de Minas Gerais,
509 região Sudeste, Brasil e os dados utilizados foram obtidos em sistemas de informações
510 virtuais. A AFE foi processada com 76 variáveis e a fatorabilidade da amostra foi confirmada
511 pela matriz de correlações de *Spearman* (ρ) com 54,73% de correlações lineares significantes
512 a nível alfa de 5%, pela medida de adequabilidade da amostra de *Kaiser-Meyer-Olkin* de
513 0,952 e pelo teste de esfericidade de *Bartlett* com *p-valor* <0,001. As suposições teóricas
514 (existência de fatores, homogeneidade e correlações conceituais) e estatísticas (linearidade,
515 homocedasticidade e normalidade) foram atendidas. A partir do critério da variância
516 explicada e do teste de *Scree Plot* foram extraídos 5 fatores que explicam 59,78% da variância
517 total dos dados. Estes foram rotulados como F1: socioambiental, F2: vulnerabilidade social,
518 F3: qualidade do ar, F4: mortalidade e F5: agropecuária. E as variáveis marcadoras foram
519 número de óbitos por neoplasia; percentual de pessoas inscritas no Cadastro Único sem
520 abastecimento de água adequado; concentração de NO₂; taxa de mortalidade por homicídio;
521 percentual de cobertura por floresta plantada e percentual de cobertura vegetal natural.
522 Portanto, a análise simultânea de múltiplas variáveis, sua estrutura latente, assim como a
523 determinação daquelas que são mais eficientes na identificação de efeitos críticos na saúde é
524 fundamental para subsidiar estratégias e políticas públicas em saúde.

525 **Palavras-chave:** Saúde pública. Indicadores ambientais. Estatística multivariada.

526 **ABSTRACT**

527 The integrated analysis of the main environmental health variables that cause adverse effects
528 on the population's health is a fundamental step for the development of public health policies.
529 In this sense, this research aimed to analyze multiple environmental variables and to identify
530 their latent structure in order to determine those that are most efficient in identifying critical
531 effects on public health. The underlying scientific theory and Exploratory Factor Analysis
532 (EFA) were used to achieve the proposed objective. Theoretical and statistical assumptions
533 for EFA were verified and calculations were performed in IBM-SPSS Statistics software,
534 version 22.0. The sample consists of 853 observations that refer to the municipalities of Minas
535 Gerais, Southeast region, Brazil and the data used were obtained from virtual information
536 systems. The EFA was processed with 76 variables and the factorability of the sample was
537 confirmed by the Spearman (ρ) correlation matrix with 54.73% of significant linear
538 correlations at the alpha level of 5%, by the measure of suitability of the sample of Kaiser-
539 Meyer-Olkin of 0.952 and by Bartlett's sphericity test with p-value <0.001 . Theoretical
540 assumptions (existence of factors, homogeneity and conceptual correlations) and statistics
541 (linearity, homoscedasticity and normality) were met. From the explained variance criterion
542 and the Scree Plot test, five factors were extracted that explain 59.78% of the total variance of
543 the data. These were named as F1: socio-environmental, F2: social vulnerability, F3: air
544 quality, F4: mortality and F5: agriculture. And their respective marker variables were the
545 number of deaths from neoplasia; percentage of people enrolled in the Single Registry without
546 adequate water supply; NO_2 concentration; homicide mortality rate; percentage of cover per
547 planted forest and percentage of natural vegetation cover. The simultaneous analysis of
548 multiple variables, their latent structure, as well as the determination of those that are most
549 efficient in identifying critical health effects is essential to support strategies and public health
550 policies.

551 **Keywords:** Environmental indicators. Public health. multivariate statistics

552 **Informações de Financiamento:**

553 Esta pesquisa foi financiada pela Coordenação de Aperfeiçoamento de Pessoal de Nível

554 Superior (CAPES) [Bolsa n° 88882.461704, 2019];

555

556 1. Introdução

557 De acordo com a Organização Mundial da Saúde mais de 13 milhões de mortes anuais
558 no mundo se relacionam a causas ambientais que poderiam ser evitáveis (OMS, 2022). A
559 relação existente entre saúde e meio ambiente pode ser vista a partir de variáveis que
560 envolvem a poluição atmosférica, saneamento básico, qualidade da água para consumo
561 humano, qualidade do solo, contaminantes químicos, impactos ambientais, ambiente de
562 trabalho, habitação, desmatamentos e vetores transmissores de doenças.

563 Como pode ser observado pelo levantamento bibliográfico realizado por Lopes e
564 Bernardino Neto (2020), estudos científicos apontam consistentes relações entre o meio
565 ambiente e a saúde pública, em que 60% das pesquisas analisadas abordaram os aspectos de
566 saúde relacionados aos determinantes ambientais, sendo os demais relacionados aos aspectos
567 sociais, econômicos e demográficos. Dentre algumas doenças que tiveram relação direta com
568 o ambiente pode-se destacar a leishmaniose, malária, dengue, leptospirose, doença de Chagas,
569 diarreia, gastroenterite e doenças respiratórias.

570 Os autores Paiva e Souza (2018) apontam relação entre condições sanitárias e
571 morbidade hospitalar por doenças de veiculação hídrica. Complementarmente, os autores
572 Gouveia *et al.* (2019), Nascimento *et al.* (2017) e Saldiva *et al.* (2020) reportam a associação
573 entre a concentração de partículas finas no ar e a mortalidade geral, doenças respiratórias
574 agudas em crianças, efeitos inflamatórios e cardiovasculares.

575 Os efeitos à saúde também estão relacionados às alterações climáticas relativas às
576 mudanças de temperatura, umidade e variação pluviométrica com o risco de disseminação de
577 doenças respiratórias, infecciosas e transmitidas por vetores (SOLOMON e ROCQUE, 2019).
578 Contudo, fatores como a exposição de trabalhadores rurais a agrotóxicos também são
579 importantes indicativos relacionados a intoxicações, distúrbios respiratórios, lesões
580 musculares e doenças mentais (LOPES e ALBUQUERQUE, 2018).

581 Diante dessa conjuntura, estudos sobre a saúde ambiental vêm se acentuando na
582 tentativa de identificar os aspectos ambientais que causam consequências adversas à saúde. O
583 tema é abrangente e envolve diversas variáveis e dimensões, assim sendo, a complexidade de
584 tal avaliação de risco deve refletir na adequabilidade metodológica de análise.

585 Nesse sentido, este trabalho tem como objetivo determinar a estrutura latente das
586 relações entre múltiplas variáveis de saúde ambiental visando inferir efeitos críticos na saúde
587 pública. A abrangência deste tema e sua crescente importância para a sustentabilidade da
588 sociedade contemporânea bem como a evidente interação dos problemas ambientais e saúde
589 pública, justificam esta pesquisa.

590 Para tanto, foi usada a técnica de Análise Fatorial Exploratória (AFE) que tem a
591 habilidade de organizar as variáveis em fatores por meio da análise de suas relações de
592 interdependência, permitindo determinar a estrutura subjacente a elas (ROGERS, 2022). A
593 AFE tem a capacidade de reduzir essas múltiplas variáveis originais a um número menor com
594 uma perda mínima de informação, mantendo a representatividade delas. Dessa forma, é
595 possível definir um conjunto mínimo de fatores que consiga explicar a maior parte da
596 variância do conjunto original (LORENZO e FERRANDO, 2020).

597 A AFE é uma técnica que exige cautela durante a sua realização, sendo indispensável
598 verificar e atender os seus critérios teóricos e estatísticos a fim de obter uma adequada
599 estrutura fatorial e evitar resultados errôneos e /ou não confiáveis.

600 A amostra de dados do conjunto original é composta por 853 observações que se
601 referem aos municípios do estado de Minas Gerais, localizados na região sudeste, Brasil e por
602 130 variáveis de saúde ambiental (APÊNDICE A). A justificativa por essa escolha se deve ao
603 fato da marca de heterogeneidade presente nesses municípios com regiões desenvolvidas e
604 outras que oferecem condições de vida precárias à sua população, sendo o retrato síntese do
605 Brasil (QUEIROZ, 2001). Ademais, essa escolha também levou em consideração o número de
606 observações para a análise estatística, no caso, o estado de Minas Gerais foi o que apresentou
607 o maior número de observações (n=853) (IBGE, 2017).

608 Assim sendo, o uso da AFE tem relevante contribuição para a área da saúde, pois ao
609 determinar as variáveis marcadoras de fatores ambientais que provocam efeitos na saúde
610 pública, podem ser subsidiadas políticas públicas de saúde e tomadas de decisões preventivas
611 e eficazes.

612

613 **2. Materiais e métodos**

614 **2.1. Delineamento do estudo e amostra de dados**

615 Trata-se de um estudo descritivo e exploratório quantitativo que, entretanto, possui base
616 teórica e conceitual visando determinar a estrutura de relações subjacente às variáveis
617 envolvidas na análise por meio de uma técnica estatística multivariada de interdependência, a
618 Análise Fatorial Exploratória (AFE) (LAKATOS e MARCONI, 2022); (ROGERS, 2022).

619 Para tal, foi realizada a revisão da literatura sobre o tema e a obtenção de dados em banco
620 virtuais que se referem às variáveis de saúde ambiental dos municípios do estado de Minas
621 Gerais, localizados na região Sudeste, Brasil (amostra n= 853 observações). Com esse
622 conjunto de dados (n=853) temos a possibilidade de inserir até 85 variáveis na Análise

623 Fatorial levando em consideração o tamanho ideal de 10 observações por variável (HAIR,
624 2009).

625

626 **2.2. Variáveis e coleta dos dados**

627 A lista inicial de indicadores obtida a partir da revisão da literatura vigente consta de 130
628 variáveis (APÊNDICE A). A coleta de dados foi realizada em sistemas de informações
629 virtuais de domínio público e irrestrito para o ano de 2017, tais como: Departamento de
630 Informática do Sistema Único de Saúde (DATASUS), Instituto Brasileiro de Geografia e
631 Estatística (IBGE), Fundação João Pinheiro (FJP), Fundação Estadual do Meio Ambiente
632 (FEAM), Sistema de Informações Ambientais Integrados a Saúde (SISAM), Sistema de
633 Informação de Vigilância da Qualidade da Água (SISAGUA), Instituto da Gestão das Águas
634 de Minas Gerais (IGAM), Sistema Nacional de Informações sobre Saneamento (SNIS) e
635 Instituto Nacional de Pesquisas Espaciais (INPE).

636

637 **2.3. Análise estatística dos dados**

638 A determinação dos constructos foi realizada por meio da Análise Fatorial Exploratória em
639 quatro estágios (HAIR, 2009): (1) seleção das variáveis envolvidas na análise, (2) avaliação
640 das suposições da Análise Fatorial, (3) determinação dos fatores e (4) interpretação e
641 reespecificação dos fatores.

642

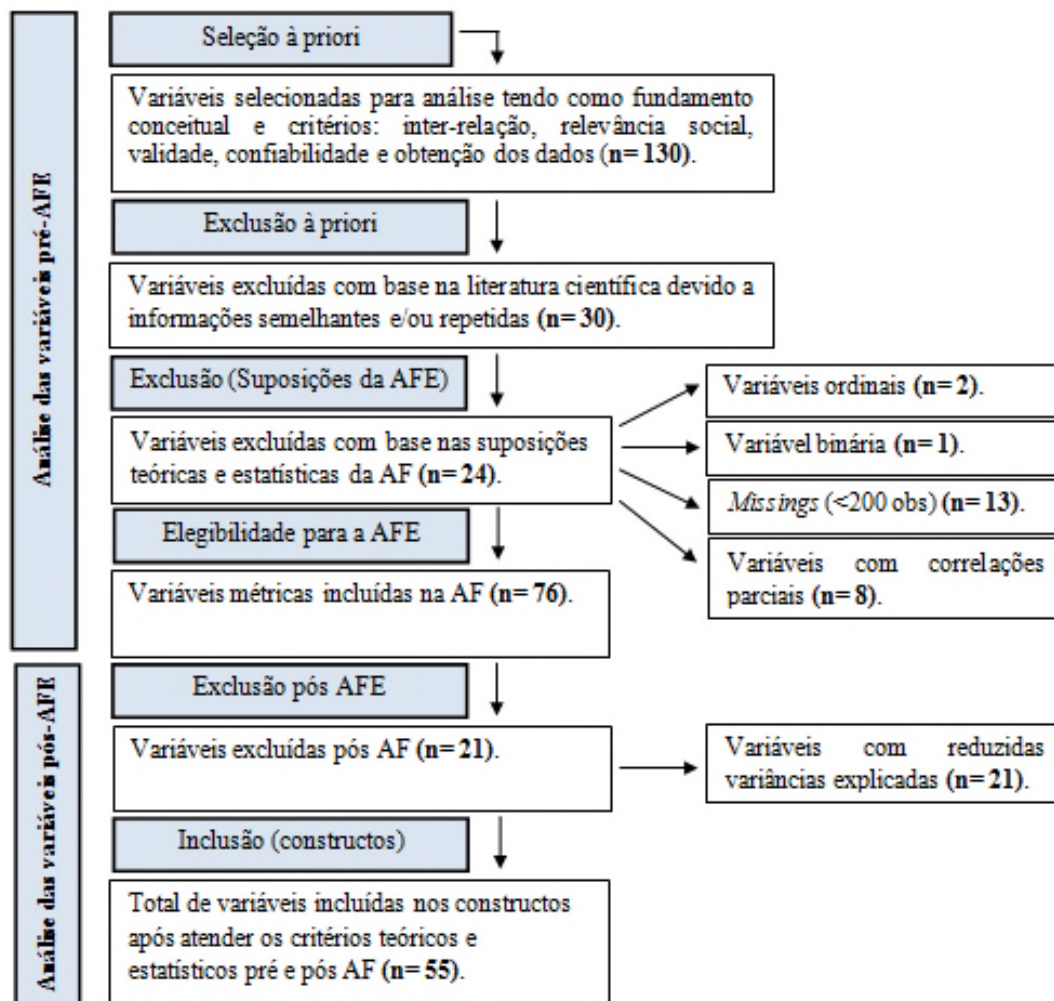
643 **2.3.1. Seleção das Variáveis envolvidas na análise**

644 A princípio, para a seleção inicial das 130 variáveis levou-se em consideração os
645 fundamentos conceituais, tendo como requisito os critérios desejáveis como: inter-relação,
646 relevância social, validade, confiabilidade, comunicabilidade e obtenção dos dados. *A priori*,
647 foram excluídas com base na literatura científica, 30 variáveis que apresentavam informações
648 semelhantes e/ou repetidas. Sucessivamente, após a verificação das suposições teóricas e
649 estatísticas da Análise Fatorial foram excluídas 24 variáveis, sendo duas de natureza ordinal,
650 uma binária, treze com dados ausentes elevados (menos de 200 observações) e oito variáveis
651 com correlações parciais. Assim sendo, foram processadas 76 variáveis métricas na AFE e
652 após a sua realização foram excluídas 21 variáveis com reduzidas variâncias explicadas
653 (variáveis do 6º fator adiante) (Figura 01).

654

655 Figura 01: Fluxograma da seleção das variáveis para estudo.

656



657
658
659

Organização: LOPES, A. F. A., 2022.

660 2.3.2. Avaliação das suposições da AFE

661 As suposições teóricas da AFE são tão importantes quanto seus requisitos estatísticos para
662 que se tenha sucesso na especificação dos constructos latentes. As suposições teóricas são:
663 existência de fatores teóricos supostos, homogeneidade das observações (municípios) em
664 relação à estrutura fatorial e existência de correlações conceituais entre as variáveis em
665 análise. As suposições estatísticas que devem ser verificadas são: fatorabilidade dos dados,
666 ausência de correlações parciais, tamanho adequado de amostra, linearidade,
667 homocedasticidade e normalidade multivariada (HAIR, 2009).

668

669 2.3.3. Determinação dos fatores

670 O método escolhido para especificação da matriz fatorial foi a Análise de Componentes
671 Principais (ACP), o método rotacional usado foi ortogonal (Varimax) e o critério de decisão
672 para o número de fatores a serem extraídos foi pautado no critério de *Scree Plot* e variância
673 explicada.

674 **2.3.4. Interpretação, rotulagem e reespecificação dos fatores**

675 A interpretação dos fatores foi feita por meio das cargas fatoriais e autovalores (raiz latente).
 676 As características das variáveis envolvidas em cada fator assim como a natureza da variável
 677 marcadora foram usadas na rotulagem dos fatores. A reespecificação dos constructos foi feita
 678 na ocorrência de variável com carga cruzada ou comunalidades baixas (menores que 0,40).

679

680 **3. Resultados**

681 As análises de constructos envolveram inicialmente 130 variáveis, a *priori*, foram
 682 excluídas 30 variáveis que apresentavam informações semelhantes e/ou repetidas. Após a
 683 verificação das suposições teóricas e estatísticas da AF também foram eliminadas 24
 684 variáveis, 03 dessas variáveis foram excluídas por consequência de sua natureza não métrica
 685 (VAR 40, VAR 41, VAR 65) e 13 variáveis por conter número elevado de dados ausentes
 686 (VAR 44, VAR 45, VAR 68, VAR 69, VAR 72, VAR 75, VAR 79, VAR 109, VAR 110,
 687 VAR 115, VAR 116, VAR 117 e VAR 120). Previamente ao processamento da análise,
 688 também foi necessário excluir as variáveis com correlações parciais, aquelas com pouca ou
 689 nenhuma correlação significativa: VAR 43, VAR 70, VAR 71, VAR 76, VAR 111, VAR 113,
 690 VAR 114, VAR 119. Desse modo, foram incluídas 76 variáveis na análise fatorial conforme
 691 tabela abaixo (Tabela 01) e suas descrições (APÊNDICE A).

692

693 Tabela 01: Identificação das variáveis excluídas e incluídas na AF

Variáveis excluídas à priori (n=30)													
5	6	8	13	15	24	25	26	27	30	39	51	54	
	56	61	64	66	67	77	83	84	85	89	90	92	
	97	99	101	102	104								
Variáveis excluídas pelas suposições da AF (n=24)													
40	41	43	44	45	65	68	69	70	71	72	75	76	
	79	109	110	111	113	114	115	116	117	119	120		
Total de variáveis excluídas da AF (n=54)													
5	6	8	13	15	24	25	26	27	30	39	40	41	
	43	44	45	51	54	56	61	64	65	66	67	68	
	69	70	71	72	75	76	77	79	83	84	85	89	
	90	92	97	99	101	102	104	109	110	111	113	114	
	115	116	117	119	120								
Variáveis incluídas na AF (n=76)													
1	2	3	4	7	9	10	11	12	14	16	17	18	
	19	20	21	22	23	28	29	31	32	33	34	35	
	36	37	38	42	46	47	48	49	50	52	53	55	
	57	58	59	60	62	63	73	74	78	80	81	82	
	86	87	88	91	93	94	95	96	98	100	103	105	
	106	107	108	112	118	121	122	123	124	125	126	127	
	128	129	130										

714
715 Fonte: LOPES, A. F. A. 2022.

716 A fatorabilidade da amostra foi confirmada pela matriz de correlações de *Spearman*
 717 com 54,73% de correlações significantes à nível de 5%, pela medida de adequabilidade da
 718 amostra de Kaiser-Meyer-Olkin (KMO) de 0,952 e o teste de esfericidade de Bartlett atingiu
 719 nível de significância *p-valor*<0,001 (Tabela 02).

720

721 Tabela 02: Teste de KMO e Bartlett.

Medida Kaiser-Meyer-Olkin de adequação de amostragem.		0,952
Teste de esfericidade de Bartlett	Aprox. Qui-quadrado	8491,840
	gl	2850
	Sig.	<0,0001

722 Fonte: LOPES, A. F. A. 2022.

723

724 A AF do tipo Análise de Componente Principal (ACP) com rotação *Varimax* resultou
 725 em 15 fatores com autovalores maiores que 1 (Tabela 03). A partir da inspeção dos
 726 autovalores, observa-se que o primeiro fator consegue explicar 36,44% da variância total dos
 727 dados, os primeiros cinco fatores atingem 59,781% e os 10 últimos fatores com autovalores
 728 maiores que 1, corroboram com apenas 16,636% de explicação da variância das variáveis.

729

730 Tabela 03: Variância Total Explicada. Método de Extração: análise de Componente Principal.

Componente	Autovalores iniciais		
	Total	% de variância	% cumulativa
1	27.695	36.440	36.440
2	8.068	10.616	47.056
3	4.105	5.402	52.458
4	3.119	4.104	56.562
5	2.446	3.219	59.781
6	1.744	2.295	62.076
7	1.640	2.157	64.234
8	1.480	1.948	66.181
9	1.267	1.666	67.848
10	1.213	1.596	69.444
11	1.115	1.467	70.912
12	1.085	1.428	72.340
13	1.061	1.396	73.736
14	1.031	1.356	75.092
15	1.007	1.325	76.417

731 Fonte: LOPES, A. F. A. 2022.

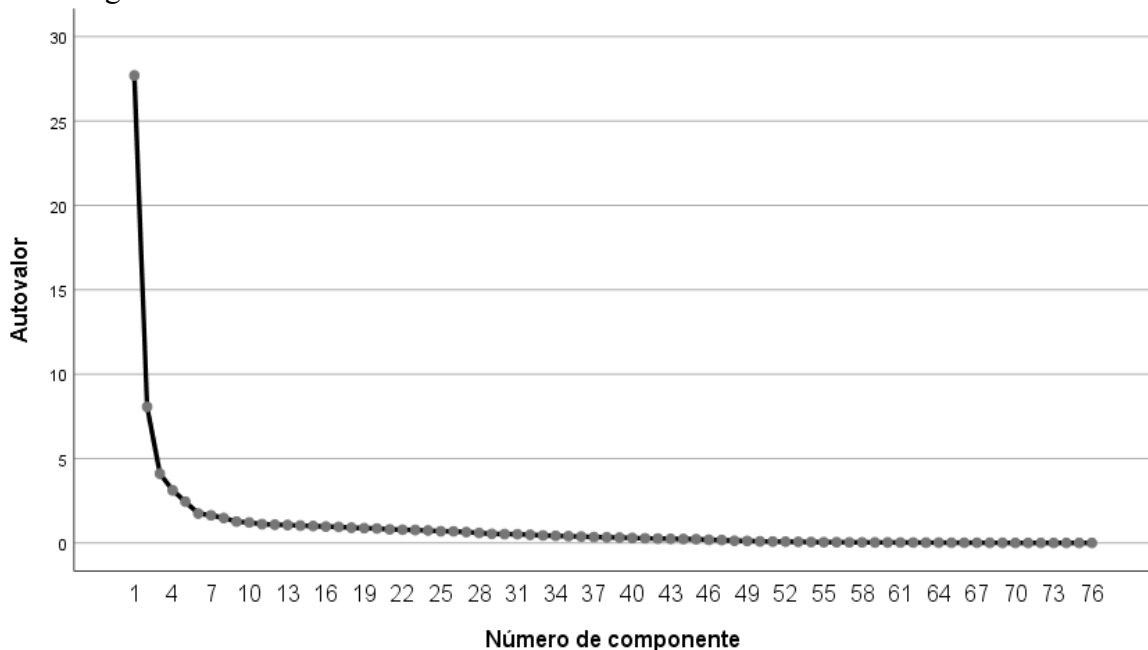
732

733 Tal resultado sugere que a variância de erro (variância específica) passa a ser
 734 substancial a partir do sexto fator, essa suposição pôde ser confirmada por meio do teste de
 735 *Scree* (Figura 02) que nos indica a extração de cinco e não 15 fatores a serem considerados.

736 A convergência de dois critérios para a decisão do número de fatores a serem extraídos
 737 não deixa dúvidas quanto ao número correto de dimensões do constructo que devemos
 738 considerar: (A) total de variância explicada em torno de pelo menos 60% (PEREIRA, 2019),
 739 (B) resultado do teste de *Scree Plot* e, por isso, o critério de extração de fatores com
 740 autovalores maiores que 1 não foi considerado neste estudo.

741

742 Figura 02: Gráfico de *Scree Plot*.



743

744

Fonte: LOPES, A. F. A. 2022.

745

746 O *Scree Plot* é o gráfico que apresenta o padrão de decréscimo dos valores de raiz
 747 latente e nos ajuda a encontrar o ponto de inflexão em que o percentual de variância única
 748 começa a aumentar (ROGERS, 2022). A inspeção do gráfico de *Scree* permite perceber que o
 749 sexto fator é o ponto de inflexão e a partir deste começa a aumentar a variância única e
 750 reduzir a variância explicada, o que justifica a escolha de cinco fatores.

751

752 A matriz de cargas fatoriais maiores que 0,40 não rotacionada foi inconclusiva, contudo
 753 a rotação pelo método Varimax definiu com clareza as variáveis de cada constructo
 754 (APÊNDICE B) e a comunalidade de todas as variáveis que entraram na análise foi maior que
 755 0,40. Com base na matriz de cargas fatoriais com rotação Varimax, as variáveis foram
 agrupadas em cinco fatores que foram assim rotulados (Tabela 04): F1: Socioambiental (31

756 variáveis); F2: Vulnerabilidade Social (10 variáveis); F3: Qualidade do ar (05 variáveis); F4:
757 Mortalidade (05 variáveis) e F5: Agropecuária (04 variáveis).

758

759 Tabela 04: Fatores extraídos e suas respectivas variáveis.

FATORES E RÓTULOS	VARIÁVEIS
F1- Socioambiental	População total (VAR 1); população urbana (VAR 2); densidade populacional (VAR 3); famílias com renda per capita até 1/2 salário mínimo (VAR 12); nº de estabelecimentos de saúde por nível de atenção (VAR 19); nº de profissionais da saúde (VAR 20); nº de Unidades Locais (VAR 28); densidade de veículos (VAR 31); cobertura por infraestrutura urbana (VAR 33); nº de óbitos por pneumonia (VAR 57); nº de óbitos por Doenças Crônicas das Vias Aéreas Inferiores (VAR 58); nº de óbitos por outras Doenças do Aparelho Respiratório (VAR 59); nº de óbitos por neoplasias malignas da traqueia, brônquios e pulmões (VAR 60); nº de óbitos por Doenças Isquêmicas do Coração (VAR 62); nº de óbitos por infarto (VAR 63); nº de óbitos por Doenças Hipertensivas (VAR 78); nº de domicílios precários (VAR 86); nº domicílios em situação de coabitação familiar (VAR 87); nº de domicílios urbanos com pelo menos um tipo de serviço básico inadequado (VAR 88); nº de óbitos por acidente de trabalho grave (VAR 112); nº de óbitos notificados por doença de Chagas (VAR 118); nº de óbitos por algumas doenças infecciosas e parasitárias (VAR 121); nº de óbitos por neoplasias (VAR 122); nº de óbitos por doenças endócrinas nutricionais e metabólicas (VAR 123); nº de óbitos por transtornos mentais e comportamentais (VAR 124); nº óbitos por doenças do sistema nervoso (VAR 125); nº de óbitos por doenças do aparelho circulatório (VAR 126); nº de óbitos por doenças do aparelho circulatório (VAR 127); nº de óbitos por doenças da pele e do tecido subcutâneo (VAR128); nº de óbitos por malformações congênitas (VAR129) e nº de óbitos por causas externas (VAR 130) (n= 31).
F2- Vulnerabilidade Social	Taxa de urbanização (VAR 4); percentual da população pobre no Cadastro Único em relação a população total (VAR 14); taxa de emprego no setor formal (VAR 16); percentual de pessoas cobertas por planos de saúde (VAR 23); percentual de pessoas em situação de vulnerabilidade pelas condições de saneamento básico (VAR 91); percentual de pessoas inscritas no Cadastro Único sem abastecimento de água adequado (VAR 93); percentual de pessoas inscritas no Cadastro Único sem esgotamento sanitário adequado (VAR 94); percentual de pessoas inscritas no Cadastro Único sem coleta de lixo adequada (VAR 95); percentual de pessoas em domicílios com abastecimento de água e esgotamento sanitário inadequados (VAR 96) e percentual de pessoas em domicílios com paredes que não sejam de alvenaria (VAR 98) (n= 10).
F3- Qualidade do ar	Concentração de CO (VAR 46); concentração de O ₃ (VAR 47); concentração de NO ₂ (VAR 48); concentração de SO ₂ (VAR 49) e temperatura do ar (VAR 52) (n=5).

F4- Mortalidade	Mortalidade proporcional da população idosa (VAR 7); taxa de crimes violentos (VAR 103); taxa de ocorrências de homicídios (VAR 105); taxa de mortalidade por homicídio (VAR 106) e taxa de mortalidade por agressão (VAR 107) (n= 5).
F5- Agropecuária	Percentual de cobertura por agropecuária (VAR 29); densidade de rebanho bovino (VAR 32); percentual de cobertura vegetal natural (VAR 34) e de cobertura por floresta plantada (VAR 35) (n=4).

760 Fonte: LOPES, A. F. A. 2022.

761

762 Os fatores foram rotulados ou nomeados de acordo com as características do conjunto
 763 dessas variáveis que os compõe. Ademais, foram identificadas as variáveis marcadoras (mais
 764 representativas) de cada fator com base em suas cargas fatoriais e evidenciadas pela literatura
 765 científica, sendo: número de óbitos por neoplasia (VAR 122); percentual de pessoas inscritas
 766 no Cadastro Único sem abastecimento de água adequado (VAR 93); concentração de NO₂
 767 (VAR 48); taxa de mortalidade por homicídio (VAR 106); percentual de cobertura por
 768 floresta plantada (VAR 35) e percentual de cobertura vegetal natural (VAR 34) (Tabela 05).

769

770 Tabela 05: Fatores (F1-F5), variáveis marcadoras e suas respectivas cargas fatoriais.

FATOR	VARIÁVEL MARCADORA	CARGA FATORIAL
F1	Óbitos por neoplasia (VAR 122);	0,993
F2	Percentual de pessoas inscritas no Cadastro Único sem abastecimento de água adequado (VAR 93);	0,896
F3	Concentração de NO ₂ (VAR 48);	0,898
F4	Taxa de mortalidade por homicídio (VAR 106);	0,953
F5	Percentual de cobertura por floresta plantada (VAR 35) e percentual de cobertura vegetal natural (VAR 34);	0,708 e 0,695

771 Fonte: LOPES, A. F. A. 2022.

772

773 Outras variáveis com carga fatorial elevada também podem ser consideradas marcadoras
 774 tendo como referência o embasamento teórico científico, assim como ocorreu no último fator
 775 em que também consideramos a segunda variável com maior carga fatorial como
 776 representativa (cobertura vegetal natural).

777

778 4. *Discussão*

779 A pressuposição de correlações conceituais entre as variáveis pôde ser confirmada por
 780 meio da base teórica subjacente que possibilitou a definição inicial de 14 dimensões: (1)
 781 Demográfico, (2) Sócio-econômico, (3) Assistência à saúde, (4) Pressão e estado do meio
 782 ambiente, (5) Água para consumo humano e saneamento básico, (6) Qualidade do ar, (7)
 783 Qualidade do solo, contaminantes ambientais e substâncias químicas, (8) Desastres naturais e

784 fatores da natureza, (9) Habitação, (10) Vulnerabilidade social, (11) Violência, (12) Ambiente
 785 de trabalho, (13) Doenças e agravos de notificação (vetores, hospedeiros, reservatórios,
 786 outros), (14) Efeitos à saúde.

787 O tamanho da amostra, que se refere ao total de municípios de Minas Gerais (N=853),
 788 foi suficiente para realizar a análise fatorial uma vez que o tamanho amostral recomendado
 789 leva em consideração a proporção de dez observações para cada variável inserida na análise
 790 (HAIR, 2009) e, com 853 observações temos a possibilidade de incluir até 85 variáveis,
 791 conquanto foram processadas 76 variáveis.

792 No que diz respeito à homogeneidade das observações (municípios) em relação à
 793 estrutura fatorial, não temos motivos para afirmar que existam subpopulações estruturalmente
 794 diferentes em relação aos fatores estudados, o que significa que as relações de
 795 interdependência entre as variáveis e seu fator são homogêneas para qualquer município
 796 envolvido na análise. Não obstante as relações entre as variáveis e entre cada variável e seu
 797 fator sejam proporcionais às suas magnitudes, elas sempre mantêm a mesma forma e sentido
 798 independente do município considerado.

799 A fatorabilidade dos dados, ponto crítico na AFE, foi verificada e confirmada por meio
 800 de três métodos: (A) A matriz de correlações de Spearman (ρ) apresentou um número
 801 substancial de correlações lineares significantes a nível alfa menor que 5% (54,73%) e esse
 802 número aumenta para 59,17% quando consideradas as correlações significantes a nível alfa
 803 menor que 10%; (B) A medida de adequação da amostra de Kaiser-Meyer-Olkin (KMO), que
 804 considera a amostra adequada a partir de valores acima de 0,60 (HAIR, 2009), foi de 0,952 e,
 805 por fim, (C) O teste de esfericidade de Bartlett atingiu nível de significância *p-valor*
 806 $<0,001$ (Tabela 02).

807 A medida de adequabilidade da amostra de KMO apresenta valores que variam de 0 a 1,
 808 escala em que valores próximos a 0 apontam que a AF não é adequada por haver apenas
 809 correlações fracas entre as variáveis em contrapartida, valores próximos de 1 indicam que seu
 810 emprego é apropriado (HONGYU, 2018). Para Hair (2009) são aceitáveis valores acima de
 811 0,60 e Fávero (2009) amplia essa avaliação de forma que os valores são classificados em
 812 muito bom (1-0,90); bom (0,80 -0,90); médio (0,70 -0,80); razoável (0,60 -0,70); ruim (0,50 -
 813 0,60) e inaceitável ($<0,50$).

814 O teste de esfericidade de *Bartlett* avalia a adequabilidade do número de correlações
 815 entre as variáveis, expressando o resultado na forma de um *p-valor*. Para que seja possível a
 816 realização da análise fatorial recomenda-se que o *p-valor* seja menor que 5% (FÁVERO,
 817 2009). O resultado obtido a partir da avaliação da amostra usada nesse trabalho (*p-valor*

818 <0,0001) garante que há relações suficientes entre as variáveis que permita a fatorabilidade,
819 rejeitando-se a hipótese nula (H_0) de que não há correlações suficientes entre as variáveis
820 (HAIR, 2009).

821 As suposições estatísticas requeridas pela AF, homocedasticidade e a normalidade
822 multivariada, são consideradas atendidas quando há número substancial de correlações
823 lineares com coeficientes acima de 0,30 (correlações moderadas à fortes), uma vez que os
824 desvios da normalidade e da homocedasticidade comprometem os resultados apenas via
825 redução das correlações lineares significantes (HAIR, 2009). Com 54,73% de correlações
826 significantes podemos garantir que a fuga da normalidade e da homocedasticidade não
827 ocorreu neste estudo.

828 A matriz de cargas fatoriais não rotacionada não oferece informação adequada das
829 variáveis analisadas quanto a matriz rotacionada porque é comum ocorrer carga fatorial alta
830 para uma mesma variável em mais de um fator, por isso foi realizada a rotação Varimax que
831 permite redistribuir as cargas das variáveis de forma que cada variável se associe mais
832 fortemente a um único constructo, reduzindo as ambigüidades e facilitando a interpretação. A
833 partir da matriz de cargas fatoriais com rotação rotacionada (Varimax) foi possível identificar
834 as variáveis que constituem cada dimensão do constructo (Tabela 04).

835 As cargas fatoriais foram consideradas altas quando atingiam valores maiores do que
836 0,40, contudo, para amostras maiores que 350 cargas fatoriais acima de 0,30 podem ser
837 consideradas sem comprometer o nível de confiança de 95% e poder de 80% (HAIR, 2009).

838 A comunalidade de uma variável expressa a variância compartilhada entre essa
839 variável e as demais variáveis nos fatores obtidos, este é um meio útil para avaliar a proporção
840 de explicação da variável pelos constructos extraídos, sendo o valor mínimo aceitável de 0,40.
841 Contudo, quanto maior a comunalidade, maior é o percentual de explicação da variável pelas
842 demais variáveis dos fatores obtidos (FÁVERO, 2009). Todas as variáveis envolvidas nessa
843 análise apresentaram uma boa variância explicada pelos fatores com comunalidades maiores
844 do que 0,40 e majoritariamente maiores do que 0,50 (APÊNDICE B).

845 Foram eliminadas as variáveis com cargas fatoriais menores do que 0,39 (correlações
846 insignificantes) e as cargas mantidas ($>0,39$) indicam aquelas que mais se relacionam com os
847 seus fatores. As variáveis com cargas fatoriais positivas apresentam relação de
848 interdependência direta com o seu fator. Por outro lado, carga fatorial negativa aponta para
849 uma relação de interdependência inversa entre os valores dessa variável e seu constructo.

850 A variável marcadora do primeiro fator refere-se ao número de óbitos por neoplasia
851 (VAR 122) com carga 0,993. Tal resultado está sonante com a literatura científica que aponta

852 para o meio ambiente como responsável por 80% dos casos de câncer (ONCOH, 2022).
853 Existem evidências da relação do ambiente de trabalho com casos de câncer de pulmão,
854 próstata, leucemia, fígado e estômago devido à exposição de substâncias químicas como
855 metais pesados, poeiras inorgânicas e agrotóxicos (BALDO *et al.* 2021).

856 O segundo fator teve como variável marcadora o percentual de pessoas inscritas no
857 Cadastro Único sem abastecimento de água adequado (VAR 93). É inegável que a falta de
858 saneamento básico, inclusive a ausência ou a precariedade do acesso à água representa
859 situação de risco para a saúde pública, corroborando para o aumento da incidência de doenças
860 infecciosas e parasitárias, principalmente em áreas que abrigam a população socialmente
861 excluída (ARRUDA e HELLER, 2022). O fornecimento de água de boa qualidade para o
862 consumo humano assegura a redução e o controle de diarreias, febre tifoide, hepatites,
863 dengue, cólera e doenças parasitárias (BRASIL, 2019).

864 No terceiro fator (qualidade do ar), a variável marcadora foi a concentração de NO₂
865 (VAR 48) com carga fatorial igual a 0,898. As evidências científicas reportam que a poluição
866 atmosférica é responsável por danos críticos à saúde e notadamente atingem o sistema
867 respiratório e cardiovascular (CHAVES *et al.* 2017). De acordo com Nascimento (2017), o
868 dióxido de nitrogênio (NO₂) é um poluente ambiental resultante da queima de combustíveis
869 fósseis com origem principalmente, no tráfego de automóvel e setor industrial. A investigação
870 científica de Godoy *et al.* (2021), que objetivou analisar os efeitos na saúde cardiorrespiratória
871 de uma população exposta a poluentes atmosféricos como material particulado (MP10),
872 dióxido de nitrogênio (NO₂), dióxido de enxofre (SO₂) e ozônio (O₃) constatou
873 estatisticamente que o NO₂ foi o poluente atmosférico que apresentou uma correlação positiva
874 mais significativa com os efeitos deletérios à saúde cardiorespiratória.

875 Outras variáveis do ambiente que impactam na saúde humana são descritas no
876 constructo 4 que se referem aos aspectos de mortalidade, com destaque para a variável
877 marcadora “taxa de mortalidade por homicídio” (VAR 106) com carga fatorial igual 0,953. Os
878 impactos causados pela violência comprometem a saúde física e emocional das famílias,
879 colaborando com essa informação, foi comprovado que o homicídio desencadeia depressão,
880 transtornos psicológicos e comportamentos agressivos (COSTA *et al.* 2017). O estudo de Sá
881 Neto *et al.* (2020) também contribui para essa discussão confirmando que a vulnerabilidade
882 social e o homicídio se mostram associados.

883 Por fim, abrangendo o tema agricultura e pecuária, nota-se que as variáveis são
884 inversamente relacionadas, ou seja, na medida em que se aumenta a cobertura por
885 agropecuária (VAR 29) e densidade de rebanho bovino (VAR 32), tem-se uma redução da

886 cobertura vegetal natural (VAR 34) e da cobertura por floresta plantada (VAR 35). Para esse
887 fator é importante salientar que os resultados da AFE apontam que a “cobertura por floresta
888 plantada” é a variável marcadora com carga fatorial (0,708). No entanto, como observado no
889 levantamento bibliográfico, a “relação da cobertura vegetal natural com efeitos à saúde” foi a
890 segunda maior carga fatorial (0,695), destarte, selecionamos as duas variáveis com maiores
891 cargas fatoriais como representativas desse fator. É evidente a relação entre o desmatamento e
892 a redução da cobertura natural sobre a incidência de doenças infecciosas causadas por vetores
893 transmissores de doenças. Os estudos de impactos ambientais relatam que o desmatamento
894 tem efeito significativo sobre o aumento da incidência da leishmaniose e da malária
895 (MORENO *et al.* 2018). Complementarmente, as evidências científicas também destacam os
896 benefícios da vegetação natural para a saúde da população como redução da temperatura e
897 ruídos, aumento da umidade e captura do material particulado (LOURENÇO *et al.* 2016).

898 Assim sendo, estas seis variáveis marcadoras supracitadas representam as 76 variáveis
899 originais inseridas na AFE com uma perda mínima de informação e mantendo a
900 representatividade do conjunto original.

901

902 **5. Conclusão**

903 O emprego da estatística multivariada por meio da Análise Fatorial Exploratória
904 permitiu agrupar as variáveis de acordo com suas correlações e gerando grupos com aquelas
905 altamente relacionadas entre si. Ademais, essa técnica realizou uma redução do conjunto
906 original de variáveis e fatores propostos inicialmente com perda mínima de informação.

907 A aplicação da AFE permitiu a extração de 5 fatores: F1) socioeconômico, demográfico,
908 pressão ambiental e efeitos à saúde; F2) vulnerabilidade social; F3) qualidade do ar; F4)
909 mortalidade e violência e o F5) Agricultura e pecuária. Estes, em conjunto, explicam 59,781%
910 da variância dos dados. Também foi possível encontrar a variável marcadora (a mais
911 representativa do fator), sendo respectivamente: número de óbitos por neoplasia; percentual
912 de pessoas inscritas no Cadastro Único sem abastecimento de água adequado; concentração
913 de NO₂; taxa de mortalidade por homicídio e cobertura vegetal natural.

914 Portanto, pode-se considerar que a análise simultânea de múltiplas variáveis, sua
915 estrutura latente, assim como a determinação daquelas que são mais eficientes na
916 identificação de efeitos críticos na saúde é fundamental para subsidiar a tomada de decisões
917 eficazes por governantes e agentes de saúde em termo de estratégias e políticas públicas em
918 saúde.

919

920 **6. Agradecimentos**

921 À Universidade Federal de Uberlândia por meio do Programa de Pós-Graduação em Ciências da
922 Saúde PPCSA/UFU e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.
923

924 **7. Referências**

925 ARRUDA, A., HELLER, L. Acesso à água e esgotos em ocupação urbana na Região Metropolitana de
926 Belo Horizonte: efeitos na saúde, qualidade de vida e relações de gênero. **Physis: Revista de Saúde**
927 **Coletiva**, v. 32, n. 2, 2022. DOI: <https://doi.org/10.1590/S0103-73312022320204>. Acesso em: 16 de
928 agosto de 2022.

929 BALDO, R. C. S. *et al.* Nexo Epidemiológico do Câncer Relacionado ao Trabalho no Município de
930 Londrina-PR. **Revista Brasileira de Cancerologia**, v. 67, n. 3, e-141328, 2021. DOI: 10.32635/2176-
931 9745. RBC.2021v67n3.1328. Acesso em: 17 de agosto de 2022.

932 BRASIL. **Saúde Brasil 2018 uma análise de situação de saúde e das doenças e agravos crônicos:**
933 **desafios e perspectivas** / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de
934 Vigilância de Doenças e Agravos Não Transmissíveis – Brasília: Ministério da Saúde, 2019.

935 CHAVES, L. E., NASCIMENTO, L. F. C., RIZOL, P.M. S. R. F. Model to estimate the number of
936 hospitalizations for asthma and pneumonia under the effects of air pollution. **Revista de Saúde**
937 **Pública**, v. 51, 2017. DOI: <https://doi.org/10.1590/S1518-8787.2017051006501>. Acesso em: 16 de
938 agosto de 2022.

939 COSTA, D. H. *et al.* Homicídios de jovens: os impactos da perda em famílias de vítimas. **Revista de**
940 **Saúde Coletiva**, v. 27, n.3, p. 685-700, 2017. DOI: [https://doi.org/10.1590/S0103-](https://doi.org/10.1590/S0103-73312017000300016)
941 [73312017000300016](https://doi.org/10.1590/S0103-73312017000300016). Acesso em: 16 de agosto de 2022.

942 FÁVERO, L.P. *et al.* **Análise de dados: modelagem multivariada para tomada de decisões.** Rio de
943 Janeiro: Elsevier, 2009.

944 GODOY, R. M. P., ABE, K. C., MIRAGLIA, S. G. E. K. Correlação entre desfechos na saúde
945 cardiorrespiratória e exposição a poluentes atmosféricos – um estudo na megacidade de São Paulo.
946 **Revista De Ciências Da Saúde**, v. 33, n. 3, p. 33–46, 2021. DOI: 10.14295/vittalle.v33i3.13390.

947 GOUVEIA, N. *et al.* Poluição do ar e impactos na saúde na Região Metropolitana de Belo Horizonte –
948 Minas Gerais, Brasil. **Ciência & Saúde Coletiva**, v. 24, n. 10, p. 3773-3781, 2019. DOI:
949 <https://doi.org/10.1590/1413-81232018.29432017>. Acesso em 16 de agosto de 2022.

950 HAIR, J. F. **Análise multivariada de dados** [recurso eletrônico] / tradução Adonai Schlup Sant’Anna.
951 – 6. ed. – Dados eletrônicos. – Porto Alegre: Bookman, 2009.

952 HONGYU, K. Exploratory Factorial Analysis: theoretical summary, application an interpretation.
953 **E&S.Engineering and Science**, v.7, n. 4, p. 88-103, 2018. DOI:10.18607/ES201877599. Acesso em:
954 27 de julho de 2022.

955 IBGE - INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Panorama Brasil**,
956 v4.6.25.1, 2017. Disponível em: <https://cidades.ibge.gov.br/brasil/panorama>. Acesso em: 02 de maio
957 de 2022.

958 LAKATOS, M. E. MARCONI, M. A. **Metodologia Científica**. 8 ed. 392p. São Paulo: Atlas, 2022.

959 LORENZO, S. U., FERRANDO, P. J. Not Positive Definite Correlation Matrices in Exploratory Item
960 Factor Analysis: Causes, Consequences and a Proposed Solution. **A Multidisciplinary Journal**,
961 28(1), 1–10, 2020. DOI: <https://doi.org/10.1080/10705511.2020.1735393>. Disponível em:
962 <https://www.tandfonline.com/doi/full/10.1080/10705511.2020.1735393>. Acesso em: 20 de setembro
963 de 2022.

964 LOPES, C. V. A., ALBUQUERQUE, G. S. S. C. Agrotóxicos e seus impactos na saúde humana e
965 ambiental: uma revisão sistemática. **Saúde em Debate**, v. 42, n. 117, p. 518-534, 2018. DOI:
966 <https://doi.org/10.1590/0103-1104201811714>. Acesso em: 17 de maio de 2022.

- 967 LOPES, A. F. A., BERNADINO NETO, M. Health, Environment and RIPSA indicators: A literature
 968 review. **Brazilian Journal of Animal and Environmental Research**, v. 3, p. 2610-2626, 2020. DOI:
 969 10.34188/bjaerv3n3-168. Acesso em: 18 de maio de 2022.
- 970 LOURENÇO, A. L. M. F. *et al.* Metr opolis, cobertura vegetal,  reas verdes e sa de. **Estudos**
 971 **Avan ados**, v. 30, n. 86, p. 113-130, 2016. DOI: <https://doi.org/10.1590/S0103-40142016.00100008>.
- 972 MORENO, E. S. *et al.* Licenciamento ambiental de grandes empreendimentos: quais os limites para
 973 avalia o de impactos diretos e indiretos em sa de? **Ci ncias Humanas**, v. 13, n. 3. p. 519-540, 2018.
 974 DOI: <https://doi.org/10.1590/1981.81222018000300003>. Acesso em: 16 de agosto de 2021.
- 975 NASCIMENTO, A. P. *et al.* Association between the concentration of fine particles in the atmosphere
 976 and acute respiratory diseases in children. **Revista de Sa de P blica**, v. 51, n. 0, 2017. Acesso em: 16
 977 de agosto de 2022. DOI: <https://doi.org/10.1590/S1518-8787.2017051006523>.
- 978 ONCOH - Organiza o Nacional de Centros de Oncologia e Hematologia. **Os fatores de riscos**
 979 **ambientais**. 2022. Dispon vel em: <https://cccancer.net/o-cancer/os-fatores-de-risco-ambientais/>.
 980 Acesso em: 16 de agosto de 2022.
- 981 OMS - ORGANIZA O MUNDIAL DE SA DE. **Coopera o estrat gica**: Sa de e
 982 Desenvolvimento. 2013. Dispon vel em: [http://www.who.int/country-cooperation/what-who-](http://www.who.int/country-cooperation/what-who-does/strategies-and-briefs/en/)
 983 [does/strategies-and-briefs/en/](http://www.who.int/country-cooperation/what-who-does/strategies-and-briefs/en/)>. Acesso em: 12 de junho de 2019.
- 984 _____ . **World Health Statistics 2022: Monitoring Health for the SDGs, Sustainable**
 985 **Development Goals**, 2022. Dispon vel em: [https://www.who.int/data/gho/publications/world-health-](https://www.who.int/data/gho/publications/world-health-statistics)
 986 [statistics](https://www.who.int/data/gho/publications/world-health-statistics). Acesso em: 20 de setembro de 2022.
- 987 PAIVA, R. F. P. S., SOUZA, F.P.S. Associa o entre condi es socioecon micas, sanit rias e de
 988 aten o b sica e a morbidade hospitalar por doen as de veicula o h drica no Brasil. **Cadernos de**
 989 **Sa de P blica**, v. 34, n. 1, 2018. DOI: <https://doi.org/10.1590/0102-311X00017316>. Acesso em: 16
 990 de junho de 2022.
- 991 PEREIRA, *et al.* **An lise Fatorial**. Texto para discuss o N  02/2019. Faculdade de Ci ncias
 992 Econ micas, Administrativas e Cont beis, Universidade de Passo Fundo. Passo Fundo, RS, 2019.
 993 Dispon vel em: <https://www.upf.br/Conteudo/textos-discussao/texto.pdf>>. Acesso em: 10 de agosto
 994 de 2022.
- 995 QUEIROZ, B.L. **Diferenciais de sal rios nas regi es mineiras**. Belo Horizonte, 2001. 191f.
 996 Disserta o (mestrado em demografia) – Faculdade de Ci ncias Econ micas, Universidade Federal de
 997 Minas Gerais, Belo Horizonte, 2001.
- 998 ROGERS, P. Best Practices for Your Exploratory Factor Analysis: A Factor Tutorial. **Revista de**
 999 **Administra o Contempor nea** [online], ISSN 1982-7849, v. 26, n. 06, e210085, 2022. DOI:
 1000 <https://doi.org/10.1590/1982-7849rac2022210085>. Dispon vel em: [https://doi.org/10.1590/1982-](https://doi.org/10.1590/1982-7849rac2022210085.en)
 1001 [7849rac2022210085.en](https://doi.org/10.1590/1982-7849rac2022210085.en) <https://doi.org/10.1590/1982-7849rac2022210085.por>>. Acesso em: 20 de
 1002 setembro de 2022.
- 1003 S  NETO, J.A. *et al.* Situational diagnosis of mortality among adolescents in conditions of
 1004 vulnerability. **Nursing Journal**, v. 28, e39279, 2020. DOI: <https://doi.org/10.12957/reuerj.2020.39279>.
 1005 Acesso em: 15 de agosto de 2022.
- 1006 SALDIVA, P.H.N. *et al.* Exercising in the urban center: Inflammatory and cardiovascular effects of
 1007 prolonged exercise under air pollution. **Chemosphere**, Sep; 254:126817, 2020. DOI:
 1008 [10.1016/j.chemosphere.2020.126817](https://doi.org/10.1016/j.chemosphere.2020.126817). Acesso em: 16 de junho de 2022.
- 1009 SOLOMON, C. G., ROCQUE, R. C. Climate Change - A Health Emergency. **New England Journal**
 1010 **of Medicine**, v. 380, n. 3, p. 209-211, 2019 Jan 17. DOI: [10.1056/NEJMp1817067](https://doi.org/10.1056/NEJMp1817067). Acesso em: 16 de
 1011 junho de 2022.
 1012

1013 APÊNDICES

1014 APÊNDICE A: Lista de variáveis originais e descrição.

VAR.	DESCRIÇÃO
VAR1	População total do município;
VAR2	População urbana do município;
VAR3	Densidade populacional (hab/km ²);
VAR4	Taxa de urbanização;
VAR5	Esperança de vida ao nascer;
VAR6	Índice de envelhecimento;
VAR7	Mortalidade proporcional da população idosa (%);
VAR8	Mortalidade até 5 anos de idade (%);
VAR9	Taxa de mortalidade infantil;
VAR10	Taxa bruta de mortalidade;
VAR11	Taxa de escolarização líquida (ensino médio) (%);
VAR12	Número de famílias com renda per capita até 1/2 salário mínimo;
VAR13	Percentual de pobres;
VAR14	Percentual da população pobre no Cadastro Único em relação à população total;
VAR15	Percentual de pessoas em idade produtiva e sem ocupação do Cadastro Único;
VAR16	Taxa de emprego no setor formal (%);
VAR17	Razão de dependência;
VAR18	Gasto per capita com atividades de saúde;
VAR19	Número geral de estabelecimentos de saúde por nível de atenção;
VAR20	Número de profissionais da saúde;
VAR21	Coberturas vacinais;
VAR22	Proporção da população atendida pela Estratégia de Saúde da Família;
VAR23	Percentual de pessoas cobertas por planos de saúde suplementar;
VAR24	Percentual de nascidos vivos com pelo menos 7 consultas de pré-natal;
VAR25	Proporção de nascidos vivos com baixo peso (%);
VAR26	Percentual de óbitos por causas mal definidas;
VAR27	Percentual de óbitos por causas mal definidas sem assistência médica;
VAR28	Número de Unidades Locais - Indústria, Construção, Comércio e Serviços;
VAR29	Cobertura por agropecuária (%);
VAR30	Cobertura por plantação de cana-de-açúcar (%);
VAR31	Densidade de veículos (unidade/km ²);
VAR32	Densidade de rebanho bovino (unidade/km ²);
VAR33	Cobertura por infraestrutura urbana (%);
VAR34	Cobertura vegetal natural (%);
VAR35	Cobertura por floresta plantada (%);
VAR36	Participação nos focos de calor do estado (%);
VAR37	Percentual de água tratada;
VAR38	Percentual de amostras fora dos padrões microbiológicos;
VAR39	Percentual de amostras fora dos padrões organolépticos;
VAR40	Índice de Qualidade da Água – Média Anual;
VAR41	Qualidade da água: Classe de contaminação por tóxicos;
VAR42	Percentual de esgoto tratado;
VAR43	Taxa de mortalidade por cólera;
VAR44	Taxa de mortalidade por diarreia e gastroenterite;

VAR45	Taxa de mortalidade por outras doenças infecciosas intestinais (febre tifóide, paratifóide, enterite por salmonela, shigelose, infecções intestinais por <i>escherichia coli</i>);
VAR46	Concentração de CO (ppb);
VAR47	Concentração de O ₃ (ppb);
VAR48	Concentração de NO ₂ (ppb);
VAR49	Concentração de SO ₂ (μg/m ³);
VAR50	Concentração de PM _{2.5} (μg/m ³);
VAR51	Concentração de PM ₁₀ (μg/m ³);
VAR52	Temperatura do ar (°C);
VAR53	Umidade relativa do ar (%);
VAR54	Precipitação acumulada (mm);
VAR55	Taxa de mortalidade por câncer de pulmão (por 100 mil habitantes);
VAR56	Número de óbitos por influenza (gripe);
VAR57	Número de óbitos por pneumonia;
VAR58	Número de óbitos por Doenças das Vias Aéreas Inferiores (bronquite, enfisema, asma);
VAR59	Número de óbitos por outras Doenças do Aparelho Respiratório (nasofaringite, sinusite, faringite, amigdalite, laringite);
VAR60	Número de óbitos por neoplasias da traquéia, brônquios e pulmões;
VAR61	Taxa de mortalidade por doenças respiratórias em crianças menor que 1 ano;
VAR62	Número de óbitos por Doenças Isquêmicas do Coração;
VAR63	Número de óbitos por Infarto Agudo do Miocárdio;
VAR64	Os resíduos possuem destino final em lixões a céu aberto?
VAR65	O município apresentou registro de áreas contaminadas?
VAR66	O município apresentou registro de acidentes ambientais?
VAR67	Número de áreas cadastradas como contaminadas e reabilitadas para municípios com 4 ou mais áreas;
VAR68	Número de óbitos por Intoxicação Exógena;
VAR69	Número de óbitos de intoxicação exógena por agrotóxico;
VAR70	Número de óbitos de intoxicação exógena por produto químico;
VAR71	Número de óbitos de intoxicação exógena por alimento e bebida;
VAR72	Número de óbitos por exposição a substâncias nocivas;
VAR73	Índice de vulnerabilidade a mudanças climáticas do território; (grau de exposição, sensibilidade e capacidade de adaptação);
VAR74	Parcela de Domicílios em Situação de Risco de Inundação;
VAR75	Número de óbitos por exposição à corrente elétrica, radiação, temperatura e pressão extrema;
VAR76	Número de óbitos por exposição às forças da natureza (calor excessivo, frio excessivo, raio, desabamento de terra, inundação);
VAR77	Índice de óbitos em decorrência de eventos hidrológicos;
VAR78	Número de óbitos por Doenças Hipertensivas (DH);
VAR79	Número de óbitos por neoplasia maligna da pele;
VAR80	Percentual da população urbana em domicílios com abastecimento de água;
VAR81	Percentual da população urbana em domicílios com esgotamento sanitário;
VAR82	Percentual da população urbana atendida por serviços de coleta de resíduos;
VAR83	Existência de coleta seletiva;
VAR84	Percentual da população em domicílios com banheiro e água encanada;
VAR85	Déficit habitacional;
VAR86	Número de domicílios precários;
VAR87	Número de domicílios em situação de coabitação familiar;

VAR88	Número de domicílios c/ pelo menos um tipo de serv. básico inadequado;
VAR89	Número de domicílios alugados com adensamento excessivo;
VAR90	Número de domicílios urbanos próprios com adensamento excessivo;
VAR91	Percentual de pessoas em situação de vulnerabilidade pelas condições de saneamento básico;
VAR92	Percentual de pessoas no Cadastro Único sem abastecimento de água, esgoto sanitário e coleta de lixo adequados;
VAR93	Percentual de pessoas no Cadastro Único sem abastecimento de água;
VAR94	Percentual de pessoas no Cadastro Único sem esgotamento sanitário;
VAR95	Percentual de pessoas no Cadastro Único sem coleta de lixo adequada;
VAR96	Percentual de pessoas em domicílios com abastecimento de água e esgotamento sanitário inadequados;
VAR97	Percentual de pessoas em domicílios sem energia elétrica;
VAR98	Percentual de pessoas em domicílios com paredes que não sejam de alvenaria;
VAR99	Número de domicílios urbanos sem banheiro;
VAR100	Percentual de pessoas em domicílios vulneráveis à pobreza e que gastam mais de 1 hora até o trabalho;
VAR101	Percentual de meninas de 10 a 14 anos de idade que tiveram filhos;
VAR102	Percentual de adolescentes de 15 a 17 anos de idade que tiveram filhos;
VAR103	Taxa de crimes violentos (por 100 mil habitantes);
VAR104	Taxa de crimes violentos contra a pessoa (por 100 mil habitantes);
VAR105	Taxa de ocorrências de homicídios dolosos (por 100 mil habitantes);
VAR106	Taxa de mortalidade por homicídio (por 100 mil habitantes);
VAR107	Taxa de mortalidade por agressão (por 100 mil habitantes);
VAR108	Taxa de mortalidade por suicídio;
VAR109	Número de óbitos por doenças relacionadas ao trabalho;
VAR110	Número de óbitos de pneumoconiose (de minérios de carvão, amianto, à poeira de sílica);
VAR111	Número de óbitos de intoxicações exógenas por exposição ao trabalho;
VAR112	Número de óbitos por acidente de trabalho grave;
VAR113	Número de óbitos de câncer relacionado ao trabalho;
VAR114	Número de óbitos notificados de dengue;
VAR115	Número de óbitos notificados de febre amarela;
VAR116	Número de óbitos notificados de Leishmaniose;
VAR117	Número de óbitos notificados de Esquistossomose;
VAR118	Número de óbitos notificados por doença de Chagas;
VAR119	Número de óbitos notificados de Leptospirose;
VAR120	Número de óbitos notificados de hepatites;
VAR121	Número de óbitos por algumas doenças infecciosas e parasitárias (cólera, gastroenterite, tuberculose, peste, difteria, coqueluche, malária, toxoplasmose);
VAR122	Número de óbitos por neoplasias;
VAR123	Número de óbitos por doenças endócrinas nutricionais e metabólicas;
VAR124	Número de óbitos por transtornos mentais e comportamentais;
VAR125	Número de óbitos por doenças do sistema nervoso;
VAR126	Número de óbitos por doenças do aparelho circulatório;
VAR127	Número de óbitos por doenças do aparelho respiratório;
VAR128	Número de óbitos por doenças da pele e do tecido subcutâneo;
VAR129	Número de óbitos por malformações congênitas, deformidades e anomalias cromossômicas;
VAR130	Número de óbitos por causas externas;

1015 **APÊNDICE B: Matriz de Componente Rotativa (Cargas fatoriais e comunalidades).**

Variável	Componente					Comunalidades
	1	2	3	4	5	
VAR1	0,978					0,965
VAR2	0,967					0,946
VAR3	0,824					0,872
VAR4		-0,831				0,782
VAR7				-0,428		0,472
VAR12	0,927					0,933
VAR14		0,533	-0,453			0,750
VAR16		-0,478				0,708
VAR19	0,973					0,960
VAR20	0,985					0,978
VAR23		-0,499				0,677
VAR28	0,988					0,985
VAR29					-0,837	0,870
VAR31	0,845					0,804
VAR32					-0,730	0,652
VAR33	0,704					0,827
VAR34					0,695	0,762
VAR35					0,708	0,631
VAR46			0,778			0,780
VAR47			0,484			0,782
VAR48			0,898			0,904
VAR49			0,854			0,833
VAR52			-0,609			0,801
VAR57	0,928					0,931
VAR58	0,975					0,964
VAR59	0,974					0,954
VAR60	0,983					0,973
VAR62	0,967					0,946
VAR63	0,933					0,925
VAR78	0,939					0,914
VAR86	0,754					0,752
VAR87	0,986					0,986
VAR88	0,678					0,738
VAR91		0,569				0,471
VAR93		0,896				0,849
VAR94		0,766				0,672
VAR95		0,855				0,836
VAR96		0,575				0,559
VAR98		0,449				0,537
VAR103				0,451		0,662
VAR105				0,866		0,769
VAR106				0,953		0,928
VAR107				0,952		0,926
VAR112	0,692					0,533
VAR118	0,762					0,690
VAR121	0,947					0,943
VAR122	0,993					0,992
VAR123	0,971					0,959
VAR124	0,969					0,959
VAR125	0,981					0,974
VAR126	0,985					0,991
VAR127	0,974					0,981
VAR128	0,938					0,886
VAR129	0,970					0,951
VAR130	0,976					0,979

1016 Fonte: LOPES, A. F. A., 2022. "Método de Extração: análise de Componente Principal. Método de
1017 Rotação: Varimax com Normalização de Kaiser." a. Rotação convergida em 17 iterações.

24 **RESUMO**

25 As pressões ambientais provocam efeitos adversos à saúde pública, sendo fundamental o
 26 desenvolvimento de pesquisas nessa área para a tomada de decisões eficazes. Diante desse
 27 cenário, esse trabalho objetiva encontrar a melhor combinação linear de variáveis ambientais
 28 que seja capaz de explicar desfechos críticos à saúde pública (óbitos). Preliminarmente, foram
 29 processados os cálculos de Análise Fatorial Exploratória (AFE), a fim de reduzir o conjunto
 30 original de variáveis e identificar as variáveis marcadoras. Em seguida, por meio de
 31 Regressão Linear Múltipla (RLM) *Stepwise* foi identificado o melhor modelo explicativo do
 32 comportamento da variável dependente ‘taxa bruta de mortalidade’ pelas variáveis
 33 independentes (marcadoras). Todas as análises estatísticas foram realizadas pelo software
 34 IBM-SPSS *Statistics*, versão 22.0. O processamento da AFE permitiu identificar 6 variáveis
 35 com elevada carga fatorial, sendo: número de óbitos por neoplasia; percentual de pessoas
 36 inscritas no Cadastro Único sem abastecimento de água adequado; concentração de NO₂; taxa
 37 de mortalidade por homicídio; percentual de cobertura por floresta plantada e percentual de
 38 cobertura vegetal natural. O melhor modelo matemático encontrado pelo método de RLM
 39 ($R_{adj}^2 = 0,126$, p-valor < 0,001) foi $Y = 7,655 + (-0,289 X_1) + (0,132 X_2) + (-0,109 X_3)$, no
 40 qual, a variação da variável dependente (taxa bruta de mortalidade) é prevista pelas variáveis
 41 ambientais: X_1 = percentual de cobertura vegetal natural ($\beta = -0,289$; p-valor = 0,000), X_2 =
 42 taxa de homicídio ($\beta = 0,132$; p-valor = 0,000) e X_3 = percentual de cobertura por floresta
 43 plantada ($\beta = -0,109$; p-valor = 0,001). As suposições de linearidade, homogeneidade de
 44 variância de erro (homocedasticidade), independência e normalidade dos resíduos e ausência
 45 de colinearidade foram verificadas e atendidas. Os resultados são capazes de suportar, por
 46 meio de um método matemático, que as variáveis ambientais (cobertura vegetal natural e
 47 cobertura por floresta plantada) estão relacionadas inversamente com a variável taxa bruta de
 48 mortalidade e a variável taxa de homicídio se relaciona diretamente com esse desfecho.
 49 Assim, políticas públicas em saúde devem ser direcionadas com maior atenção para estas
 50 variáveis como boas indicadoras de tendência da taxa bruta de mortalidade, sendo uma forma
 51 eficiente de monitorar, antever e estudar cenários de saúde pública.

52 **Palavras-chave:** Meio Ambiente. Saúde Pública. Estatística Multivariada

53

54

55 **ABSTRACT**

56 Environmental pressures cause adverse effects on public health, and the development of
 57 research in this area is essential for effective decision-making. Given this scenario, this work
 58 aims to find the best linear combination of environmental variables that is capable of
 59 explaining critical public health outcomes (deaths). Preliminarily, Exploratory Factor
 60 Analysis (EFA) calculations were processed in order to reduce the original set of variables
 61 and identify the marker variables. Then, through Stepwise Multiple Linear Regression
 62 (MLR), the best explanatory model of the behavior of the dependent variable 'crude mortality
 63 rate' by the independent variables (markers) was identified. All statistical analyzes were
 64 performed using the IBM-SPSS Statistics software, version 22.0. The processing of the EFA
 65 allowed the identification of 6 variables with a high factorial load, namely: number of deaths
 66 from neoplasia; percentage of people enrolled in the Single Registry without adequate water
 67 supply; NO₂ concentration; homicide mortality rate; percentage of cover by planted forest and
 68 percentage of natural vegetation cover. The best model mathematical found by the RLM
 69 method ($R_{adj}^2 = 0.126$, p-value < 0.001) was $Y = 7.655 + (-0.289 X_1) + (0.132 X_2) + (-0.109$
 70 $X_3)$, in which the variation of the variable dependent (gross mortality rate) is predicted by
 71 environmental variables: X_1 = percentage of natural vegetation cover ($\beta = -0.289$; p-value =
 72 0.000), X_2 = homicide rate ($\beta = 0.132$; p-value = 0.000) and X_3 = percentage of coverage by
 73 planted forest ($\beta = -0.109$; p-value = 0.001). The assumptions of linearity, homogeneity of
 74 error variance (homoscedasticity), independence and normality of residuals and absence of
 75 collinearity were verified and met. The results are able to support, through a mathematical
 76 method, that the environmental variables (natural vegetation cover and planted forest cover)
 77 are inversely related to the gross mortality rate variable and the homicide rate variable is
 78 directly related to this outcome. Thus, public health policies should pay greater attention to
 79 these variables as good indicators of trends in the gross mortality rate, being an efficient way
 80 of monitoring, anticipating and studying public health scenarios.

81 **Keywords:** Environment. Public health. Multivariate Statistics

82

83

84 **Informações de Financiamento:**

85 Esta pesquisa foi financiada pela Coordenação de Aperfeiçoamento de Pessoal de Nível
86 Superior (CAPES) [Bolsa n° 88882.461704, 2019];

87 **1. Introdução**

88 A despeito do modo complexo, a carga de doenças em populações está relacionada às
89 pressões ambientais e exposições a múltiplos fatores de agravos físicos, químicos e/ou
90 biológicos.

91 Estas relações são comprovadas a partir da associação entre a poluição do ar e doenças
92 cardiovasculares ou Infecções Respiratórias Agudas (IRA) (SALDIVA, 2020). O
93 desmatamento e a redução da cobertura vegetal natural também corroboram com doenças que
94 são transmitidas por vetores como dengue, febre amarela e malária (MORENO *et al.*, 2018).

95 As doenças infecciosas e parasitárias também estão associadas ao saneamento básico
96 inadequado, principalmente em áreas de habitações insalubres (ARRUDA e HELLER, 2022).
97 A má qualidade da água para consumo humano está diretamente ligada às doenças de
98 veiculação hídrica como febre tifóide, cólera, hepatite A, amebíase, giardíase, ascaridíase e
99 doenças diarréicas agudas (CASTRO, CRUVINEL e OLIVEIRA, 2020).

100 Ademais, é evidenciada a associação entre a exposição humana às substâncias químicas
101 como agrotóxicos e casos de intoxicações, problemas respiratórios, câncer, doenças mentais,
102 neurológicas, endócrinas e renais (NOGUEIRA, SZWARCOWALD e DAMACENA, 2020).
103 Outra dimensão envolve os ambientes inseguros, violentos e que apresentam elevadas taxas
104 de homicídios que comprometem negativamente a saúde física e emocional das famílias das
105 vítimas (COSTA *et al.*, 2017).

106 Diante da consistente relação entre o meio ambiente e a saúde, essa pesquisa tem como
107 objetivo determinar o melhor modelo matemático de combinação linear de variáveis
108 ambientais que seja capaz de explicar desfecho óbito em saúde pública. Para tanto, foram
109 usadas análises estatísticas multivariadas de dados. Primeiramente, a Análise Fatorial
110 Exploratória (AFE) foi processada para compreender as relações de interdependência dessas
111 variáveis e, a partir do constructo delas, reduzir o conjunto original de variáveis e identificar
112 as marcadoras, que entraram na próxima fase. Em seguida, neste estudo, foi usado o método
113 *stepwise* de Regressão Linear Múltipla para chegar ao objetivo final proposto de identificar a
114 melhor combinação linear de variáveis ambientais (variáveis independentes) que seja capaz de
115 inferir o comportamento da variável desfecho analisada (taxa bruta de mortalidade).

116 Assim, esse trabalho mostra originalidade e relevância ao determinar quais são as
117 variáveis ambientais melhores indicadoras da variação de comportamento da taxa de
118 mortalidade e viabilizar suporte teórico para as tomadas de decisão em políticas públicas

119 voltadas ao amplo objetivo de proteger as condições de vida das populações e garantir
120 ambientes saudáveis.

121

122 **2. Materiais e métodos**

123 **2.1. Delineamento do estudo**

124 Trata-se de um estudo observacional transversal, quantitativo e analítico que, utilizou o
125 método *stepwise* de Regressão Linear Múltipla (RLM) para tratamento dos dados e
126 determinação do melhor modelo matemático preditor do comportamento da variável óbito
127 (LAKATOS e MARCONI, 2022); (HAIR, 2009).

128

129 **2.2. Definição das variáveis independentes (x) e dependentes (y)**

130 As seis variáveis independentes usadas na estimação do melhor modelo de regressão
131 linear múltipla foram definidas a partir de Análise Fatorial Exploratória (AFE) de um
132 conjunto de 130 variáveis originais, do qual foram extraídos 5 fatores. Foram selecionadas as
133 variáveis marcadoras de cada um dos fatores e as duas variáveis com maiores cargas fatoriais
134 do quinto fator (percentual de cobertura por floresta plantada e percentual de cobertura
135 vegetal natural).

136 Desse modo, as variáveis que entraram na RLM foram: (i) número de óbitos por
137 neoplasia; (ii) percentual de pessoas inscritas no Cadastro Único sem abastecimento de água
138 adequado; (iii) concentração de NO₂; (iv) taxa de mortalidade por homicídio; (v) percentual
139 de cobertura por floresta plantada e (vi) percentual de cobertura vegetal natural. A variável
140 dependente (resposta) cujo comportamento está sendo estudado foi a taxa bruta de
141 mortalidade.

142

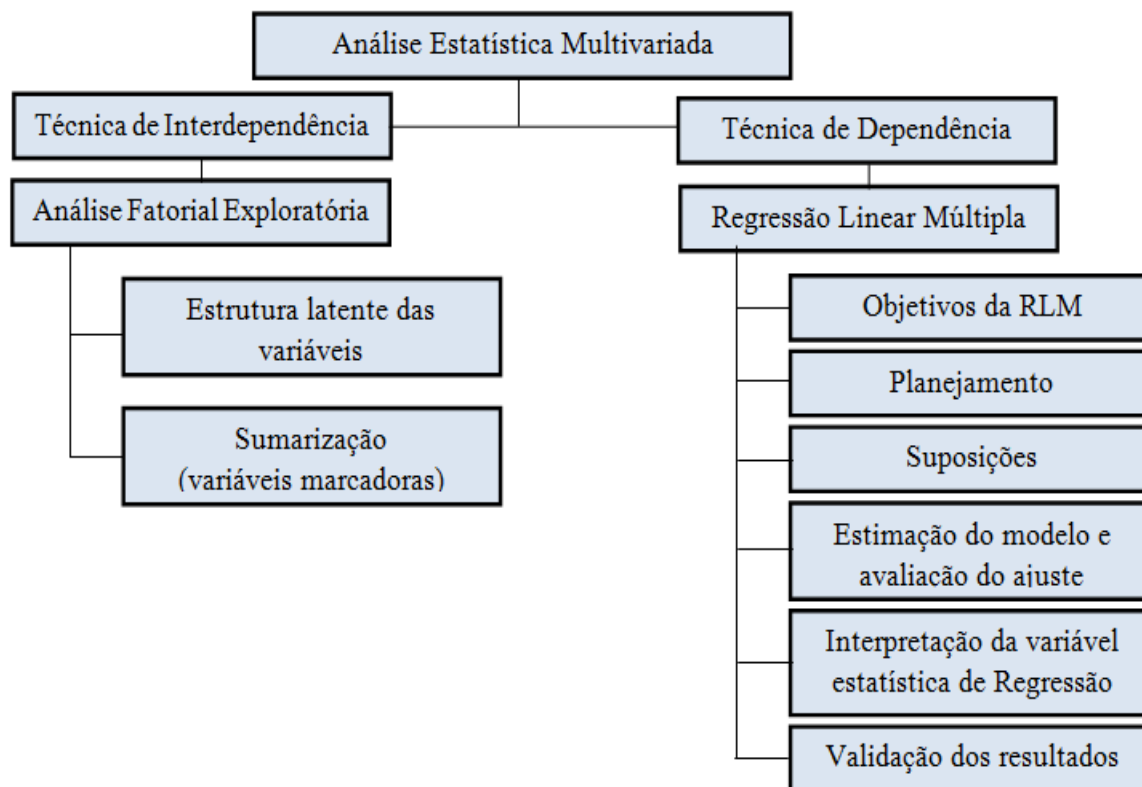
143 **2.3. Análise estatística dos dados**

144 A análise estatística dos dados foi realizada por meio de duas técnicas (Figura 01), uma
145 de interdependência entre as variáveis envolvidas na análise (AFE) e a outra técnica de
146 dependência (RLM) (HAIR, 2009).

147 A sumarização do conjunto original de dados, a estrutura latente e a identificação das
148 variáveis marcadoras foram realizadas preliminarmente por meio da AFE. A partir de então,
149 com as variáveis marcadoras identificadas pela AFE, este estudo usou a Regressão Linear
150 Múltipla (RLM) para determinação da melhor combinação linear destas variáveis que seja
151 capaz de explicar a variável dependente (taxa bruta de mortalidade).

152

153 Figura 01: Análises estatísticas multivariadas usadas



154

155 Organização: LOPES, A. F. A., 2022.

156

157 A realização da RLM, pautou-se em 6 estágios (HAIR, 2009): (A) Objetivos da RLM,
 158 (B) Planejamento, (C) Suposições, (D) Estimação do modelo e avaliação do ajuste, (E)
 159 Interpretação da variável estatística de Regressão e (F) Validação dos resultados.

160

161 ***(A) Definição do objetivo***

162 O objetivo da aplicação da RLM foi encontrar a melhor combinação linear das 6
 163 variáveis obtidas por meio da AFE que seja capaz de explicar com o mínimo de erro o
 164 comportamento da variável dependente desfecho (óbitos).

165

166 ***(B) Planejamento***

167 O tamanho amostral corresponde à N=853 observações, que se referem aos municípios de
 168 Minas Gerais, atende com suficiência ao número de observações requerido pela RLM que, de
 169 modo conservador, é de 10 a 20 observações por variável e todas as variáveis que entraram na
 170 análise são de natureza métrica e medidas em escala de razão (HAIR, 2009).

171

172 ***(C) Suposições em Regressão Múltipla***

173 As suposições de linearidade, homogeneidade de variância de resíduos
174 (homocedasticidade), normalidade multivariada e independência dos termos de erro foram
175 verificadas e atendidas. A ausência de colinearidade foi examinada por meio das estatísticas
176 do fator de inflação de variância (VIF) e teste de tolerância.

177

178 ***(D) Estimação do modelo e avaliação do ajuste***

179 A estimação do modelo foi realizada a partir do método *Stepwise* acrescentando ou
180 retirando variáveis com posterior avaliação da influência sobre os parâmetros de avaliação do
181 modelo de estimação da variável dependente. Como se trata de um modelo de explicação,
182 para a avaliação do ajuste foram verificados (i) a significância (*p valor*) do modelo e dos
183 coeficientes de correlação betas (β), além da (ii) avaliação dos resíduos.

184

185 ***(E) Interpretação da variável estatística de regressão***

186 Para a interpretação do modelo foram usados os parâmetros: coeficiente de determinação
187 ajustado (R_{adj}^2), soma dos quadrados dos resíduos, coeficientes de regressão padronizados
188 (β_i), além da significância do modelo e dos parâmetros supracitados.

189

190 ***(F) Validação dos Resultados***

191 Os resultados foram validados a partir de uma subamostragem da amostra total (30%),
192 e ao comparar os dois modelos foi verificado as mesmas variáveis em ambas as amostras,
193 comprovando-se a validação dos resultados do modelo.

194

195 **3. Resultados**

196 A RLM foi processada com seis variáveis independentes marcadoras (explicativas) que se
197 referem respectivamente: (i) número de óbitos por neoplasia; (ii) percentual de pessoas
198 inscritas no Cadastro Único sem abastecimento de água adequado; (iii) concentração de NO₂;
199 (iv) taxa de mortalidade por homicídio; (v) percentual de cobertura por floresta plantada e (vi)
200 percentual de cobertura vegetal natural. Estas variáveis foram definidas a partir das cargas
201 fatoriais extraídas de um modelo de Análise Fatorial Exploratória (APÊNDICE A). E a
202 variável desfecho analisada refere-se à taxa bruta de mortalidade.

203 A avaliação da existência de correlações lineares entre as variáveis independentes e a
204 variável taxa bruta de mortalidade mostrou três das seis correlações significantes a nível de α

205 < 0,01 sendo, uma correlação significativa com $\alpha = 0,011$ e uma correlação significativa a nível
 206 $\alpha < 0,10$. Apenas uma das correlações não mostrou significância com $\alpha = 0,186$ (Tabela 01).

207

208 Tabela 01: Correlações de Pearson entre as variáveis independentes e Taxa bruta de
 209 mortalidade

	r	p-valor
Percentual de cobertura vegetal	-0,307	0.000**
Cobertura de floresta plantada	-0.184	0.000**
Concentração NO ₂ (ppb)	0.079	0.011*
Percentual de pessoas sem abastecimento de água adequado	-0.050	0.074***
Taxa de mortalidade por homicídio	0.131	0.000**
Óbitos por neoplasia	-0.031	0.186

* significativa a nível $\alpha < 0.05$; ** significativa a nível $\alpha < 0.01$; *** significância a nível $\alpha < 0.10$

210 Fonte: LOPES, A. F. A., 2022.

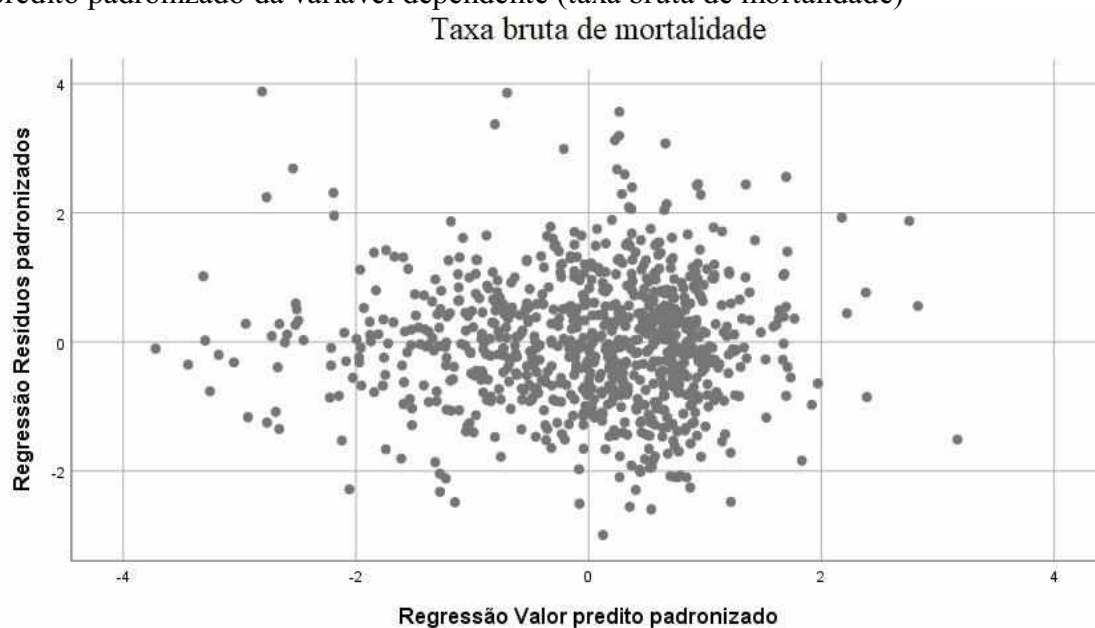
211

212 A partir da matriz de correlação de Pearson podemos identificar as variáveis
 213 independentes que apresentam correlações positivas (dependência direta) e aquelas com
 214 correlações negativas (dependência inversa) com a taxa bruta de mortalidade (HAIR, 2009).

215 A avaliação gráfica da homocedasticidade foi feita por meio da dispersão dos resíduos
 216 padronizados do modelo geral, que envolve todas as seis variáveis independentes em análise
 217 *versus* o valor predito padronizado da variável dependente (Figura 02).

218

219 Figura 02: Gráfico de Dispersão de Resíduos padronizados do modelo geral *versus* o valor
 220 predito padronizado da variável dependente (taxa bruta de mortalidade)



221

222 Fonte: LOPES, A. F. A., 2022.

223 Destarte, a suposição da normalidade da distribuição dos termos de erro foi avaliada por
 224 meio da garantia de linearidade entre as variáveis e a homocedasticidade dos resíduos (HAIR,
 225 2009). A independência dos termos de erro foi verificada a partir do padrão de dispersão
 226 gráfica dos resíduos padronizados e do teste de *Durbin-Watson (DW)* que teve resultado igual
 227 a 2,046.

228 A colinearidade foi verificada pelo teste tolerância e Fator de Inflação de Variância (VIF).
 229 Os resultados mostraram que os valores de tolerância foram iguais e próximos a 1,00 (Tabela
 230 02). Para o Fator de Inflação de Variância (VIF) está diretamente relacionado ao valor de
 231 tolerância ($VIF=1/TOL$), todos os resultados mostraram valores muito inferiores a $VIF < 10$.

232

233 Tabela 02: Avaliação da Colinearidade no melhor modelo obtido pelo método *Stepwise* de
 234 RLM

Modelo		Estatísticas de colinearidade	
		Tolerância	VIF
3	(Constante)		
	percent_cobert_vegetal	0,932	1,073
	tx_mort_homicidio	1,000	1,000
	cobert_florest_plantada	0,932	1,073

235 Fonte: LOPES, A. F. A., 2022.

236

237 Entre todas as variáveis inseridas na análise, a variável que melhor explica a variância da
 238 variável dependente taxa bruta de mortalidade, foi o percentual de cobertura vegetal, que
 239 constituiu o primeiro modelo de regressão ($p\text{-valor} < 0,001$; $SQR = 2052,180$). Entre as
 240 variáveis excluídas do primeiro modelo, a taxa de mortalidade por homicídio foi a segunda
 241 melhor variável explicadora da variável resposta estudada e foi, por isso, incluída no segundo
 242 modelo de regressão ($p\text{-valor} < 0,001$; $SQR = 2282,082$), por fim, a terceira e última variável
 243 com capacidade significativa de aumentar a variância explicada da taxa bruta de mortalidade,
 244 percentual de cobertura de floresta plantada, foi incluída no terceiro modelo ($p\text{-valor} < 0,001$;
 245 $SQR = 1987,038$) (Tabela 3).

246 Tabela 03: Variáveis incluídas e excluídas em cada modelo preditor da Taxa bruta de mortalidade.

Modelo ^a		Coeficientes		SQregressão	SQresíduos	MQresíduos	Stresíduos	p-valor ^b
Variáveis excluídas	Variável incluída	β_{padr}	p-valor					
	percent_cobert_vegetal	-0,317	0,000	229,902	2052,180	2,437	2282,082	0,000
1	Cobertura de floresta plantada	-0,109	0,001					
	Concentração NO ₂ (ppb)	0,021	0,528					
	Percentual de pessoas sem abastecimento de água adequado	0,032	0,344					
	Taxa de mortalidade por homicídio	0,132	0,000					
	Óbitos por neoplasia	-0,046	0,157					
	percent_cobert_vegetal	-0,318	0,000	269,662	2012,419	2,393	2282,082	0,000
	Taxa de mortalidade por homicídio	0,132	0,000					
2	Cobertura de floresta plantada	-0,109	0,001					
	Concentração NO ₂ (ppb)	0,044	0,191					
	Percentual de pessoas sem abastecimento de água adequado	0,032	0,340					
	Óbitos por neoplasia	-0,050	0,121					
	percent_cobert_vegetal	-0,289	0,000					
	Taxa de mortalidade por homicídio	0,132	0,000	295,044	1987,038	2,366	2282,082	0,000
	Cobertura de floresta plantada	-0,109	0,000					
3	Concentração NO ₂ (ppb)	0,034	0,308					
	Percentual de pessoas sem abastecimento de água adequado	0,029	0,382					
	Óbitos por neoplasia	-0,052	0,107					

a. Variável Dependente: taxa bruta de mortalidade

b. Teste ANOVA

247 A capacidade preditiva de cada modelo, avaliada pelo coeficiente de determinação (R
 248 quadrado ajustado) para o modelo 1 foi de 0,100, para o modelo 2 foi de 0,117 e para o
 249 modelo 3 foi de 0,126 (Tabela 04).

250

251 Tabela 04: Capacidade preditiva do modelo

Modelo	R	R quadrado	R quadrado ajustado
1	0,318 ^a	0,101	0,100
2	0,344 ^b	0,119	0,117
3	0,359 ^c	0,129	0,126

a. Preditores: (Constante), percentual de cobertura vegetal

b. Preditores: (Constante), percentual de cobertura vegetal, taxa de mortalidade por homicídio

c. Preditores: (Constante), percentual de cobertura vegetal, taxa de mortalidade por homicídio, percentual de cobertura por floresta plantada

d. Variável Dependente: taxa bruta de mortalidade

252 Fonte: LOPES, A. F. A., 2022.

253

254 Diante destes resultados, observa-se que o terceiro modelo (percentual de cobertura
 255 vegetal, taxa de mortalidade por homicídio e cobertura por floresta plantada) foi o que
 256 apresentou o maior percentual de variância explicada (12,6%).

257

258 **4. Discussão**

259 O tamanho amostral correspondente a 853 observações foi superior ao requerido para
 260 realizar a Regressão Linear Múltipla, uma vez que, o tamanho da amostra ideal leva em
 261 consideração a proporção de dez observações para cada variável inserida na análise (HAIR,
 262 2009) e, com 853 observações temos a possibilidade de incluir até 85 variáveis, conquanto
 263 foram processadas 6 variáveis independentes (explicativas) que representam o conjunto
 264 original de dados nessa análise.

265 A suposição de linearidade foi atendida a partir da existência de correlações lineares
 266 entre as variáveis independentes e a variável dependente (Tabela 01). E através do gráfico de
 267 dispersão dos resíduos padronizados observou-se que os resíduos possuem uma distribuição
 268 homogênea (Figura 02), confirmando a existência de homogeneidade de variância.

269 A suposição da normalidade da distribuição dos termos de erro foi atendida tendo em
 270 vista a existência de linearidade entre as variáveis e a homocedasticidade dos resíduos que
 271 pode ser verificada a partir do padrão de dispersão gráfica dos resíduos padronizados. A
 272 inexistência de autocorrelação entre os termos de erro foi confirmada pelo teste de *Durbin-*

273 *Watson (DW)* com resultado igual a 2,046. Esta estatística de teste está no intervalo [0,4],
274 sendo que para erros não autocorrelacionados *DW* é igual a 2, para autocorrelação positiva
275 *DW* aproxima-se de 0 e para autocorrelação negativa *DW* aproxima-se de 4 (JACOBI, 2020).

276 A ausência de colinearidade que se refere à relação entre duas ou mais
277 (multicolinearidade) variáveis independentes foi comprovada por meio do teste de Tolerância
278 e o Fator de Inflação de Variância (VIF) (Tabela 02). Os valores de tolerância iguais e
279 próximos a 1,00 apontaram a ausência de colinearidade entre as variáveis, uma vez que se o
280 valor de tolerância for $> 0,1$ considera-se que a multicolinearidade não afeta decisivamente os
281 resultados (FREUND e WILSON, 2006). Os valores de VIF elevados indicam alto grau de
282 colinearidade (ou multicolinearidade) entre as variáveis independentes (HAIR, 2009), o que
283 não ocorreu neste estudo, uma vez que todos os resultados mostraram valores muito inferiores
284 a $VIF < 10$.

285 Para estimação da melhor combinação linear das variáveis independentes em estudo
286 capaz de explicar o desfecho taxa bruta de mortalidade foi usada a técnica de Regressão
287 Linear Múltipla do tipo *Stepwise*, que permitiu hierarquizar as variáveis, ou seja, ao
288 acrescentar ou eliminar seletivamente variáveis do modelo foi possível identificar aquelas que
289 mais contribuem para a explicação do desfecho e encontrar o melhor modelo com mínimos
290 quadrados de resíduos.

291 Esta técnica *Stepwise* gerou três modelos de regressão sendo que o primeiro usou a
292 variável independente ‘percentual de cobertura vegetal’, que foi a variável com maior
293 capacidade explicativa significativa da taxa bruta de mortalidade. O segundo modelo envolveu
294 as variáveis ‘percentual de cobertura vegetal’ e ‘taxa de mortalidade por homicídio’ e, por
295 fim, o último modelo usou as seguintes três variáveis ambientais simultaneamente:
296 ‘percentual de cobertura vegetal’, ‘taxa de mortalidade por homicídio’ e ‘cobertura por
297 floresta plantada’.

298 Cada modelo foi avaliado pelo valor de significância, soma de quadrados de regressão
299 (SQregressão), soma de quadrados de resíduos (SQresíduos), média quadrática dos resíduos
300 (MQresíduos) e soma total de resíduos (STresíduos) (Tabela 03). Com base em todos esses
301 parâmetros, o melhor modelo para a explicação do comportamento da variável ‘taxa de
302 mortalidade’ pelas variáveis independentes foi o terceiro modelo, com maior soma dos
303 quadrados de regressão, menor soma dos quadrados do resíduo e menor média quadrática dos
304 resíduos. Todos os três modelos testados tiveram significância estatística ($p\ valor < 0,001$), ou
305 seja, nível de confiança de 99,99%.

306 O terceiro modelo ainda foi o que gerou o maior coeficiente de determinação ($R_{adj}^2 =$
307 0,126, p-valor < 0,001) (Tabela 04), o que significa que este modelo é capaz de explicar
308 12,6% a variação da taxa bruta de mortalidade. Devemos considerar a relevância deste
309 percentual de variância explicada em face da complexidade da variável dependente ‘taxa de
310 mortalidade’. A taxa bruta de mortalidade envolve amplo número de dimensões inter-
311 relacionadas além dos aspectos ambientais. Fatores como genéticos, nutricionais, alimentares,
312 hábitos, recursos públicos e coberturas médico-hospitalares entre diversas outras, influenciam
313 de modo direto ou mediado a mortalidade, constituindo uma rede intrincada de relações de
314 dependência e interdependência dos diversos fatores. Nesse cenário, a habilidade das
315 variáveis ambientais de conseguir explicar 12,6% da variação da taxa de mortalidade usando
316 apenas três variáveis, a despeito de todas as demais dimensões e fatores relacionados à
317 mortalidade, é um resultado muito relevante e denota a importância das variáveis ambientais
318 na tomada de decisões em saúde pública.

319 A partir dos resultados dos coeficientes betas (β) para o melhor modelo foi possível
320 identificar o peso e o sentido da influência das variáveis independentes na variação da
321 variável dependente (taxa bruta de mortalidade), sendo: cobertura vegetal natural ($\beta = -0,289$;
322 $p = 0,000$), taxa de homicídio ($\beta = 0,132$; $p = 0,000$) e cobertura por floresta plantada ($\beta = -0,109$;
323 $p = 0,001$). E a equação de regressão obtida para o melhor modelo é dada por (HAIR, 2009):
324 $Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 \dots + \beta_n X_n$, portanto, o modelo matemático de regressão linear
325 que descreve essa relação é: $Y = 7,655 + (-0,289 X_1) + (0,132 X_2) + (-0,109 X_3)$. Em que β_0 é o
326 intercepto, β_1 é a peso da variável X_1 na determinação dos valores da taxa de mortalidade, β_2 é
327 o peso da variável X_2 e β_3 é o peso da variável X_3 , nessa determinação (HAIR, 2009).

328 As variáveis independentes ‘percentual de cobertura vegetal’ e ‘percentual de cobertura
329 por floresta plantada’ apresentam betas negativos (relação inversa destas variáveis com a
330 variável dependente), ou seja, à medida que se reduz a cobertura vegetal natural e a cobertura
331 por floresta plantada há um aumento da taxa de mortalidade. Esse resultado é absolutamente
332 importante, uma vez que este trabalho partiu de dados numéricos de variáveis ambientais e se
333 pautou em cada fase sempre em tratamentos matemáticos e estatísticos. Todas as conclusões
334 aqui apresentadas foram suportadas por parâmetros matemáticos exatos, livres de
335 subjetividade na interpretação.

336 As áreas verdes (vegetação) contribuem para a captura do material particulado e a
337 minimização de efeitos críticos à saúde causados pela poluição do ar, tais como Doenças
338 Cerebrovasculares (DCV), Doenças Isquêmica do Coração (DIC), Doença Pulmonar

339 Obstrutiva Crônica (DPOC), infecção respiratória aguda e câncer de pulmão (SALDIVA,
340 2020). Ademais, o controle do desmatamento é fundamental para mitigar as doenças causadas
341 por vetores como malária, febre amarela e dengue (MORENO *et al.*, 2018); (LOPES e
342 BERNADINO NETO, 2020).

343 A vegetação natural também é importante para a redução da temperatura, redução de
344 ruídos e aumento da umidade (LOURENÇO *et al.* 2016). Os níveis baixos de umidade do ar
345 facilitam o surgimento de gripes, agravamento de alergias, asma, bronquite crônica e Doença
346 Pulmonar Obstrutiva Crônica (DPOC)(SOLOMON e ROCQUE, 2019). Em conjunto com a
347 poluição do ar, a variação da temperatura e da umidade relativa do ar são os principais
348 motivos para o aumento do risco de internações e mortes provocadas por doenças
349 cardiovasculares e respiratórias no mundo (MORAES, *et al.* 2019).

350 A variável taxa de homicídio, por sua vez, apresenta relação direta com a taxa bruta de
351 mortalidade. A taxa de mortalidade por homicídio é uma variável que reflete direta ou
352 indiretamente outras variáveis de iniquidades sociais, tais como ausência de infraestrutura,
353 desemprego, falta de segurança pública e saneamento básico. Neste aspecto, fica evidente que
354 ambientes que possuem infraestrutura precária, desemprego, insegurança pública e
355 saneamento inadequado tendem estão relacionados com taxas de homicídios elevados.

356 A pesquisa de Sá Neto *et al.* (2020) contribui para essa discussão apontando que a
357 vulnerabilidade social e o homicídio se mostram associados. Assim sendo, as pessoas que
358 vivem nestes ambientes estão não apenas mais susceptíveis a riscos de exposição a efeitos
359 adversos à saúde com maiores chances de se adoecerem e morrerem por algum agravo à
360 saúde, mas também expostas à maior violência com risco de morte por homicídio. As
361 disparidades sociais em renda, escolarização e habitação também são apontadas como
362 aspectos marcantes dessa relação entre fatores ambientais e saúde (SOUZA, *et al.*, 2017).
363 Portanto, podemos ponderar que a variável taxa de mortalidade por homicídio é indicadora
364 (representante) de outras variáveis que envolvem aspectos de vulnerabilidade social.

365

366 **5. Conclusão**

367 A Análise de Regressão Linear Múltipla pelo método *Stepwise* resultou em um modelo
368 matemático estatisticamente significativo com *p-valor* do modelo e dos coeficientes betas
369 ($<0,001$) indicando que as variáveis ambientais (cobertura vegetal natural e cobertura por
370 floresta plantada) estão relacionadas inversamente com a variável taxa bruta de mortalidade e
371 a variável taxa de homicídio se relaciona diretamente com esse desfecho.

372 Portanto, o melhor modelo matemático encontrado pelo método de RLM ($R_{adj}^2 =$
 373 0,126, p-valor < 0,001) foi $Y = 7,655 + (-0,289 X_1) + (0,132 X_2) + (-0,109 X_3)$, no qual, a
 374 variação da variável dependente (taxa bruta de mortalidade) é prevista pelas variáveis
 375 ambientais: $X_1 =$ percentual de cobertura vegetal natural ($\beta = -0,289$; p-valor = 0,000), $X_2 =$
 376 taxa de homicídio ($\beta = 0,132$; p-valor = 0,000) e $X_3 =$ percentual de cobertura por floresta
 377 plantada ($\beta = -0,109$; p-valor = 0,001). Assim sendo, estas três variáveis marcadoras foram
 378 capazes corroborar com a explicação da variância na variável dependente (taxa bruta de
 379 mortalidade).

380 O emprego combinado da Análise Fatorial Exploratória (AFE) foi fundamental para a
 381 seleção das variáveis ambientais, eliminando, a priori, variáveis duvidosas, com correlações
 382 parciais, insignificantes e com pouca capacidade preditiva. Portanto, a AFE contribuiu de
 383 modo importante com modelo de RLM por meio da identificação do menor número de
 384 variáveis que, em conjunto, são capazes de representar todo o conjunto original com mínima
 385 perda de informação.

386 Os resultados encontrados nessa pesquisa são de fundamental relevância para a gestão
 387 pública, pois comprovam a partir de modelo matemático de regressão linear múltipla, livre de
 388 interpretações subjetivas, que as variáveis que expressam nível de vegetação existente estão
 389 inversamente relacionadas à taxa bruta de mortalidade e que a taxa de homicídio reflete
 390 diversos aspectos ambientais de desequilíbrio e fragilidade social. Estas variáveis contribuem
 391 para a explicação do desfecho mortalidade em saúde pública, e são capazes de orientar ações
 392 governamentais em medidas efetivas e políticas públicas em saúde, a fim de minimizar os
 393 riscos e proteger as condições de vida da população.

394

395 **6. Agradecimentos**

396 À Universidade Federal de Uberlândia por meio do Programa de Pós-Graduação em Ciências da
 397 Saúde PPCSA/UFU e à Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES.
 398

399 **7. Referências**

400 ARRUDA, A.; HELLER, L. Acesso à água e esgotos em ocupação urbana na Região Metropolitana de
 401 Belo Horizonte: efeitos na saúde, qualidade de vida e relações de gênero. *Physis: Revista de Saúde*
 402 *Coletiva*, v. 32, n. 2, 2022. ISSN 1809-4481. DOI: <https://doi.org/10.1590/S0103-73312022320204>.
 403 Disponível em: <<https://doi.org/10.1590/S0103-73312022320204>>. Acesso em: 16 de agosto de 2022.
 404 CASTRO, R. S. de, CRUVINEL, V. R. N.; OLIVEIRA, J. L. da M. Correlação entre qualidade da
 405 água e ocorrência de diarreia e hepatite A no Distrito Federal/Brasil. *Saúde em Debate* [online]. 2019,
 406 v. 43, n. spe3, pp. 8-19, ISSN 2358-2898, Epub 13 Jan 2020. DOI: [https://doi.org/10.1590/0103-](https://doi.org/10.1590/0103-11042019S301)
 407 [11042019S301](https://doi.org/10.1590/0103-11042019S301). Disponível em: <<https://doi.org/10.1590/0103-11042019S301>>. Acesso em: 13 de
 408 setembro de 20220.

- 409 COSTA, D. H. et al. Homicídios de jovens: os impactos da perda em famílias de vítimas^{1,2}. **Revista**
410 **de Saúde Coletiva**, 2017, v. 27, n.3, p. 685-70, ISSN 1809-4481. DOI: [https://doi.org/10.1590/S0103-](https://doi.org/10.1590/S0103-73312017000300016)
411 [73312017000300016](https://doi.org/10.1590/S0103-73312017000300016). Disponível em: <<https://www.scielo.br/>>. Acesso em: 16 de agosto de 2022.
- 412 FREUND, R. J.; WILSON, W. J.; SA, P. **Regression analysis – Statistical Modeling of a response**
413 **variable**. Elsevier, Inc., San Diego, 459p, 2006.
- 414 HAIR, J. F. **Análise multivariada de dados** [recurso eletrônico] / tradução Adonai Schlup Sant’Anna.
415 – 6. ed. – Dados eletrônicos. – Porto Alegre: Bookman, 2009.
- 416 JACOBI, L. F. et al. Análise dos resíduos em modelos de regressão não linear ajustados aos dados de
417 crotalaria juncea por diferentes métodos de estimação. **Ciência e Natura**, vol. 42, ed. 4, 2020. DOI:
418 <https://doi.org/10.5902/2179460X40542>. Disponível em:<<https://periodicos.ufsm.br/cienciaenatura>>.
419 Acesso em: 13 de setembro de 2022.
- 420 LAKATOS, M. E. MARCONI, M. A. **Metodologia Científica**. 8 ed. 392p. São Paulo: Atlas, 2022.
- 421 LOPES, A. F. A.; BERNADINO NETO, M. **Health, Environment and RIPSA indicators: A literature**
422 **review**. **Brazilian Journal of Animal and Environmental Research**, v. 3, p. 2610-2626, 2020. DOI:
423 [10.34188/bjaerv3n3-168](https://doi.org/10.34188/bjaerv3n3-168). Disponível em: <<https://brazilianjournals.com/ojs/index.php/BJAER>>.
424 Acesso em: 18 de maio de 2022.
- 425 LOURENÇO, A. L. M. F. et al. Metrôpoles, cobertura vegetal, áreas verdes e saúde. **Estudos**
426 **Avançados**, 2016, v. 30, n. 86, pp. 113-130, ISSN 1806-9592. DOI: [https://doi.org/10.1590/S0103-](https://doi.org/10.1590/S0103-40142016.00100008)
427 [40142016.00100008](https://doi.org/10.1590/S0103-40142016.00100008).
- 428 MORAES, S.L.;ALMENDRA, R.;SANTANA, P.;GALVANI, E. Variáveis meteorológicas e poluição
429 do ar e sua associação com internações respiratórias em crianças: estudo de caso em São Paulo, Brasil.
430 **Cadernos de Saúde Pública**, ISSN 1678-4464 35 n°.7, Rio de Janeiro, Julho 2019. DOI:
431 <http://dx.doi.org/10.1590/0102-311X00101418>. Disponível em: <<http://cadernos.ensp.fiocruz.br/>>.
432 Acesso em: 13 de setembro de 2022.
- 433 MORENO, E. S. et al. Licenciamento ambiental de grandes empreendimentos: quais os limites para
434 avaliação de impactos diretos e indiretos em saúde?. **Ciências Humanas**, 2018, v. 13, n. 3. pp. 519-
435 540, ISSN 2178-2547. DOI: <https://doi.org/10.1590/1981.81222018000300003>. Disponível em:
436 <<https://www.scielo.br/j/bgoeldi>>. Acesso em: 16 de agosto de 2021.
- 437 NOGUEIRA, F.A.M., SZWARCOWALD, C. L. e DAMACENA, G. N. Exposição a agrotóxicos e
438 agravos à saúde em trabalhadores agrícolas: o que revela a literatura?. **Revista Brasileira de Saúde**
439 **Ocupacional** [online]. 2020, v. 45, e36, ISSN 2317-6369, Epub 27 Nov 2020. DOI:
440 <https://doi.org/10.1590/2317-6369000041118>. Disponível em: <[https://doi.org/10.1590/2317-](https://doi.org/10.1590/2317-6369000041118)
441 [6369000041118](https://doi.org/10.1590/2317-6369000041118)>. Acesso em: 13 de setembro de 2022.
- 442 SALDIVA, P.H.N. et al. *Exercising in the urban center: Inflammatory and cardiovascular effects of*
443 *prolonged exercise under air pollution*. **Chemosphere**, 2020 Sep; 254:126817. DOI:
444 [10.1016/j.chemosphere.2020.126817](https://doi.org/10.1016/j.chemosphere.2020.126817). Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/32339794/>>.
445 Acesso em: 16 de junho de 2022.
- 446 SÁ NETO J.A. et al. *Situational diagnosis of mortality among adolescents in conditions of*
447 *vulnerability*. UERJ **Nursing Journal**, Rio de Janeiro, 2020; 28:e39279. DOI:
448 <https://doi.org/10.12957/reuerj.2020.39279>. Disponível em: <<https://www.e-publicacoes.uerj.br/>>.
449 Acesso em: 15 de agosto de 2022.
- 450 SOLOMON C. G, ROCQUE R. C. *Climate Change - A Health Emergency*. **New England Journal of**
451 **Medicine**, 2019 Jan 17; 380(3):209-211. DOI: [10.1056/NEJMp1817067](https://doi.org/10.1056/NEJMp1817067). Disponível em:
452 <https://pubmed.ncbi.nlm.nih.gov/32339794/>. Acesso em: 16 de junho de 2022.
- 453 SOUSA, I. V. et al. Diagnóstico participativo para identificação de problemas de saúde em
454 comunidade em situação de vulnerabilidade social. **Ciência & Saúde Coletiva** [online]. 2017, v. 22,
455 n. 12, pp. 3945-3954, ISSN 1678-4561. DOI: [10.1590/1413812320172](https://doi.org/10.1590/1413812320172). Disponível em:
456 <www.scielo.com.br>. Acesso em: 13 de setembro de 2022.

457 **APÊNDICE A: Matriz de Componente Rotativa (Cargas fatoriais e comunalidades).**

Variável	Componente					Comunalidades
	1	2	3	4	5	
VAR1	0,978					0,965
VAR2	0,967					0,946
VAR3	0,824					0,872
VAR4		-0,831				0,782
VAR7				-0,428		0,472
VAR12	0,927					0,933
VAR14		0,533	-0,453			0,750
VAR16		-0,478				0,708
VAR19	0,973					0,960
VAR20	0,985					0,978
VAR23		-0,499				0,677
VAR28	0,988					0,985
VAR29					-0,837	0,870
VAR31	0,845					0,804
VAR32					-0,730	0,652
VAR33	0,704					0,827
VAR34					0,695	0,762
VAR35					0,708	0,631
VAR46			0,778			0,780
VAR47			0,484			0,782
VAR48			0,898			0,904
VAR49			0,854			0,833
VAR52			-0,609			0,801
VAR57	0,928					0,931
VAR58	0,975					0,964
VAR59	0,974					0,954
VAR60	0,983					0,973
VAR62	0,967					0,946
VAR63	0,933					0,925
VAR78	0,939					0,914
VAR86	0,754					0,752
VAR87	0,986					0,986
VAR88	0,678					0,738
VAR91		0,569				0,471
VAR93		0,896				0,849
VAR94		0,766				0,672
VAR95		0,855				0,836
VAR96		0,575				0,559
VAR98		0,449				0,537
VAR103				0,451		0,662
VAR105				0,866		0,769
VAR106				0,953		0,928
VAR107				0,952		0,926
VAR112	0,692					0,533
VAR118	0,762					0,690
VAR121	0,947					0,943
VAR122	0,993					0,992
VAR123	0,971					0,959
VAR124	0,969					0,959
VAR125	0,981					0,974
VAR126	0,985					0,991
VAR127	0,974					0,981
VAR128	0,938					0,886
VAR129	0,970					0,951
VAR130	0,976					0,979

458 Fonte: LOPES, A. F. A., 2022. "Método de Extração: análise de Componente Principal. Método de
459 Rotação: Varimax com Normalização de Kaiser." a. Rotação convergida em 17 iterações.

5. CONSIDERAÇÕES FINAIS

A partir da revisão de literatura sobre a temática saúde e meio ambiente foi possível realizar o levantamento, seleção e organização de múltiplas variáveis de saúde ambiental pautados no modelo Força Motriz/Pressão/Situação/Exposição/Efeito (FPSEE) recomendado pela Organização Mundial da Saúde e seus supostos fatores, assim como uma exploração das metodologias existentes de organização de múltiplas variáveis.

A combinação da Análise Fatorial Exploratória (AFE) e da Regressão Linear Múltipla (RLM) também foram ferramentas fundamentais de análise estatística multivariada de dados que permitiram criar um modelo matemático baseado em múltiplas variáveis ambientais capaz de estimar mortalidade em saúde pública.

A AFE permitiu explorar o comportamento das variáveis e identificar a sua estrutura latente. Ademais, o conjunto original de variáveis e fatores foi reduzido com uma perda mínima de informação. A fatorabilidade dos dados foi confirmada por três métodos: A) pela matriz de correlações de *Spearman* (ρ) que apresentou 54,73% de correlações lineares significantes entre as variáveis ($\alpha < 0,05$), percentual que aumenta para 59,17% quando consideradas as correlações significantes a nível $\alpha < 0,10$. B) pelo teste de esfericidade de Bartlett que apresentou $p\text{-valor} < 0,001$ garantindo as relações suficientes entre as variáveis e C) pela medida de adequação da amostra de Kaiser-Meyer-Olkin (KMO) com resultado igual a 0,952 indicando um valor muito bom para o emprego da análise fatorial, uma vez que valores próximos de 1 indicam que há correlações fortes entre as variáveis.

A partir da matriz de cargas fatoriais rotacionada (*varimax*) e com base no teste *Scree Plot* e o percentual de variância explicada, foram extraídos 5 fatores que, juntos, explicam 59,78% da variância total dos dados. Estes 5 fatores foram rotulados como F1) socioambiental; F2) vulnerabilidade social; F3) qualidade do ar; F4) mortalidade e F5) agropecuária. E as respectivas variáveis marcadoras foram: número de óbitos por neoplasia; percentual de pessoas inscritas no Cadastro Único sem abastecimento de água adequado; concentração de NO₂; taxa de mortalidade por homicídio e por fim, percentual de cobertura por floresta plantada e percentual de cobertura vegetal natural.

Por sua vez, a RLM do tipo *Stepwise* possibilitou identificar o melhor modelo matemático de combinação linear entre estas variáveis marcadoras sendo capaz de maximizar a previsibilidade da variável desfecho. O melhor modelo matemático encontrado pelo método de RLM ($R_{adj}^2 = 0,126$, $p\text{-valor} < 0,001$) foi $Y = 7,655 + (-0,289 X_1) + (0,132 X_2) + (-0,109 X_3)$, no qual, a variação da variável dependente (taxa bruta de mortalidade) é prevista pelas

variáveis ambientais: X_1 = percentual de cobertura vegetal natural ($\beta = -0,289$; p-valor = 0,000), X_2 = taxa de homicídio ($\beta = 0,132$; p-valor = 0,000) e X_3 = percentual de cobertura por floresta plantada ($\beta = -0,109$; p-valor = 0,001).

Ressalta-se que estas três variáveis ambientais conseguiram explicar 12,6% da variação da taxa bruta de mortalidade. Podemos considerar que este percentual de variância explicada é relevante tendo em vista a complexidade de análise da variável dependente, uma vez que, o estudo da taxa bruta de mortalidade envolve não apenas aspectos ambientais, mas outros múltiplos fatores como genéticos, nutricionais, alimentares, hábitos, recursos públicos, atendimento médico-hospitalares, renda, escolarização, dentre outros fatores que apresentam intrínseca relação de dependência e interdependência com a mortalidade.

Portanto, podemos considerar que foi constatado por meio de métodos matemáticos livres de interpretações subjetivas e seguindo todos os critérios teóricos e estatísticos que as variáveis ambientais (percentual de cobertura vegetal natural e percentual de cobertura por floresta plantada) estão relacionadas inversamente com a variável taxa bruta de mortalidade e a variável taxa de homicídio se relaciona diretamente com esse desfecho sendo indicadora de outras variáveis que envolvem aspectos de fragilidade ambiental.

Assim sendo, diante das evidências encontradas nessa pesquisa pode-se afirmar que a análise simultânea de vários indicadores de saúde ambiental por meio das análises estatísticas multivariadas de dados aumenta a capacidade de explicar desfechos críticos na saúde pública a partir da identificação de variáveis marcadoras e do melhor modelo matemático linear explicativo.

Para tanto, os resultados encontrados nesta tese são ferramentas essenciais para auxiliar o desenvolvimento de estratégias e processos decisórios para a prevenção de riscos em saúde pública e estruturação de ações voltadas principalmente para os grupos de risco. No que se refere às limitações é importante a consolidação de um modelo operacional que gere integração com os dados de todos os sistemas de informações em saúde e com as demais vigilâncias em saúde para facilitar o acesso e análise dos dados.

E para trabalhos futuros sugere-se a criação de uma escala de risco à saúde pública baseada em métodos de Modelagem de Equações Estruturais (MEE), semelhantes ao usado nessa pesquisa (combinação da Análise Fatorial e da Regressão Linear Múltipla). O escore gerado por escalas tem a habilidade de expressar cenários de maior ou menor risco à saúde das populações medida por meio de variáveis ambientais e que também pode contribuir para medidas de intervenção sobre as condições de saúde.

REFERÊNCIAS

- ALENCAR, N. M. *et. al.*, *Environmental health and its influence on quality of life: an integrative review*. **Brazilian Journal of Development**, 6(6), 33093–33105, 2020. Disponível em: <https://brazilianjournals.com>>. Acesso em 05 de setembro de 2022. <https://doi.org/10.34117/bjdv6n6-021>
- ARANTES, K. M.; PEREIRA, B. B. Levantamento, análise e seleção de indicadores ambientais e socioeconômicos como subsídio para o fortalecimento das estratégias de controle da dengue no município de Uberlândia- MG. **JHBS - Journal of Health and Biological Sciences**, v.5, n. 1, 2017. Disponível em: <<https://periodicos.unichristus.edu.br/jhbs/article>>. Acesso em 16 de maio de 2022. <http://dx.doi.org/10.12662/2317-3076jhbs.v5i1.1104.p86-94.2017>.
- ARRUDA, A.; HELLER, L. Acesso à água e esgotos em ocupação urbana na Região Metropolitana de Belo Horizonte: efeitos na saúde, qualidade de vida e relações de gênero. **Physis: Revista de Saúde Coletiva**, v. 32, n. 2, 2022. ISSN 1809-4481. Disponível em: <<https://doi.org/10.1590/S0103-73312022320204>>. Acesso em 16 de agosto de 2022. <https://doi.org/10.1590/S0103-73312022320204>.
- BAKKE, H.A; LEITE, A. S. M; SILVA, L.B. Estatística Multivariada: Aplicação da Análise Fatorial na Engenharia de Produção. **Revista Gestão Industrial**, Paraná, v. 04, n. 04, p. 01-14, 2008. <https://doi.org/10.3895/S1808-04482008000400001>
- BORROR, C. M. *Statistical decision making. The Certified Quality Engineer Handbook*. 3rd ed. Milwaukee, WI: ASQ Quality Press. p. 418–472, 2009.
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Vigilância ambiental em saúde: textos de epidemiologia** / Ministério da Saúde. Secretaria de Vigilância em Saúde. 132 p. Brasília: Ministério da Saúde, 2004. Disponível em: https://bvsm.sau.gov.br/bvs/publicacoes/vigilancia_ambiental_saude_textos_epidemiologia.pdf . Acesso em 20 de abril de 2020.
- _____. Ministério da Saúde. Conselho Nacional de Saúde. **Subsídios para construção da Política Nacional de Saúde Ambiental**. Brasília: Ministério da Saúde, 2007. 56 p.
- _____. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância em Saúde Ambiental e Saúde do Trabalhador. **Saúde ambiental: guia básico para construção de indicadores** / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância em Saúde Ambiental e Saúde do Trabalhador. Brasília: Ministério da Saúde, 2011. Disponível em: https://bvsm.sau.gov.br/bvs/publicacoes/saude_ambiental_guia_basico.pdf. Acesso em 20 de abril de 2020.
- _____. Secretaria de Estado da Saúde de São Paulo. Centro de Vigilância Epidemiológica “Prof. Alexandre Vranjac”. **Caderno de Vigilância Epidemiológica – Vigilância Epidemiológica em Saúde Ambiental**/ Secretaria de Estado da Saúde, Centro de Vigilância Epidemiológica “Prof. Alexandre Vranjac” – Divisão de Doenças ocasionada pelo Meio Ambiente – São Paulo, 2013. 135 p. – (Série A. Normas e Manuais Técnicos). Disponível

em: https://bvsmms.saude.gov.br/bvs/publicacoes/guia_vigilancia_epidemiologica_7ed.pdf. Acesso em 20 de abril de 2020.

_____. **Saúde Brasil 2018 uma análise de situação de saúde e das doenças e agravos crônicos: desafios e perspectivas** / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância de Doenças e Agravos Não Transmissíveis – Brasília: Ministério da Saúde, 2019. Disponível em: https://bvsmms.saude.gov.br/bvs/publicacoes/saude_brasil_2018_analise_situacao_saude_doencas_agravos_cronicos_desafios_perspectivas.pdf. Acesso em: 09 de abril de 2020.

_____. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Vigidesastres**, 2020. Disponível em: <<https://www.saude.gov.br/vigilancia-em-saude/vigilancia-ambiental/vigidesastres>>. Acesso em 27 de junho de 2020.

_____. Ministério da Saúde. Secretaria de Vigilância em Saúde. **Acidentes com Produtos Perigosos**, 2020. Disponível em: <<https://www.saude.gov.br/vigilancia-em-saude/vigilancia-ambiental/vigidesastres/acidentes-com-produtos-quimicos-perigosos#:~:text=A%20Vigil%C3%A2ncia%20em%20Sa%C3%BAde%20Ambiental,bem%20como%2C%20realizar%20a%20vigil%C3%A2ncia>>. Acesso em 27 de junho de 2020.

COSTA *et al.* Processos de saúde-doença: diálogos entre as teorias psicanalítica, cognitivo-comportamental e sistêmica. **Rev. SPAGESP**, Ribeirão Preto, v. 21, n. 2, p. 111-125, dez. 2020. Disponível em <http://pepsic.bvsalud.org/scielo.php?script=sci_arttext&pid=S1677-29702020000200009&lng=pt&nrm=iso>. Acesso em 05 setembro de 2022.

FÁVERO, Luiz; BELFIORE, Patrícia; SILVA, Fabiana Lopes da; CHAN, Betty Lilian. **Análise de dados: modelagem multivariada para tomada de decisões**. 5. ed. Rio de Janeiro: Elsevier, 2009. 646 p

FERREIRA, F. L. **Análise dos indicadores municipais de sustentabilidade ambiental utilizados no grande ABC**. 165 f. Dissertação (Mestrado em Administração) – Universidade Municipal de São Caetano do Sul. São Caetano do Sul, 2011.

FURLETTI, L. A. *et al.* Aplicação de regressão linear múltipla para analisar a relação entre buscas por palavras-chave na internet e casos de COVID-19 no Brasil. **Exacta**, São Paulo, v. 20, n.3, 2022. Disponível em: <https://periodicos.uninove.br/exacta>. Acesso em 28 de setembro de 2022. <https://doi.org/10.5585/exactaep.2022.20401>.

GODOY, R. M. P., ABE, K. C., MIRAGLIA, S. G. E. K.. Correlação entre desfechos na saúde cardiorrespiratória e exposição a poluentes atmosféricos – um estudo na megacidade de São Paulo. **Revista De Ciências Da Saúde**, 33(3), 33–46, 2021. Disponível em: <https://periodicos.furg.br/vittalle/article/view/13390>. Acesso em 05 setembro de 2022. <https://doi.org/10.14295/vittalle.v33i3.13390>

HAIR, J. F. **Análise multivariada de dados** [recurso eletrônico] / tradução Adonai Schlup Sant’Anna. – 6. ed. – Dados eletrônicos. – Porto Alegre: Bookman, 2009.

HONGYU, K. Exploratory Factorial Analysis: theoretical summary, application and interpretation. **E&S.Engineering and Science**, v.7, n. 4, p. 88-103, 2018. Acesso em: 27 de julho de 2022. <https://doi.org/10.18607/ES201877599>.

LINDENAU, J.D.R.; GUIMARAES, L.S.P. Calculating effect size in SPSS. *Clinical Biomedical Research*, 32(3), 2012. Disponível em: <https://seer.ufrgs.br/hcpa>. Acesso em 05 de setembro de 2022.

LOPES, C. V. A.; ALBUQUERQUE, G. S. S. C. Agrotóxicos e seus impactos na saúde humana e ambiental: uma revisão sistemática. *Saúde em Debate*, ISSN 2358-2898, v. 42, n. 117, pp. 518-534, 2018. Disponível em: <<https://www.scielo.br/j/sdeb>>. Acesso em: 17 de maio de 2022. <https://doi.org/10.1590/0103-1104201811714>.

LOPES, A. F. A.; BERNADINO NETO, M. *Health, Environment and RIPSA indicators: A literature review. Brazilian Journal of Animal and Environmental Research*, v. 3, p. 2610-2626, 2020. Disponível em: <<https://brazilianjournals.com/ojs/index.php/BJAER>>. Acesso em: 18 de maio de 2022. <https://doi.org/10.34188/bjaerv3n3-168>.

MANLY, B. J. F. **Métodos estatísticos multivariados**: uma introdução. 3 ed. Porto Alegre: Bookman, 2008. 229 p.

MINGOTI, S. A. **Análise de dados através de métodos de estatística multivariada: uma abordagem aplicada**. Belo Horizonte: Editora UFMG, 2005.

MINISTÉRIO DO MEIO AMBIENTE (MMA). Painel Nacional de Indicadores Ambientais - PNIA 2012. **Painel Nacional de Indicadores Ambientais - Referencial teórico, composição e síntese dos indicadores da versão-piloto**. Ministério do Meio Ambiente. Brasília. 2014. 107 p.

MORAES, M. B. C. **Aula: Introdução à Análise Multivariada**. Adaptado do Material de Aula de Corrar, Paulo e Dias Filho (2007). Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, São Paulo, 2017. RCC0305 – Métodos Quantitativos I.

OLIVEIRA, J.B.; ERVILHA, G.T. Serviços de saneamento básico em Minas Gerais e seus determinantes locais, demográficos e socioeconômicos. *Revista Brasileira de Estudos Regionais e Urbanos (RBERU)*, v.13, n.2, p.243-267, 2019. Disponível em: <https://revistaaber.org.br/rberu/article/view/447>. Acesso em 27 de junho de 2020. <http://orcid.org/0000-0002-0952-3663>

OMS. ORGANIZAÇÃO MUNDIAL DE SAÚDE. *Environmental Health Indicators: Development of a Methodology for the WHO European Region*. Dinamarca, 2001. Disponível em: <https://apps.who.int/iris/handle/10665/108366>. Acesso em 20 de abril de 2020.

_____. Ministério da Saúde. **Mortes devido à poluição aumentam 14% em dez anos no Brasil**. 2019. Disponível em: <http://www.saude.gov.br/noticias/agencia-saude/45500-mortes-devido-a-poluicao-aumentam-14-em-dez-anos-no-brasil>. Acesso em: 25 de setembro de 2019.

_____. *Global air quality guidelines. Particulate matter (PM2.5 and PM10), ozone, nitrogen dioxide, sulfur dioxide and carbon monoxide*. Geneva: World Health Organization; 2021. Disponível em: <https://apps.who.int/iris/bitstream/handle>. Acesso em 27 de setembro de 2022.

_____. *World Health Statistics 2022: Monitoring Health for the SDGs, Sustainable Development Goals*, 2022. Disponível em: <https://www.who.int/data/gho/publications/world-health-statistics>. Acesso em 20 de setembro de 2022.

OMS, *et.al.*; **One Health Joint Plan of Action (2022-2026)**. Working together for the health of humans, animals, plants and the environment. Rome, 2022. <https://doi.org/10.4060/cc2289en>

OPAS. ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE. **Indicadores de Saúde: Elementos Conceituais e Práticos**. Washington, D.C.: OPAS, 2018.

PATINO, M.C.; FERREIRA, J.C. *Internal and External Validity: Can you apply research results to your patients?* **Brazilian Journal of Pulmonology**; v. 44, n.3, p.183-183, São Paulo, 2018. Disponível em: https://www.scielo.br/pdf/jbpneu/v44n3/pt_1806-3713-jbpneu-44-03-00183.pdf. Acesso em: 21 de abril de 2021. <http://dx.doi.org/10.1590/s180637562018000000164>.

RADICCHI, A.L.A.; LEMOS, A. F. **Saúde ambiental**. Belo Horizonte: Nescon/UFMG, Coopmed, 2009. 76p.

RIPSA. REDE INTERAGENCIAL DE INFORMAÇÃO PARA A SAÚDE. Indicadores Básicos para a Saúde no Brasil: Conceitos e Aplicações. Rede Interagencial de Informação para a Saúde - Ripsa. – 2. ed. – Brasília: Organização Pan-Americana da Saúde, 2008. Disponível em: <http://tabnet.datasus.gov.br/tabdata/livroidb/2ed/indicadores.pdf>. Acesso em: 08 de abril de 2020.

REZENDE, M.J.; MORAES, A.V.; PERINI, E.G. **Seara de Asclépio : uma visão diacrônica da medicina** [Recurso eletrônico] – 2. ed. – Goiânia : Editora UFG, 2018.

SALDIVA, P.H.N. et al. *Exercising in the urban center: Inflammatory and cardiovascular effects of prolonged exercise under air pollution*. **Chemosphere**, 2020 Sep; 254:126817. Disponível em: <<https://pubmed.ncbi.nlm.nih.gov/32339794/>>. Acesso em 16 de junho de 2022. <https://doi.org/10.1016/j.chemosphere.2020.126817>.

SOBRAL, A; FREITAS, C.M. Modelo de Organização de Indicadores para Operacionalização dos Determinantes Socioambientais da Saúde. **Saúde soc.**, São Paulo , v. 19, n. 1, p. 35-47, Mar. 2010. Disponível em: www.scielo.br/pdf/sausoc/v19n1/04.pdf. Acesso em: 03 de julho de 2019. <http://dx.doi.org/10.1590/S0104-12902010000100004>.

SOLOMON C. G, ROCQUE R. C. *Climate Change - A Health Emergency*. **New England Journal of Medicine**, 2019 Jan 17; 380(3):209-211. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/30650319/>. Acesso em 16 de junho de 2022. <https://doi.org/10.1056/NEJMp1817067>.

TRIOLA, M.F. **Introdução à Estatística**: atualização da tecnologia / tradução e revisão técnica Ana Maria L. de Farias, Vera Regina L.F. e Flores. Rio de Janeiro: LTC, 2013.

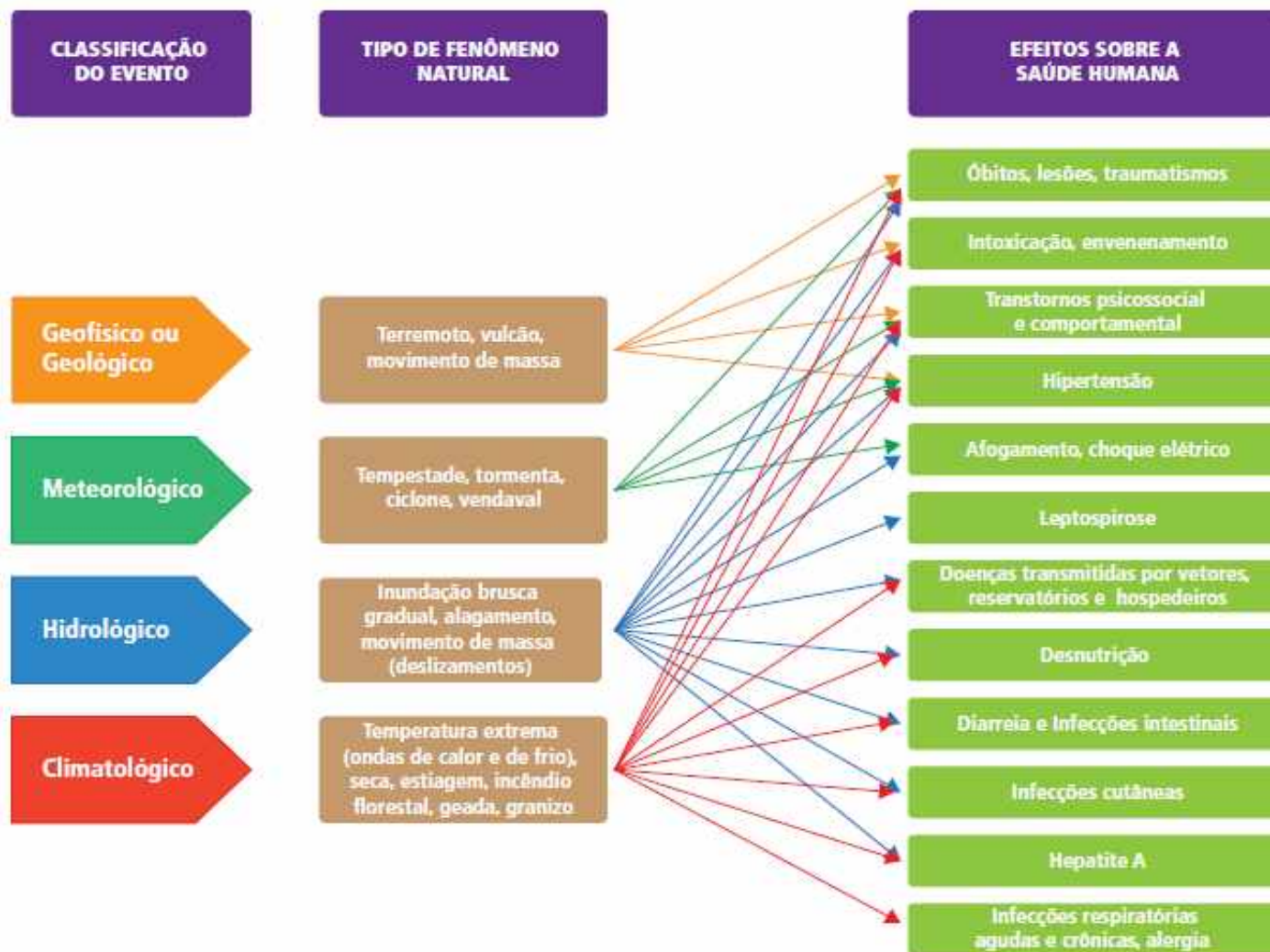
VEIGA, J. E. **Saúde e sustentabilidade**. Estudos Avançados, 2020, v. 34, n. 99, pp. 303-310, ISSN 1806-9592. Disponível em: <<https://www.scielo.br/j/ea/a/xvdzDcvs5mdrcZ7VMCzL6gf/?lang=pt>>. Acesso em: 16 de maio de 2022. <https://doi.org/10.1590/s0103-4014.2020.3499.018>.

VIEIRA, S. **Introdução à Bioestatística**. 4 ed. Elsevier, 2008.

VICINI, L. **Análise multivariada da teoria à prática**. 215f. Monografia (Programa de Pós-Graduação em Engenharia de Produção) – Universidade Federal de Santa Maria: UFSM, Santa Maria, 2005.

ANEXO 01

Anexo 01: Tipos de fenômenos Naturais e seus efeitos sobre a saúde humana



Fonte: OPAS/MS, 2015.