

UNIVERSIDADE FEDERAL DO PAMPA

LUIS FELIPE MACHADO DA SILVA

**ANÁLISE PREDITIVA BASEADA EM INTELIGÊNCIA ARTIFICIAL: UM CAMINHO
PARA A TRANSFORMAÇÃO DO MODELO DE VIGILÂNCIA DAS DOENÇAS
CRÔNICAS NÃO TRANSMISSÍVEIS**

Sant'Anna do Livramento
2023

LUIS FELIPE MACHADO DA SILVA

**ANÁLISE PREDITIVA BASEADA EM INTELIGÊNCIA ARTIFICIAL: UM CAMINHO
PARA A TRANSFORMAÇÃO DO MODELO DE VIGILÂNCIA DAS DOENÇAS
CRÔNICAS NÃO TRANSMISSÍVEIS**

Dissertação apresentada ao Programa de Pós-Graduação em Administração da Fundação Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Administração.

Linha de Pesquisa: Estratégia e Sistemas.

Orientador: Prof. Dr. Rafael Camargo Ferraz

Sant'Anna do Livramento,
2023

Ficha catalográfica elaborada automaticamente com os dados
pelo(a) autor(a) através do Módulo de Biblioteca do
Sistema GURI (Gestão Unificada de Recursos)

D229a Da Silva, Luís Felipe Machado
Análise Preditiva baseada em Inteligência Artificial: Um
caminho para a Transformação do Modelo de Vigilância das
Doenças Crônicas Não Transmissíveis / Luís Felipe Machado Da
Silva.
140 p.

Dissertação (Mestrado)-- Universidade Federal do Pampa,
MESTRADO EM ADMINISTRAÇÃO, 2023.

"Orientação: Rafael Camargo Ferraz".

1. Sistemas de Apoio à Decisão. 2. Tecnologia e Saúde. 3.
Inteligência Artificial. 4. Alkire-Foster. 5. DCNTs. I.
Título.

LUIS FELIPE MACHADO DA SILVA

**ANÁLISE PREDITIVA BASEADA EM INTELIGÊNCIA ARTIFICIAL: UM CAMINHO
PARA A TRANSFORMAÇÃO DO MODELO DE VIGILÂNCIA DAS DOENÇAS
CRÔNICAS NÃO TRANSMISSÍVEIS**

Dissertação apresentada ao Programa de Pós-graduação em Administração da Universidade Federal do Pampa, como requisito parcial para obtenção do Título de Mestre em Administração.

Dissertação defendida e aprovada em: 21 de julho de 2023.

Banca examinadora:

Prof. Dr. Rafael Camargo Ferraz

Orientador

UNIPAMPA

Profa. Dra. Raquel Potter Garcia

UNIPAMPA

Prof. Dr. Paulo Vanderlei Cassanego Junior

UNIPAMPA

Profa. Dra. Sibeles Vasconcelos de Oliveira

UFSM



Assinado eletronicamente por **SIBELE VASCONCELOS DE OLIVEIRA, Usuário Externo**, em 24/08/2023, às 11:54, conforme horário oficial de Brasília, de acordo com as normativas legais aplicáveis.



Assinado eletronicamente por **RAFAEL CAMARGO FERRAZ, PROFESSOR DO MAGISTERIO SUPERIOR**, em 24/08/2023, às 15:01, conforme horário oficial de Brasília, de acordo com as normativas legais aplicáveis.



Assinado eletronicamente por **RAQUEL POTTER GARCIA, Coordenador(a) do Curso de Enfermagem**, em 24/08/2023, às 15:37, conforme horário oficial de Brasília, de acordo com as normativas legais aplicáveis.



Assinado eletronicamente por **PAULO VANDERLEI CASSANEGO JUNIOR, PROFESSOR DO MAGISTERIO SUPERIOR**, em 24/08/2023, às 16:05, conforme horário oficial de Brasília, de acordo com as normativas legais aplicáveis.



A autenticidade deste documento pode ser conferida no site https://sei.unipampa.edu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1188349** e o código CRC **1BA01B58**.

À minha avó Valda e às minhas tias Marucha e Oscila.

(in memoriam).

Essas mulheres incríveis me amaram, cuidaram e protegeram. Sem elas a vida não seria tão especial e as minhas recordações não seriam tão incríveis.

Obrigado por tudo!

AGRADECIMENTOS

A Deus, ele é o meu refúgio, a minha fortaleza e nele eu confiarei.

Aos meus pais, Soares e Iolanda, obrigado por serem minha família e meus maiores fãs. E por serem meu alicerce e a certeza de ter com quem contar.

Ao meu orientador, Professor Dr. Rafael Camargo Ferraz, por toda a paciência em ter me orientado durante esta jornada. Obrigado pela confiança no meu trabalho, pelo respeito, por me ensinar, pela compreensão e pelos sábios conselhos. Fui orientado por um altruísta, que em breve será um grande pai. Foi uma honra e um privilégio que ficarão registrados na minha memória.

Aos professores membros da banca de qualificação – Professor Dr. Paulo Cassanego Junior e Professora Dra. Raquel Potter Garcia – pelos pertinentes apontamentos que engrandeceram este estudo, e aos membros da banca examinadora, Professora Dra. Sibeles Vasconcelos de Oliveira, Professor Dr. Paulo Cassanego Junior e Professora Dra. Raquel Potter Garcia pelo interesse e disponibilidade.

À Professora Dra. Kathiane Benedetti Corso, pelas inovadoras aulas das disciplinas de TIS e Ateliê de Pesquisa, que impulsionaram o meu interesse pelo estudo no campo da administração da informação.

Aos professores do Programa de Pós-graduação em Administração, pelos ensinamentos que certamente contribuíram positivamente para a minha formação. E, aos técnicos em administração, pelo pronto atendimento sempre que solicitado.

Às colegas doutorandas (egressas do PPGA) Marina Valim Bandeira, por sua boa vontade em me ensinar o passo-a-passo do primeiro artigo publicado, e Cristiane Ferreira de Souza Araújo por me ensinar com maestria a arte de elaborar uma Revisão Sistemática de Literatura, sempre com paciência e extrema educação.

Aos meus colegas de trabalho que não pouparam esforços para que eu pudesse me dedicar exclusivamente ao estudo.

Aos meus colegas de mestrado: Daniel, Rosany, Geovana, Gabriel e Caroline Prestes, cujo apoio e amizade estiveram presentes em todos os momentos.

À Universidade Federal do Pampa, pela oportunidade de engrandecer meus conhecimentos.

E desta maneira, Senhor, dou aqui a Vossa Alteza do que nesta vossa terra vi. E, se algum pouco me alonguei, Ela me perdoe, que o desejo que tinha, de Vos tudo dizer, mo fez assim pôr pelo miúdo. Deste Porto Seguro, da vossa Ilha da Vera Cruz, hoje, sexta-feira, 1º dia de maio de 1500.

(Trecho da Carta de Pero Vaz de Caminha a O Rei D Manual, 1º de maio de 1500)

RESUMO

Essa dissertação, apresentada no formato de artigos/capítulos, teve, como objetivo, identificar modelos de predição baseados em Inteligência Artificial (IA) que auxiliem na vigilância das Doenças Crônicas Não Transmissíveis (DCNTs) no Brasil. O Capítulo I apresenta, como objetivo, analisar a qualidade de vida (QV) da população brasileira com o emprego do Método Alkire-Foster aos microdados do Inquérito de Vigilância de Fatores de Risco e Proteção para DCNTs. Foram selecionados onze indicadores distribuídos em três dimensões. A QV é um conceito intensamente marcado pela subjetividade, envolvendo todos os componentes essenciais da condição humana, quer seja físico, psicológico, social, cultural ou espiritual. Sabe-se que estas doenças comprometem muito além do que apenas a esfera biológica, e acabam interferindo de diferentes formas no próprio estilo de vida das pessoas acometidas, do seu grupo familiar e social. Os resultados encontrados indicam uma alta proporção de privação de QV entre os respondentes do inquérito. A QV parece ser fortemente determinada por fatores de ordem socioeconômica, uma vez que a doença crônica interrompe ou dificulta a inserção no processo produtivo, diminuindo as possibilidades de acesso aos bens de consumo. A partir dos resultados obtidos, sugere-se a importância de se considerar a multidimensionalidade do conceito de QV na assistência das pessoas com DCNTs. A partir deste argumento, construído no primeiro artigo, foi proposto um outro estudo teórico, que se encontra no Capítulo II, cujo objetivo é apresentar uma síntese de artigos científicos mundiais publicados sobre modelos preditivos para DCNTs, e analisar como esse tema tem sido abordado na Administração da Informação (ADI), sobre a área de Tecnologia e Saúde. Para isso, foi realizada uma revisão sistemática de literatura (RSL), que buscou por publicações sobre o tema, nas plataformas de busca SCOPUS e WEB of SCIENCE, entre 2012 e 2022. Este tipo de revisão contribui para mapear e, ao mesmo tempo, sintetizar, de forma a documentar as descobertas recentes sobre o tema pesquisado. O conjunto dos 40 artigos apresenta como o tema foi tratado na ADI, mostrando a evolução da temática. A justificativa e relevância para esta pesquisa está na necessidade de compreender como os pesquisadores exploram as diversas aplicações das técnicas de predição na vigilância destas doenças. Observa-se a contemporaneidade do uso da Inteligência Artificial (IA) na saúde, através de um arcabouço de pesquisas sobre o tema. Entre as pesquisas futuras, sugere-se a proposição de uma Agenda de Pesquisa com propostas de temas a serem explorados dentro da área de Sistemas de Informação. Segue-se para o Capítulo III, que apresenta um estudo empírico, utilizando-se modelos de previsões baseados em algoritmos de Aprendizado de Máquina (AM) para predição de risco em Obesidade, Diabetes ou ambas, a partir de variáveis colhidas no inquérito VIGITEL 2021. Buscou-se modelar matematicamente um conjunto de dados de 24.217 mil respondentes do inquérito, a partir de treze variáveis de entrada. Como modelo de saída, realizou-se três testes. Os modelos com melhor desempenho ficaram classificados na seguinte ordem: Obesidade, Obesidade e Diabetes, Diabetes. Concluiu-se, que a técnica de AM pode melhorar a prontidão da saúde, pois permite identificar, digitalmente, os sujeitos passíveis de maior risco de desenvolvimento destas doenças, recomendando intervenção antes de seu agravamento.

Palavras-chave: Sistemas de Apoio à Decisão, Tecnologia e Saúde, Inteligência Artificial, Alkire-Foster, DCNTs.

ABSTRACT

This dissertation, presented in the format of articles/chapters, aimed to identify prediction models based on Artificial Intelligence (AI) that assist in the surveillance of Chronic Noncommunicable Diseases (NCDs) in Brazil. The objective of Chapter I is to analyze the quality of life (QL) of the Brazilian population using the Alkire-Foster Method on microdata from the Surveillance Survey of Risk and Protective Factors for NCDs. Eleven indicators distributed across three dimensions were selected. QoL is a concept intensely marked by subjectivity, involving all essential components of the human condition, whether physical, psychological, social, cultural or spiritual. It is known that these diseases compromise much more than just the biological sphere and end up interfering in different ways with the lifestyle of the people affected, their family and social group. The results found indicate a high proportion of QoL deprivation among survey respondents. QoL appears to be strongly determined by socioeconomic factors, since chronic illness interrupts or hinders integration into the production process, reducing the possibilities of access to consumer goods. Based on the results obtained, the importance of considering the multidimensionality of the concept of QOL in the care of people with NCDs is suggested. Based on this argument, constructed in the first article, another theoretical study was proposed, which is found in Chapter II, whose objective is to present a synthesis of worldwide scientific articles published on predictive models for NCDs, and analyze how this topic has been addressed in Information Administration (ADI), on the area of Technology and Health. For this, a systematic literature review (RSL) was carried out, which searched for publications on the topic, on the SCOPUS and WEB of SCIENCE search platforms, between 2012 and 2022. This type of review contributes to mapping and, at the same time, synthesizing, in order to document recent discoveries on the researched topic. The set of 40 articles presents how the topic was treated at ADI, showing the evolution of the topic. The justification and relevance for this research lies in the need to understand how researchers explore the various applications of prediction techniques in the surveillance of these diseases. The contemporary use of Artificial Intelligence (AI) in healthcare is observed, through a research framework on the topic. Among future research, it is suggested to propose a Research Agenda with proposals for topics to be explored within the area of Information Systems. We move on to Chapter III, which presents an empirical study, using prediction models based on Machine Learning (ML) algorithms to predict risk in Obesity, Diabetes or both, based on variables collected in the VIGITEL 2021 survey. The aim was to mathematically model a data set of 24,217 thousand survey respondents, based on thirteen input variables. As an output model, three tests were carried out. The models with the best performance were classified in the following order: Obesity, Obesity and Diabetes, Diabetes. It was concluded that the AM technique can improve health preparedness, as it allows digitally identifying subjects at greater risk of developing these diseases, recommending intervention before their worsening.

Keywords: Decision Support Systems, Technology and Health, Artificial Intelligence, Alkire-Foster, NCDs.

LISTA DE FIGURAS

Figura 1 – Condução da Revisão Sistemática de Literatura	66
Figura 2 – Distribuição anual do corpus textual em número de artigos.....	70
Figura 3 – Distribuição Anual do Corpus Textual em Percentual	72
Figura 4 – Distribuição dos Autores do Corpus Textual por País.....	73
Figura 5 – Decomposição de autoria do corpus textual	74
Figura 6 – Representação da Elite de Pesquisa do Corpus Textual	75
Figura 7 – Rede de cocitação de autores com 'nós' calculados segundo a quantidade de citações	77
Figura 8 – Rede de documentos acoplados com 'nós' calculados segundo a quantidade de citações	78
Figura 9 – Análise de Palavras. Extraído do software IRAMUTEQ	83
Figura 10 – Análise de Similitude. Extraído do software IRAMUTEQ	84
Figura 11– Rede de agrupamento de palavras	86
Figura 12 – Estrutura do Modelo Multi Layer Perceptron (MLP).	102

LISTA DE QUADROS

Quadro 1 – Passos a serem seguidos	33
Quadro 2 – Dimensões, indicadores e linhas de corte para estimação do IPQV	35
Quadro 3 – Resultados da pesquisa de privação multidimensional das grandes regiões brasileiras para $k \geq 6$ indicadores em 2001 e 2011	47
Quadro 4 – Strings de busca.....	65
Quadro 5 – Proposta de agenda de pesquisa sobre modelos preditivos doenças crônicas não transmissíveis	88
Quadro 6 – Classificação do índice de desempenho	104
Quadro 7 – Variáveis utilizadas no estudo	106

LISTA DE TABELAS

Tabela 1 – Frequência de tabagismo	40
Tabela 2 – Tabagismo – Frequências de tabagismo nas principais capitais dos estados brasileiros e Distrito Federal	40
Tabela 3 – Frequência de sobrepeso e obesidade por idade e sexo	41
Tabela 4 – Percentual* de adultos (≥ 18 anos) com excesso de peso (IMC ≥ 25 kg/m ²), por sexo, segundo as capitais dos estados brasileiros e o Distrito Federal	42
Tabela 5 – Percentual* de adultos (≥ 18 anos) com obesidade (IMC ≥ 30 kg/m ²), por sexo, segundo as capitais dos estados brasileiros e o Distrito Federal	43
Tabela 6 – Maiores e menores frequências de consumo de hortaliças, verduras e frutas por capital e sexo	44
Tabela 7 – Maiores e menores frequências de consumo de refrigerantes por capital e sexo.....	44
Tabela 8 – Maiores e menores frequências de prática de atividade física por capital e sexo.....	45
Tabela 9 – Maiores e menores frequências de consumo de bebidas alcoólicas nos últimos trinta dias por capital e sexo	46
Tabela 10 – Documentos acoplados com 'nós' calculados segundo a quantidade de citações	79
Tabela 11 – Indicadores bibliométricos dos periódicos do corpus textual.....	80
Tabela 12 – Distribuição dos Journals por País	82
Tabela 13 – Diagrama de frequência da amostra VIGITEL ano 2021.....	107
Tabela 14 - Treinamento e validação dos RNA's obtidos para estimar indivíduos com diabetes ou obesidade ou diabetes e obesidade.....	108

SUMÁRIO

CAPÍTULO 1 – INTRODUÇÃO GERAL	12
REFERÊNCIAS CAPÍTULO 1	21
CAPÍTULO 2 – O FARDO ECONÔMICO DAS DOENÇAS CRÔNICAS NÃO TRANSMISSÍVEIS: UMA APLICAÇÃO DO MÉTODO ALKIRE FOSTER (AF) A PARTIR DO INQUÉRITO VIGITEL 2021	24
1 INTRODUÇÃO.....	25
1.1 <i>INSIGHTS</i> SOBRE O CONCEITO DE QUALIDADE DE VIDA.....	26
1.2 CONDICIONANTES DA QUALIDADE DE VIDA: ESFORÇOS ANALÍTICOS E MENSURAÇÃO	29
2 PROCEDIMENTOS METODOLÓGICOS	32
3 ANÁLISE E DISCUSSÃO DOS RESULTADOS	39
4 CONCLUSÃO	51
REFERÊNCIAS CAPÍTULO 2	54
CAPÍTULO 3 – MACHINE LEARNING PARA ANÁLISES PREDITIVAS EM DOENÇAS CRÔNICAS NÃO TRANSMISSÍVEIS: UMA REVISÃO SISTEMÁTICA DE LITERATURA	58
1 INTRODUÇÃO.....	59
1.1 <i>MACHINE LEARNING</i> PARA ANÁLISES PREDITIVAS EM DCNTS.....	60
2 PROCEDIMENTOS METODOLÓGICOS	62
2.1 PRIMEIRA FASE: PLANEJAMENTO DA REVISÃO	63
2.2 SEGUNDA FASE: CONDUÇÃO DA REVISÃO.....	63
2.3 TERCEIRA FASE: DISSEMINAÇÃO DO CONHECIMENTO.....	67
2.3.1 Primeira subfase: análise detalhada.....	67
2.3.2 Segunda subfase: análise temática.....	68
3 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS.....	69
4 CONSIDERAÇÕES FINAIS.....	89
REFERÊNCIAS CAPÍTULO 3	91

CAPÍTULO 4 – MODELAGEM E PREVISÃO DE OBESIDADE E DIABETES: UM ESTUDO A PARTIR DO INQUÉRITO TELEFÔNICO VIGITEL 2021		95
1	INTRODUÇÃO.....	96
2	REFERENCIAL TEÓRICO	98
2.1	RELAÇÃO ENTRE OBESIDADE E DIABETES	98
2.2	USO DE INTELIGÊNCIA ARTIFICIAL.....	100
3	METODOLOGIA.....	101
3.1	CARACTERÍSTICAS DA PESQUISA.....	101
3.2	MODELOS DE PREDIÇÃO E AVALIAÇÃO DE DESEMPENHO PARA OBESIDADE E DIABETES	102
3.3	DADOS SECUNDÁRIOS – BANCO DE DADOS	105
4	RESULTADOS	106
5	CONSIDERAÇÕES FINAIS.....	110
REFERÊNCIAS CAPÍTULO 4		111
CAPÍTULO 5 – CONSIDERAÇÕES FINAIS DO ESTUDO.....		113

CAPÍTULO 1 – INTRODUÇÃO GERAL

Observa-se o crescimento da incorporação de tecnologias na realidade cotidiana das pessoas e do mundo corporativo em geral, o que tem levado a avanços que exigem maior dependência de algoritmos de aprendizado para facilitar a integração de informações em diversos setores da sociedade (TARAPANOFF, 2001). Na área de gestão de informação, as tecnologias associadas à Inteligência Artificial (IA)¹, e seus respectivos subcampos: aprendizado de máquina (*machine learning*) e aprendizado profundo (*deep learning*) desempenham um papel relevante para que haja desenvolvimento em várias aplicações de capacidade analítica, operacional e de tomada de decisões, apresentando embasamento para novas soluções e perspectivas em diferentes segmentos (FEUERRIEGEL et al., 2022).

A IA compreende a simulação de processos de inteligência humana por máquinas, especialmente sistemas computacionais (PEREIRA et al., 2022a), que pode ser útil conforme auxilia no aperfeiçoamento do fluxo de trabalho e de soluções mais eficazes para necessidades existentes (ARAUJO FILHO et al., 2019). A contribuição da IA tem se ampliado paulatinamente na gestão de saúde (seja ela pública ou privada), tendo em vista a gama de aplicações de IA na medicina, em áreas médicas concentradas em práticas clínicas, diagnósticas, reabilitativas, cirúrgicas e preditivas, ou ainda, na tomada de decisões clínicas, no diagnóstico e nos controles (preventivos e preditivos) de doenças (FREITAS et al., 2021), principalmente as Doenças Crônicas Não Transmissíveis (DCNTs).

A informação compreende uma fonte valiosa para o desenvolvimento cotidiano de atividades de forma profissional ou pessoal. É utilizada em tudo o que se faz, se usa, se pensa, e representa um grande volume de conhecimento para as pessoas e empresas. Entretanto, o excesso de informação que é gerado nas diferentes plataformas utilizando Big Data traz uma sobrecarga de informações, dificultando a tomada de decisão.

Percebe-se que há um desafio premente em gerir a informação, e escolher entre o que é ou não relevante, o que é qualificado e pertinente, e que efetivamente

¹ IA: Essa tecnologia parte da premissa de que sistemas podem aprender com os dados para identificar padrões, tomar decisões informadas e melhorar a si mesmos com o mínimo de intervenção humana (TOTVS, 2023).

ajudará na tomada de decisão efetiva. Estendendo-se à área da saúde, a tomada de decisão orientada por dados utiliza a IA, facilitando, entre outras coisas, por exemplo, os processos, diagnósticos e direcionamentos de tratamentos.

Para Netto et al. (2021a), atualmente se assiste à transição do que se chama de *volume care* – a tomada de decisões baseadas em análise de big data – para *value-based care*, que consiste na implementação de tecnologias baseadas em ciência a partir dos dados em saúde e inteligência computacional para criar métodos de avaliação do desempenho e a eficácia dos tratamentos de saúde. Tão difícil como gerir informações de forma confiável, é a expectativa de reunir informações confiáveis e relevantes para a gestão na área da saúde, visando encontrar soluções pertinentes e eficazes, associadas ao uso da IA, principalmente para intensificar as ações voltadas às DCNTs.

O início do Século XX mostrava como principais causas de óbito na população mundial as doenças infecciosas (LEMOS, 2017). Entretanto, avanços nas condições socioeconômicas e de saúde trouxeram novas causas de óbitos (ZIVIANI, 2019), foram substituídas pelas DCNTs, que compreendem as doenças associadas aos sistemas cardiovasculares, ao câncer, ao diabetes e às doenças do aparelho respiratório (ALWAN et al., 2010; BRASIL, 2021a).

As DCNTs são caracterizadas como um conjunto que engloba patologias de diferentes causas e/ou dos fatores que representam risco à saúde, que apresentam uma latência de períodos alongados e um curso prolongado. Sua origem, especificamente, não ocorre de forma infecciosa, e pode levar a resultados de incapacidade funcional (BRASIL, 2008). Essas doenças alcançam um número de indivíduos elevado, pertencentes às mais diferentes classes sociais e econômicas. Sobretudo, atingem mais intensamente as pessoas que se enquadram em grupos de vulnerabilidade social, entre eles os idosos e indivíduos de baixa escolaridade e renda. Tais fatos elevam as condições de privação em que se encontram essas pessoas, que são acometidas por diferentes incapacidades, por limitações, e enfrentam uma redução em sua força produtiva (FIGUEIREDO; CECCON; FIGUEIREDO, 2021; GUIMARÃES, 2017; MALTA et al., 2019).

Evidências mostram que as DCNTs compartilham fatores de risco comuns e, modificáveis, que em sua grande maioria estão relacionadas ao comportamento e estilo de vida, como: tabagismo, alimentação pouco saudável, falta de atividade física;

além do uso abusivo do álcool, que levam ao excesso de peso e à obesidade, aumento da pressão arterial e aumento do colesterol (BRASIL, 2021a).

Entre as DCNTs, uma das mais preocupantes é a obesidade, que provoca a morte de cerca de cinco milhões de indivíduos em cada ano no mundo, principalmente em países com renda *per capita* baixa e média. Com aumento dos riscos de hospitalização e de desenvolvimento de doenças graves, no Brasil, há perspectiva de aumento exponencial da obesidade, com estimativas de que 41% dos adultos sofrerão com essa doença até 2035. As taxas de crescimento da doença na faixa infantil são de 4,4% a.a., e entre os adultos, é de 2,8% a.a. (ATLAS MUNDIAL DA OBESIDADE, 2023).

Incluindo-se a mortalidade, as DCNTs acarretam outros prejuízos à saúde e morbidades relacionadas, como internações, casos de amputações, perda da mobilidade e disfunções neurológicas (BRASIL, 2013). Economicamente, o prejuízo alcança os gastos com saúde pública, absenteísmo² e presenteísmo³ em razão de doenças. No ano de 2020 foram mais de 975.400 óbitos, com aproximadamente 74% destes, decorrentes das DCNTs no Brasil, e 17% dessas mortes se deram de forma prematura, ou seja, atingiram indivíduos que tinham entre 30 e 69 anos de idade (SUPLICI et al., 2021). As mortes acometem principalmente as populações de baixa renda, que não recebem recursos suficientes para tratamentos eficazes, através das políticas públicas no combate à letalidade (FIGUEIREDO et al., 2021).

Os custos econômicos envolvidos nas ações que visam minimizar as consequências das DCNTs oneram os sistemas de saúde e a própria sociedade, representando retrocessos para o desenvolvimento dos países. Envolvem ainda despreparo dos agentes de saúde em relação ao atendimento dessas doenças, falta de programas públicos eficientes de combate às DCNT na educação e deficiências no aprimoramento dos profissionais da área (FIGUEIREDO et al., 2021).

Dados divulgados pelo *Global Burden of Disease* (GBD, 2020) mostram que os gastos governamentais com saúde no Brasil, em 2019, foram de US\$ 639/pessoa, e a previsão para 2050 estimada é de US\$ 1.099/cidadão. Em termos de Produto Interno Bruto (PIB), os gastos totais em saúde (gastos em saúde e assistência ao desenvolvimento da saúde) em 2019 foram equivalentes a 9,3% do PIB, e para 2050, estima-se que atinja 14,8% do PIB (MICAH et al., 2021).

² Absenteísmo - dias de trabalho perdidos devido a doença ou lesão

³ Presenteísmo - redução da produtividade durante o trabalho

Estudos evidenciam que têm declinado as taxas referentes à mortalidade prematura em razão de DCNTs no Brasil (MALTA et al., 2019). Entretanto, a Organização Mundial da Saúde (OMS) destaca que ainda são necessárias ações que envolvam políticas públicas, incentivem o consumo de maior número de alimentos naturais e saudáveis e moderação de itens nocivos - como excesso de sal ou de alimentos industrializados -, incentivo à atividade física em espaços ao ar livre e de acesso ao público em geral, combate ao tabagismo, ao consumo de álcool, entre outros (BRASIL, 2021a; MALTA et al., 2019).

O uso de tecnologias na rede de assistência básica e nos registros sobre as DCNTs e sua evolução, consiste em ferramenta que auxilia no cuidado aos pacientes acometidos por tais doenças. Ainda, a realização de estudos longitudinais, de forma integral, permite a observação de dados tanto qualitativos quanto quantitativos, para elaborar medidas contínuas de rastreamento dos pacientes ao longo do tratamento de DCNTs. Autocuidado, acompanhamento e criação de vínculos junto aos profissionais de atenção em saúde consistem em incentivadores para o curso contínuo e prolongado no tratamento de DCNTs (MALTA et al., 2019).

Na Agenda 2030, que visa ao Desenvolvimento Sustentável e que foi adotada em 2015 pelas Nações Unidas, reconhece-se que as DCNTs são grande desafio de saúde pública, devido aos multi-fatores que as influenciam, principalmente os fatores genéticos e ambientais, dificultando a prevenção e tratamento de maneira eficaz (ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE – OPAS, 2016). Os arquivos médicos e de saúde, que conjugam fatores como estilo de vida, dados clínicos e biológicos em seu rol, são importantes fontes ilimitadas de informações acerca de doenças, excedendo a capacidade humana de entendê-las (ALANAZI, 2022).

Aplicando técnicas de aprendizado de máquina, os problemas nos conjuntos de dados de alta dimensão, podem ser superados. A IA oferece o potencial de analisar conjuntos de dados considerados grandes e complexos a fim de melhorar a medicina preditiva, preventiva e personalizada – Medicina 3P's (ALANAZI, 2022).

Grande parte da população em escala global pode desenvolver DCNTs, o que leva muitos países e seus governos a assumirem compromissos no intuito de realizar monitoramento, enfrentamento e redução dessas doenças e de seus agravos (OPAS, 2016). Na agenda de políticas públicas brasileiras, existem ações de comprometimento federal assumidas até 2030, como o Plano de Ações Estratégicas

para o Enfrentamento das DCNTs, que conduz as ações, bem como políticas públicas, com o objetivo atenuar os fatores agravantes que ocorrem em razão das DCNTs, assumindo um compromisso para atualizar e renovar a pauta escolhida para dez anos (BRASIL, 2021a).

Já a segunda ação compreende a construção do Relatório da Vigilância de Fatores de Risco e Proteção para Doenças Crônicas em Inquérito Telefônico (VIGITEL), que é realizado anualmente através de pesquisa telefônica, em forma de inquérito, que abrange todos as capitais dos Estados brasileiros, incluindo-se o Distrito Federal. Nesse estudo, estima-se com que frequência e como se distribuem os principais fatores de riscos e proteção para cada DCNT (BRASIL, 2021b). Um fato relevante na pesquisa VIGITEL é que o estudo consiste em coletar os dados e descrever a evolução anual dos índices da amostra populacional, como: obesidade, hipertensão, diabetes e mais (BRASIL, 2021b).

Apesar dos esforços governamentais mostrados pelas iniciativas do Plano de Dant e do Relatório VIGITEL, há preocupações relativas à gestão das informações coletadas, de forma integrada. Para adotar soluções com abordagem de prevenção e controle às DCNTs, é preciso focar em aspectos preditivos e preventivos, para que se amplie a possibilidade ao sistema de saúde em realizar a detecção de riscos e demais complicações, ou ainda, óbitos (RAHIMLOO; JAFARIAN, 2016).

Como têm um comportamento silencioso quando acometem as pessoas, as DCNTs não são percebidas ou sentidas (SUPLICI et al., 2021); precisam de tratamentos direcionados ao acompanhamento com equipes de profissionais e diagnóstico precoce. Dependem, ainda, de ações para desenvolver adesão à terapêutica e envolvimento com a promoção da saúde, auxiliando na mitigação das DCNTs. De acordo com Netto et al. (2021a), pesquisas que permitam a estimativa de probabilidade para identificação dos sujeitos passíveis de maior risco de desenvolvimento de DCNTs apresentam condições para auxiliar na prevenção e tratamento das doenças antes de seu agravamento, minimizando, inclusive, os custos sistêmicos⁴ associados aos tratamentos a longo prazo (COSTA; GOUVEIA, 2022).

Nesse contexto, as pesquisas realizadas na área de Tecnologia da Informação têm demonstrado que os novos estudos em *machine learning* (ML) preveem caminhos

⁴ Em razão das possíveis consequências que podem se mostrar irreversíveis, incapacitantes, impactando a vida laboral, o bem-estar do indivíduo e as atividades sociais, decorrem outras consequências como: os efeitos econômicos que afetam as famílias, a saúde pública e a saúde privada (SUPLICI et al., 2021).

futuros para que se utilizem ferramentas para diagnóstico baseados nesta técnica, atuando ao lado de profissionais da saúde, qualificando a tomada de decisão no diagnóstico mais preciso das doenças, além de determinar um caminho de tratamento com maior efetividade. Com relação aos sistemas informatizados, considera-se que sua utilização adequada é particularmente efetiva para detecção das anomalias, e mesmo para detectar doenças raras (COSTA; GOUVEIA, 2022; NETTO et al., 2021a).

De certa forma, percebe-se certa obsolescência do modelo nacional de vigilância das DCNTs. Sabe-se que é possível ampliar as ações a partir de um trabalho interdisciplinar que deve unir não somente profissionais da administração, da computação, da estatística, da matemática, como também os profissionais da saúde. A partir das experiências de cada um nas especialidades a que se dedicam, constrói-se um conhecimento técnico em prol de realizar análises e interpretações de dados adquiridos.

Atualmente, os dados colhidos na rede pública através do Sistema Único de Saúde brasileiro e os dados obtidos pelo Relatório VIGITEL, além de registros locais em hospitais, laboratórios e clínicas da rede pública, não recebem o tratamento de integração e de aplicação de técnicas de IA, o que prejudica a prevenção e predição das DCNTs. Pode-se identificar algumas ações iniciais bem planejadas, de forma estratégica, como e-SUS Atenção Primária (e-SUS APS), envolvendo um sistema de Coleta de Dados Simplificada (CDS) e o Prontuário Eletrônico do Cidadão (PEC), que busca integrar os sistemas já implantados, qualificando a gestão da informação em saúde, no qual será uma ferramenta importante nos estudos e aplicações de soluções (BRASIL, 2023).

Entre os problemas identificados, demonstra-se que o modelo de vigilância atual utilizado pelo governo brasileiro, é incapaz de identificar quais são os pacientes que possuem maior probabilidade de desenvolver DCNT, e de que forma realizar uma intervenção efetiva antes desses problemas se agravarem.

Um segundo aspecto, é que a própria administração pública, apesar de deter inúmeras bases de dados, parece ignorar a riqueza contida nos sistemas de informação da saúde, em especial oriundo da coleta de dados dos atendimentos realizados em Unidades Básicas de Saúde (UBS) (PEREIRA et al., 2022b). Portanto, tendo em vista o exposto, levantou-se o seguinte problema de pesquisa para este

estudo: como os modelos de predição baseados em inteligência artificial podem auxiliar na vigilância das Doenças Crônicas Não Transmissíveis no Brasil?

Nessa dissertação, o objetivo geral é identificar modelos de predição baseados em inteligência artificial (IA) que auxiliem na vigilância das Doenças Crônicas Não Transmissíveis (DCNTs) no Brasil. Como objetivos específicos, pode-se destacar:

- a) Identificar e sintetizar as evidências relevantes disponíveis sobre estudos de predição em DCNTs, através da revisão sistemática da literatura;
- b) Definir as variáveis que podem contribuir na definição de modelos de predição;
- c) Avaliar o desempenho de modelos de inteligência artificial com capacidade de predição de obesidade e diabetes no Brasil;
- d) Analisar a qualidade de vida da população brasileira com o emprego do Método Alkire-Foster⁵ aos microdados do Inquérito de Vigilância de Fatores de Risco e Proteção para DCNTs.

Justifica-se o estudo, pois em termos acadêmicos, o tema de utilização de IA e *machine learning* para predição de DCNTs é relevante, e busca investigar as vantagens da adoção de técnicas modernas de estatística, visando à extração de conhecimento útil, completude de informações e de seu potencial na utilização para pesquisas científicas, assim, proporcionando uma melhor tomada de decisão, baseada mais em dados, e menos em intuição (NETTO et al., 2021a). Além disso, colabora para o debate da ciência à medida em que provoca a reflexão sobre questões de permeiam a formação futura dos profissionais da saúde.

Em termos empíricos, cria-se uma solução alternativa, que permite incluir, de forma preditiva, um volume valioso de informações desconhecidas das bases de dados de saúde e, a partir disso, extrair conhecimento ao descobrir relações ocultas, padrões e, então, estabelecer regras para predizer e correlacionar dados que podem ajudar as organizações nas tomadas de decisões mais rápidas ou, até mesmo, a tomar decisões com maior confiança (OLIVEIRA et al., 2019). Reforça-se a concepção de Ziviani (2019) para quem o próprio governo federal não usa adequadamente a

⁵ A estimativa de um índice de qualidade de vida pelo método Alkire-Foster está inserida num conjunto de ações de vigilância das DCNTs, que possibilitam conhecer a distribuição, magnitude e tendência dessas doenças e de seus fatores de risco na população, identificando seus condicionantes sociais, econômicos e ambientais, com o objetivo de subsidiar o planejamento, execução e avaliação da prevenção e controle das DCNTs. O resultado desta estimativa permite aperfeiçoar os algoritmos preditivos, o que ajuda na promoção de qualidade de vida, prevenção e controle das DCNTs (MALTA et al., 2019).

riqueza de dados que possui em suas bases, gerindo “silos isolados de informação” que não têm uso integrado na base de dados.

No contexto social, o tema pode ter um papel importante na efetividade das políticas públicas, especialmente no caso de superar dificuldades de abordagens mais tradicionais, como no estabelecimento de estudos quase experimentais e por meio da realização de análises preditivas baseada em IA para a alocação prioritária de recursos. Em termos de Políticas Públicas, o estudo está em consonância com as metas definidas no Plano Enfrentamento DCNT 2021-2030 elaborado pelo Ministério da Saúde do Brasil, bem como a Estratégia de Saúde Digital para o Brasil 2020-2028 (BRASIL, 2021b, p. 66).

Em termos econômicos, as DCNT's representam gastos onerosos aos governos. Em acordo ao relatório divulgado pela Organização Mundial da Saúde (OMS, 2018), a cada US\$ 1 investido para ampliar as ações de tratamento das DCNTs nos países que apresentam baixa e média renda, poderá ser encontrado, como retorno ao investimento, ao menos US\$ 7, revertidos em prol da sociedade em forma de aumento do número de empregos, maior produtividade e longevidade.

Nilson et al. (2020) destacam os custos totais no Brasil referentes à hipertensão, diabetes e obesidade no SUS, que alcançaram R\$ 3,45 bilhões no ano de 2018, considerando-se gastos com hospitalizações, com procedimentos ambulatoriais, e, ainda, com medicamentos. Explica-se que 59% desses custos foram relacionados para tratamento da hipertensão, outros 30% são pertinentes ao tratamento de diabetes e 11% à obesidade. Ainda, em 2018, ocorreram, no SUS, 1.829.779 internações em decorrência de causas associadas às doenças como hipertensão arterial, diabetes e obesidade, correspondendo a 16% do total do período das internações hospitalares, ocorridas no SUS.

Segundo o Atlas Mundial da Obesidade (2023), somente o impacto de indivíduos com sobrepeso sobre o valor do PIB brasileiro tende a 3% até 2035, o que é demasiado representativo. Soma-se a isso, um histórico de gastos relacionados às pessoas acometidas com sobrepeso, em assistência médica, que passou de R\$ 64,3 milhões no ano de 2020, estimando-se à perspectiva de R\$ 100 milhões no ano de 2035. Estima-se, ainda, uma elevação mundial da obesidade infantil levará mais de 400 milhões de crianças a conviver com a obesidade e suas consequências em 2035, o que afeta a probabilidade de incremento das DCNTs.

Para Amaral (2016), como país no mundo que tem população de mais de 100 milhões de habitantes, e único a contar com sistema público de saúde, o Brasil deve considerar qualquer solução que otimize recursos e resultados para fins de estudo. Além disso, compreender a dinâmica e os critérios que conduzem à tomada de decisão dos gestores públicos é algo de suma importância, pois permitirá que se identifiquem os conceitos que apoiam o processo decisório, segundo as análises de possíveis lacunas, excessos e avanços.

Desta forma, a pesquisa norteia-se no conceito de análise preditiva baseada em IA, que utiliza de dados de pacientes para identificar aqueles que possuem maior probabilidade de desenvolver DCNTs, e, assim, intervir anteriormente ao agravamento dos problemas.

Após a apresentação e introdução ao tema, bem como a problemática, os objetivos e a justificativas no Capítulo I, esta dissertação foi elaborada no formato de artigos, apresentados nos Capítulos II, III e IV, integrados para o desenvolvimento da presente pesquisa.

O primeiro estudo, no Capítulo II, teve como objetivo analisar a qualidade de vida (QV) da população brasileira, com o emprego do Método Alkire-Foster (AF) aos microdados do Inquérito da Vigilância de Fatores de Risco e Proteção para Doenças Crônicas (VIGITEL) 2021. No segundo estudo, no Capítulo III, o objetivo era identificar o conhecimento produzido sobre o uso de modelos de predição para DCNTs, verificando a elite do *corpus* textual, distribuída segundo a reputação de periódicos em nove países. No Capítulo IV, buscou-se avaliar o desempenho de modelos de redes neurais artificiais com capacidade de predição de obesidade e diabetes no Brasil a partir de variáveis colhidas no inquérito VIGITEL 2021.

Ao final, são apresentadas as considerações finais do presente estudo.

REFERÊNCIAS CAPÍTULO 1

ALANAZI, R. Identification and prediction of chronic diseases using machine learning approach. **Journal Of Healthcare Engineering**, v. 25, Feb, 2826127, 2022.

ALWAN, A.; MACLEAN, D.R.; RILEY, L.M.; D'ESPAIGNET, E.T.; MATHERS, C.D.; STEVENS, G.A.; BETTCHER, D. Monitoring and surveillance of chronic non-communicable diseases: progress and capacity in high-burden countries. **Lancet**, v.27, n.376 (9755), p.1861- 1868, 2010.

AMARAL, F. **Introdução à ciência de dados: mineração de dados e Big Data**. São Paulo: Alta Books Editora, 2016.

ARAUJO-FILHO, J. A. B.; PINTO, F. M. I.; NOMURA, C. H. Inteligência artificial e diagnóstico por imagem - o futuro chegou? **Revista da Sociedade de Cardiologia do Estado de São Paulo (SOCESP)**, v. 29, p. 346-349, 2019.

ATLAS MUNDIAL DA OBESIDADE (2023). Disponível em: <https://pt.worldobesityday.org/resources/entry/world-obesity-atlas-2023>

BRASIL (2008). Ministério Da Saúde (MS). **Diretrizes e recomendações para o cuidado integral de doenças crônicas não-transmissíveis**: promoção da saúde, vigilância, prevenção e assistência. Brasília: MS; 2008.

BRASIL (2013). Ministério Da Saúde (MS). **Diretrizes para o cuidado das pessoas com doenças crônicas nas redes de atenção à saúde e nas linhas de cuidado prioritárias** /Ministério Da Saúde, Secretaria De Atenção À Saúde, Departamento de Atenção Básica. Brasília: Ministério Da Saúde, 2013.

BRASIL (2021a). Ministério Da Saúde (MS). **Plano de ações estratégicas para o enfrentamento das doenças crônicas e agravos não transmissíveis no Brasil 2021-2030** [Recurso Eletrônico] / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Análise em Saúde e Vigilância de Doenças Não Transmissíveis. Brasília: Ministério da Saúde, 2021.

BRASIL (2021b). Ministério da Saúde (MS). **Vigitel Brasil 2021**: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico: estimativas sobre frequência e distribuição sociodemográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal em 2021. Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Análise em Saúde e Vigilância de Doenças Não Transmissíveis. Brasília: Ministério da Saúde, 2021.

BRASIL (2023). **Estratégia e-SUS atenção primária. Em busca de um SUS eletrônico**. Disponível em: <https://sisaps.saude.gov.br/esus/>. Acesso em: 14.mai.2023.

COSTA, O.; GOUVEIA, L. B. Uma proposta para um sistema inteligente de previsão do risco de doenças crônicas. “no prelo”. **XIX Congresso Brasileiro de Informática em Saúde – CBIS**. Campinas/SP, Brasil. 29/11 A 02/12 De 2022 - Campinas/SP – Brasil, 2022.

FEUERRIEGEL, S.; SHRESTHA, Y.R.; VON KROGH, G.; ZHANG, C. E. Bringing artificial intelligence to business management. **Nature Machine Intelligence**, v.4, p. 611–613, 2022.

FIGUEIREDO, A. E. B.; CECCON, R. F.; FIGUEIREDO, J. H.C. Doenças crônicas não transmissíveis e suas implicações na vida de idosos dependentes. **Ciência & Saúde Coletiva**, v.26, n.1, p.77– 88, 2021.

FREITAS, A.L.S.; IEKER, A.S.D; TEIXEIRA, H.M.P.; PINHEIRO, J.M.; RINALDI, W. Aprendizado de máquina aplicado à predição de doenças cardiometabólicas com utilização de indicadores metabólicos e comportamentais de risco à saúde. **XII Computer on the Beach**, 7 a 9 de Abril de 2021, Online, SC, Brasil.

GBD – GLOBAL BURDEN OF DISEASE (2020). Disponível em: <https://www.healthdata.org/gbd>. Acesso em 18.abr.2023.

GUIMARÃES, M.C.S. Dados abertos: uma perspectiva a partir dos estudos sociais da ciência. In: Encontros pré-confoa: dados abertos, ciência de dados aplicada à saúde, **CIDACS**, v.5, Rio de Janeiro, 2017, 17 p.

LEMOS, I. G. S. (2017). **Saúde cardíaca dos jovens universitários: análise da percepção de risco**. Coord. Hayasida, N.M.A. Manaus: EDUA, 2017.

MALTA, D. C.; ANDRADE, S. S. C. A.; OLIVEIRA, T. P.; MOURA, L.; PRADO, R. R.; SOUZA, M. F. M. (2019). Probabilidade de morte prematura por doenças crônicas não transmissíveis, brasil e regiões: projeções para 2025. **Revista Brasileira de Epidemiologia**, v. 22, p. e190030.

MICAH, A.E.; COGSWELL, I. E.; CUNNINGHAM, B. et al. Tracking development assistance for health and for COVID-19: a review of development assistance, government, out-of-pocket, and other private spending on health for 204 countries and territories, 1990–2050. **Lancet**, v.398, p. 1317-1343, 2021.

NETTO, A.V.; BERTON, L.; TAKAHATA, A.K. **Ciência de dados e a inteligência artificial na área da saúde**. Rio de Janeiro: Editora dos Editores, 2021a.

NILSON, E. A. F., ANDRADE, R. C. S., BRITO, D. A.; OLIVEIRA, M. L. Custos atribuíveis à obesidade, hipertensão e diabetes no sistema único de saúde, Brasil, 2018. **Revista Panamericana De Salud Publica**, v.44, n. 32, p. 1-7, 2020.

OLIVEIRA, R. R.; OLIVEIRA, R. R.; ZIVIANI, F. Processo de inovação no contexto de transformação digital: framework para gestão da inovação. **Anais do Congresso Internacional de Conhecimento e Inovação – CIKI**, v.1, n. 1, 2019.

OMS. Organização Mundial da Saúde. Relatório: *Saving lives, spending less: A strategic response to noncommunicable diseases*. **World Health Organization**, 2018. <https://iris.who.int/bitstream/handle/10665/272534/WHO-NMH-NVI-18.8-eng.pdf>

OPAS – ORGANIZAÇÃO PAN-AMERICANA DA SAÚDE. **fatores de risco para doenças crônicas não transmissíveis nas américas: considerações sobre o fortalecimento da capacidade regulatória**. Documento de Referência Técnica REGULA. Washington, DC; OPAS, 2016.

PEREIRA, A. L.; REZENDE, S. O.; MARCACINI, R.; CALDEIRA, D. M. Ciência de dados na administração pública: desafios e oportunidades. **Revista da CGU**, [s. l.], v. 14, n. 26, 2022a.

PEREIRA, K. S. S.; MELO, D. R. A.; VILELA JUNIOR, D. C.; RODRIGUES, L. G. (2022). Fatores que influenciam a aceitação de tecnologias de inteligência artificial na saúde. **Revista Gestão & Saúde**, [s. l.], v. 13, n. 01, p. 02–20, 2022b.

RAHIMLOO, P.; JAFARIAN, A. Prediction of diabetes by using artificial neural network logistic regression statistical model and combination of them. **Bulletin De La Société Royale des Sciences de Liège** [en ligne], v. 85, 2016, Actes De Colloques, Special Edition, p. 1148 - 1164.

SUPLICI, S. E. R.; MEIRELLES, B. H. S.; SILVA, D. M. G. V.; BOELL, J. E. W. Adesão ao autocuidado de pessoas com Diabetes Mellitus na atenção primária: estudo de método misto. **Escola Anna Nery**, v. 25, n.5, p. e20210032, 2021.

TARAPANOFF, Kira (org.). **Inteligência organizacional e competitiva**. Brasília: Ed. da Universidade de Brasília, 2001. 343 p.

ZIVIANI, A. Desafios de ciência de dados aplicada à saúde digital. **Reflexões Sobre Ensino E Pesquisa No SUS**, v.2, p. 1877, 2019.

CAPÍTULO 2 – O FARDO ECONÔMICO DAS DOENÇAS CRÔNICAS NÃO TRANSMISSÍVEIS: UMA APLICAÇÃO DO MÉTODO ALKIRE FOSTER (AF) A PARTIR DO INQUÉRITO VIGITEL 2021

RESUMO

As Doenças Crônicas Não Transmissíveis (DCNTs) representam um desafio crescente para os sistemas de saúde em todo o mundo, com impacto significativo para a saúde da população e para a economia. Este artigo teve como objetivo analisar a qualidade de vida (QV) da população brasileira, com o emprego do Método Alkire-Foster (AF) aos microdados do Inquérito da Vigilância de Fatores de Risco e Proteção para Doenças Crônicas (VIGITEL) 2021. Os objetivos específicos procuraram examinar a relação entre as DCNTs e o status socioeconômico das pessoas, considerando indicadores como renda, escolaridade e acesso aos serviços de saúde; e estimar um índice de qualidade de vida para a população respondente do VIGITEL 2021. A importância deste estudo reside na necessidade de informar políticas públicas mais eficazes e sustentáveis para lidar com as DCNTs. Assim, este estudo fornece subsídios para a formulação de estratégias direcionadas, visando mitigar os impactos econômicos e melhorar a saúde da população brasileira. Em média, 36,21% dos multidimensionalmente privados apresentam algum grau de privação, sendo que, dos onze indicadores avaliados, em média, os indivíduos são privados em oito deles. E, do total de indicadores em que os multidimensionalmente privados sofrem alguma privação, esse índice representa 27,24% do máximo possível na amostra em estudo.

Palavras-chave: Doenças Crônicas Não Transmissíveis, Fardo Econômico, Método Alkire-Foster.

THE ECONOMIC BURDEN OF NON-COMMUNICABLE DISEASES: AN APPLICATION OF THE ALKIRE FOSTER METHOD (AF) FROM THE VIGITEL 2021 SURVEY

ABSTRACT

Chronic Noncommunicable Diseases (NCDs) represent a growing challenge for health systems around the world, with a significant impact on the health of the population and the economy. This article aimed to analyze the quality of life (QOL) of the Brazilian population, using the Alkire-Foster Method (AF) on microdata from the Surveillance Survey of Risk and Protective Factors for Chronic Diseases (VIGITEL) 2021. The objectives Specific studies sought to examine the relationship between NCDs and people's socioeconomic status, considering indicators such as income, education and access to health services; and estimate a quality-of-life index for the VIGITEL 2021 respondent population. The importance of this study lies in the need to inform more effective and sustainable public policies to deal with NCDs. Thus, this study provides support for the formulation of targeted strategies, aiming to mitigate economic impacts and improve the health of the Brazilian population. On average, 36.21% of multidimensionally deprived people present some degree of deprivation, and, of the eleven indicators evaluated, on average, individuals are deprived in eight of them. And, of the total indicators in which the multidimensionally deprived suffer some deprivation, this index represents 27.24% of the maximum possible in the sample under study.

Keywords: Noncommunicable Diseases, Economic Burden, Alkire-Foster Method.

1 INTRODUÇÃO

As Doenças Crônicas Não Transmissíveis (DCNTs) consistem em um grupo de condições de saúde, que têm longa duração e, geralmente, progressão lenta, caracterizadas por apresentarem causas complexas e múltiplos fatores de risco, e podem ter efeitos debilitantes ao longo da vida (WORLD HEALTH ORGANIZATION, 2018). Essas doenças são responsáveis por um grande impacto na saúde da população mundial e incluem condições como doenças cardiovasculares, diabetes, câncer e doenças respiratórias crônicas. Ao contrário das doenças infecciosas, as DCNTs não são transmitidas de pessoa para pessoa, mas geralmente resultam da combinação de fatores genéticos, ambientais e comportamentais (WORLD HEALTH ORGANIZATION, 2018).

Muitas vezes, as DCNTs são progressivas, ou seja, pioram gradualmente com o tempo, resultando em danos irreversíveis aos órgãos afetados. Os principais fatores de risco para o desenvolvimento de DCNTs incluem o tabagismo, consumo excessivo de álcool, alimentação não saudável, falta de atividade física, exposição a poluentes ambientais e histórico familiar da doença (WORLD HEALTH ORGANIZATION, 2018). Esses fatores de risco são amplamente influenciados pelo estilo de vida e pelas condições socioeconômicas em que as pessoas vivem. Portanto, as DCNTs estão intimamente ligadas às desigualdades sociais, econômicas e ambientais, afetando de forma desproporcional as populações mais vulneráveis (WORLD HEALTH ORGANIZATION, 2018).

Tendo em vista o exposto, infere-se que o impacto das DCNTs na saúde pública e nos sistemas de saúde é significativo. Essas doenças são responsáveis por parcela substancial da carga global de doenças, contribuindo para a morbidade e mortalidade prematura. Além disso, as DCNTs impõem grande ônus econômico, resultando em altos custos de tratamento médico, perda de produtividade no trabalho e impactos financeiros nas famílias e na sociedade como um todo (WORLD HEALTH ORGANIZATION, 2018; BLOOM et al., 2018). A prevenção e o controle das DCNTs são, portanto, desafios importantes para os formuladores de políticas de saúde.

De acordo com a Global Burden of Disease Study, em 2019, as DCNTs foram responsáveis por aproximadamente 71% de todas as mortes no mundo e por cerca de 85% da carga de doença ajustada por incapacidade (ROTH et al., 2020). Além

disso, as DCNTs impõem carga financeira significativa sobre os sistemas de saúde e a economia como um todo. Estima-se que os custos diretos de tratamento médico das DCNTs sejam extremamente elevados, alcançando trilhões de dólares anualmente (BLOOM et al., 2018).

Diante desse cenário, é fundamental compreender o fardo econômico das DCNTs em nível nacional e desenvolver estratégias eficazes para enfrentamento do problema. Nesse contexto, o Método Alkire-Foster (AF) tem sido aplicado como uma abordagem para mensurar privações, permitindo uma análise mais abrangente das condições de vida das pessoas (ALKIRE; FOSTER, 2011). A utilização desse método pode oferecer *insights* valiosos sobre as múltiplas dimensões da privação em qualidade de vida que a população brasileira está submetida face à ocorrência das DCNTs e seu impacto econômico.

Desta forma, concentra-se como objetivo geral deste estudo, analisar a qualidade de vida (QV) da população brasileira, com o emprego do Método Alkire-Foster (AF) aos microdados do Inquérito da Vigilância de Fatores de Risco e Proteção para Doenças Crônicas (VIGITEL) 2021. Os objetivos específicos são: i) examinar a relação entre as DCNTs e o status socioeconômico das pessoas, considerando indicadores como renda, escolaridade e acesso aos serviços de saúde; e ii) estimar um índice de qualidade de vida para a população respondente do VIGITEL 2021.

A realização deste estudo é justificada pela necessidade de compreender melhor o impacto das DCNTs na economia e na sociedade brasileira, visando informar políticas públicas mais eficazes e sustentáveis no enfrentamento dessas doenças. A aplicação do Método Alkire-Foster e a utilização dos dados do Inquérito VIGITEL 2021 permitirão uma análise atualizada e embasada do fardo econômico das DCNTs, fornecendo subsídios para a formulação de estratégias direcionadas e efetivas.

1.1 *Insights* sobre o conceito de qualidade de vida

Compreender com efetividade o conceito de Qualidade de Vida tem sido objeto de estudo e debate ao longo da história, refletindo a evolução das perspectivas filosóficas, sociológicas e científicas sobre o bem-estar humano. Farquhar (1995), destaca que não há um consenso sobre a definição do conceito de QV; Minayo, Hartz e Buss (2000) descrevem o conceito de QV como polissêmico, com diferentes

nuances, contextos e entendimentos. O desenvolvimento histórico desse conceito pode ser rastreado desde as antigas civilizações até os estudos contemporâneos.

As definições desse conceito estão intimamente ligadas às diferentes necessidades humanas: as necessidades podem ser básicas que, uma vez satisfeitas, permanecem insatisfeitas; outras necessidades não desaparecem e se concentram nos relacionamentos com os outros e nos sentimentos sobre si mesmo. Nesse contexto, Gómez e Sabeh (2001) argumentam que os indivíduos estão imersos em um mundo social no qual associam suas condições de vida e satisfação pessoal aos meios materiais necessários para sua sobrevivência, e no qual suas expectativas de vida e sistema de valores são influenciadas por normas e políticas sociais.

No contexto da filosofia clássica grega, o filósofo Aristóteles trouxe importantes reflexões sobre a busca pela Eudaimonia, um estado de bem-estar integral e plenitude. Em sua obra "Ética a Nicômaco", Aristóteles defende que a qualidade de vida está intrinsecamente ligada à realização de atividades virtuosas e ao desenvolvimento das capacidades humanas. Essas ideias fundamentais influenciaram o pensamento ocidental ao longo dos séculos (ARISTÓTELES, 350 a.C.).

Ao longo da história, outras correntes filosóficas contribuíram para a compreensão do conceito de Qualidade de Vida. No período iluminista, pensadores como John Locke e Immanuel Kant enfatizaram a importância dos direitos individuais e da liberdade como pilares para o bem-estar humano. Locke, em sua obra "Ensaio sobre o Entendimento Humano", destacou a necessidade de proteger a vida, a liberdade e a propriedade como elementos essenciais para a busca da felicidade (LOCKE, 1689). Kant, por sua vez, argumentou que a dignidade humana deve ser respeitada e que as pessoas devem ser tratadas como fins em si mesmos, e não como meros meios para alcançar determinados fins (KANT, 1785).

Ainda, a qualidade de vida deriva de um interesse em avaliar o bem-estar de um indivíduo em relação às suas expectativas, normas e preocupações, com o contexto social do indivíduo refletindo a lacuna entre as expectativas e a experiência real, a fim de alcançar um estado de satisfação geral. No campo da sociologia, a noção de Qualidade de Vida ganhou destaque durante o século XX, impulsionada por teóricos como Max Weber e Émile Durkheim. Weber, em sua obra "A Ética Protestante e o Espírito do Capitalismo", investigou a influência das crenças religiosas e dos valores éticos na qualidade de vida das pessoas (WEBER, 1904-1905). Já Durkheim,

em "O Suicídio", analisou os fatores sociais que podem afetar negativamente a qualidade de vida, como a falta de coesão social e a anomia (DURKHEIM, 1897).

Esses indicadores evoluíram, inicialmente referindo-se a condições econômicas e sociais objetivas e, depois, começando a levar em conta os elementos subjetivos do indivíduo (GÓMEZ; SABEH, 2001). Nas décadas de 1930 e 1940, Palomino e López (PALOMINO; LÓPEZ, 1999) argumentaram que as tendências econômicas deveriam ser medidas em termos de produto nacional bruto (PNB), que fornecia o valor monetário de toda a riqueza gerada em um país.

Além disso, outro indicador econômico foi adicionado para medir as condições de bem-estar, a saber, a renda per capita, obtida pela divisão do PNB pela população total do país. Entretanto, estudos posteriores mostraram (QUARTOS; GARZÓN, 2013) que esses dois indicadores são limitados para medir o bem-estar social alcançado pela população de um país e a qualidade de vida de seus habitantes.

A esse respeito, o desenvolvimento e o refinamento dos indicadores sociais na década de 1970 e no início da década de 1980 permitiram avançar na diferenciação entre indicadores sociais e qualidade de vida, sendo que esse último termo começou a ser definido como um conceito integrador que leva em conta todas as áreas da vida (de natureza multidimensional), envolvendo tanto condições objetivas quanto componentes subjetivas (QUARTOS; GARZÓN, 2013).

Nos tempos contemporâneos, o estudo da Qualidade de Vida tornou-se multidimensional, abarcando aspectos físicos, psicológicos, sociais e ambientais. A Organização Mundial da Saúde (OMS) desempenhou papel central na definição desse conceito moderno, definindo-a como "a percepção do indivíduo de sua posição na vida, no contexto da cultura e do sistema de valores nos quais ele vive, e em relação a seus objetivos, expectativas, padrões e preocupações" (OMS, 1994). Logo, essa perspectiva leva em conta não apenas indicadores objetivos, como saúde física e condições materiais, mas também fatores subjetivos, como emoções, relações sociais, satisfação no trabalho e sentido de propósito.

A Organização das Nações Unidas (ONU) tem desempenhado papel importante na definição e no monitoramento da qualidade de vida em nível global. O Relatório Mundial de Desenvolvimento Humano, publicado pelo Programa das Nações Unidas para o Desenvolvimento (PNUD), revela a importância da qualidade de vida como um dos principais objetivos do desenvolvimento humano. O relatório enfatiza a

necessidade de ir além dos indicadores econômicos tradicionais e considerar fatores como educação, saúde, igualdade de gênero, participação social e liberdades individuais (PNUD, 2020).

A abordagem do bem-estar subjetivo tem se destacado na compreensão da qualidade de vida. O bem-estar subjetivo refere-se à avaliação subjetiva que os indivíduos fazem de sua própria vida, levando em conta aspectos emocionais e cognitivos. Nessa perspectiva, a satisfação com a vida, o afeto positivo e a ausência de afeto negativo são considerados indicadores-chave de qualidade de vida (DIENER et al., 2018). Estudos nessa área têm demonstrado que o bem-estar subjetivo está associado a diversos aspectos da vida, como saúde física e mental, relacionamentos interpessoais e desempenho acadêmico e profissional.

1.2 Condicionantes da qualidade de vida: esforços analíticos e mensuração

A qualidade de vida, como um conceito abrangente, envolve uma variedade de elementos e características que influenciam o bem-estar e a satisfação das pessoas. Diversos estudos têm identificado elementos-chave que compõem a qualidade de vida, fornecendo uma base para a compreensão dessas características.

Um dos aspectos frequentemente associados à qualidade de vida é a saúde física. A presença de boa saúde e a ausência de doenças crônicas está relacionada a uma melhor qualidade de vida. Além disso, ter acesso a cuidados médicos adequados e manter a salubridade em seu estilo de vida, incluindo uma dieta equilibrada e atividade física regular, são fatores importantes para a promoção da qualidade de vida (JIA et al., 2018).

Outro elemento essencial é a saúde mental. A presença de um estado mental saudável, incluindo bem-estar emocional e ausência de transtornos psicológicos, é essencial à boa qualidade de vida. A capacidade que o indivíduo tem em lidar com seu estresse, a presença de emoções positivas e a autopercepção de uma boa saúde mental são indicadores relevantes nesse contexto (HUPPERT; COOPER, 2014).

A qualidade dos relacionamentos interpessoais é outro componente crucial da qualidade de vida. Ter relações sociais saudáveis e satisfatórias, incluindo vínculos familiares, amigos e redes de apoio, está associado a um maior bem-estar e

satisfação na vida. A presença de apoio social, suporte emocional e interações positivas com os outros contribuem para a qualidade de vida (THOITS, 2011).

Além disso, a realização pessoal e o desenvolvimento de habilidades e talentos são elementos que compõem a qualidade de vida. A capacidade de perseguir metas e interesses pessoais, ter um senso de propósito e autonomia, e desenvolver habilidades e competências ao longo da vida, são aspectos que contribuem ao aprimoramento da melhor qualidade de vida (KEYES et al., 2002). Tendo em vista o exposto, a medição da qualidade de vida no Brasil envolve a utilização de diversas pesquisas e instrumentos que fornecem uma compreensão abrangente dos indicadores de bem-estar da população. Essas abordagens contribuem para o monitoramento e a formulação de políticas públicas direcionadas ao aprimoramento das condições de vida da sociedade.

Um exemplo de fonte de dados para a mensuração da qualidade de vida no Brasil é a Pesquisa Nacional por Amostra de Domicílios (PNAD), que é conduzida pelo Instituto Brasileiro de Geografia e Estatística (IBGE). Ela abrange uma ampla gama de temas socioeconômicos, como educação, saúde, trabalho e renda, fornecendo informações valiosas sobre as condições de vida da população (IBGE, 2021).

Dos indicadores que representam o fenômeno, cita-se o Índice de Desenvolvimento Humano (IDH), desenvolvido pelo Programa das Nações Unidas para o Desenvolvimento (PNUD). O IDH é um indicador composto que combina dados sobre educação, renda e longevidade, permitindo uma avaliação multidimensional da qualidade de vida no Brasil e em outros países (PNUD, 2020).

No campo da saúde, a Organização Mundial da Saúde (OMS) desenvolveu a World Health Survey (WHS), uma pesquisa utilizada em diversos países, incluindo o Brasil, para avaliar a qualidade de vida relacionada à saúde. Esse instrumento aborda questões como acesso a serviços de saúde, condições crônicas, bem-estar emocional e funcionamento físico, proporcionando uma visão abrangente do bem-estar da população (OMS, 2002).

O WHOQOL-BREF, uma versão abreviada do instrumento de avaliação da qualidade de vida da OMS, é amplamente utilizado no Brasil. Esse questionário avalia diferentes domínios da qualidade de vida, como saúde física, saúde psicológica, relações sociais e ambientes, permitindo uma análise mais aprofundada dos aspectos que compõem o bem-estar (FLECK et al., 1999).

Essas pesquisas e instrumentos têm sido fundamentais para a compreensão da qualidade de vida no Brasil, fornecendo informações valiosas para a tomada de decisões em políticas públicas e para a implementação de intervenções direcionadas ao aprimoramento do bem-estar da população. Sobretudo, argumenta-se que a avaliação crítica da qualidade das medições de qualidade de vida é um aspecto crucial para garantir a confiabilidade e validade dos resultados obtidos.

Assim sendo, diversas abordagens têm sido adotadas para examinar de forma sistemática as limitações e as qualidades dos instrumentos e pesquisas utilizados nesse contexto. Uma das estratégias comuns é a análise psicométrica, que envolve a avaliação das propriedades psicométricas dos instrumentos de medição, como validade, confiabilidade e sensibilidade ao contexto cultural. Essa análise visa identificar potenciais vieses e limitações na forma como os construtos de qualidade de vida são definidos e mensurados (REVICKI et al., 2006).

Outra abordagem utilizada é a comparação entre diferentes instrumentos e pesquisas, buscando identificar convergências e divergências nos resultados obtidos. Isso permite avaliar a consistência e a comparabilidade dos dados gerados pelas diferentes ferramentas de medição (COONS et al., 2007).

A perspectiva dos participantes também é fundamental na avaliação crítica da qualidade das medições de qualidade de vida. A inclusão da voz dos indivíduos que estão sendo avaliados é essencial para compreender a adequação e a relevância dos itens e das perguntas utilizadas nos instrumentos de medição (PATRICK; ERICKSON, 2016).

Uma abordagem complementar é a revisão sistemática da literatura, que analisa estudos anteriores sobre a medição da qualidade de vida e identifica as limitações e as lacunas existentes. Essa análise crítica contribui para o avanço do conhecimento e orienta o desenvolvimento de instrumentos mais robustos e confiáveis (APÓSTOLO et al., 2018).

É importante ressaltar que a avaliação crítica da qualidade das medições de qualidade de vida é um processo contínuo e evolutivo, pois as necessidades e as percepções das populações mudam ao longo do tempo. Portanto, é essencial que as pesquisas e os instrumentos de medição sejam revisados e atualizados regularmente para garantir que eles continuem sendo representativos e relevantes para a sociedade. Nessa seara, destaca-se também o Relatório da Vigilância de Fatores de Risco e

Proteção para Doenças Crônicas em Inquérito Telefônico VIGITEL, que conjuga diferentes hábitos de saúde e bem-estar da população brasileira.

2 PROCEDIMENTOS METODOLÓGICOS

Sabina Alkire e James Foster (2007) desenvolveram o Método Alkire-Foster (AF), que consiste em 12 etapas, com o objetivo de mensurar privações sob uma abordagem multidimensional (ALKIRE; FOSTER, 2009). Esse método serviu como base para a construção do Índice de Privação Multidimensional (MPI) apresentado pelo Programa das Nações Unidas para o Desenvolvimento (PNUD). O AF se tornou uma referência acadêmica amplamente utilizada para mensurar privações em diferentes perspectivas.

Uma das vantagens desse método é a sua facilidade de interpretação, pois ele explicita os passos para identificar privações simultâneas e oferece diversas formas de apresentar essas privações por meio da mensuração da intensidade (FERES; VILLATORO, 2013). Dessa forma, o método é flexível, permitindo que o pesquisador defina os cortes e pesos das dimensões e indicadores de acordo com o contexto estudado.

Uma fortaleza dessa abordagem é a identificação de dimensões e indicadores que permitem a análise e agregação proposta por Sen (1976), possibilitando a mensuração da privação com base em diferentes aspectos, além do aspecto monetário. Alkire e Foster (2007) estabeleceram o Método AF, composto por 12 etapas, para medir a privação multidimensional, integrando diversas dimensões e indicadores para uma abordagem abrangente da privação em diferentes níveis, como indivíduos, domicílios e comunidades, considerando o contexto em que vivem e seu acesso a recursos (Quadro 1):

Passo	Descrição
1º	Escolha da unidade de análise, que pode ser por indivíduo, domicílio, comunidade, etc.
2º	Escolha das dimensões que serão consideradas para conceituar a privação, podem ser saúde, educação, habitação, etc.
3º	Escolha dos indicadores, ou seja, dentro de cada dimensão são eleitos pontos importantes que devem ser considerados para conceituar a privação.
4º	Escolha dos indicadores, ou seja, dentro de cada dimensão são eleitos pontos importantes que devem ser considerados para conceituar a privação.
5º	Aplicação da primeira linha da privação. Considerando cada limite determinado no passo anterior, é identificado o status de cada indivíduo, domicílio, comunidade, etc. entre privado e não privado.
6º	Contagem do número de privações para cada indivíduo, domicílio, comunidade etc.
7º	Definição da segunda linha de corte de privação, identificada por "k", que define quantos indicadores os indivíduos, domicílios, comunidades etc. devem ser privados para serem considerados privados de maneira multidimensional.
8º	Aplicação da linha "k" para obter a lista de indivíduos privados e separá-los dos não privados. As informações dos não privados são substituídas por zero, e os privados, considerados com privações, assumem o valor de um.
9º	Cálculo do percentual de privados, "H", dividindo o total de pessoas que são privados em "k" ou mais indicadores pelo total geral da amostra, obtendo-se assim a proporção de pessoas privadas em relação à população da amostra.
10º	Cálculo do hiato da privação média, "A". Soma-se as proporções das privações de todas as pessoas privadas e divide-se pelo número total de privados, encontrando o número médio de privações que os indivíduos, domicílios, comunidades, etc. considerados privados sofrem.
11º	Cálculo da incidência ajustada, "M0". Multiplica-se "H" por "A", resultando em um valor que varia de um, quando a população se aproxima da privação extrema, a zero, quando a população se afasta da situação de privação.
12º	Decomposição por grupos e/ou dimensões, em que "M0" pode ser decomposto para cada subgrupo da população. É possível analisar a contribuição de cada dimensão para a privação geral. Para isso, calcula-se "A" dividido pelo número de privados na dimensão "j" (considerando qualquer uma das dimensões consideradas), resultando em "Aj". Então, "Aj" multiplicado por "H" permite encontrar a dimensão ajustada que mostra a participação da dimensão "j" na privação global, "M0j".

Quadro 1 – Passos a serem seguidos

Fonte: adaptado de Alkire e Foster (2007).

A possibilidade de medir a privação com base nas etapas de identificação e agregação propostas por Sen (1976) é a principal vantagem do método Alkire-Foster,

pois oferece flexibilidade ao pesquisador para definir os cortes e pesos das dimensões que estão sendo analisadas (FERES; VILLATORO, 2013). Neste estudo, com vistas ao contexto da estimativa de um Índice de Privação de Qualidade de Vida (IPQV) para os brasileiros, o método Alkire-Foster foi adaptado para manipular os dados da Pesquisa VIGITEL 2021, um inquérito telefônico realizado pelo Ministério da Saúde. O VIGITEL monitora anualmente a frequência e a distribuição dos principais fatores de risco e proteção associados às Doenças Crônicas Não Transmissíveis (DCNT) em todas as capitais dos 26 estados brasileiros e no Distrito Federal.

É importante esclarecer que, de acordo com a disponibilidade das informações, a unidade de análise considerada são os indivíduos com 18 anos ou mais, residentes nas capitais dos 26 estados brasileiros e no Distrito Federal, que foram entrevistados na pesquisa VIGITEL 2022, realizada durante o ano de 2021 e parte de 2022. Com relação à escolha das dimensões, é importante mencionar que não há consenso sobre os critérios analíticos adotados na literatura, incluindo a seleção das dimensões e sua importância (ALKIRE, 2008). As escolhas das dimensões estão principalmente relacionadas à experiência de vida de cada indivíduo e à crença nas variáveis relevantes que têm influência direta e permitem o acesso a diferentes aspectos da vida, o que pode levar a considerá-las multidimensionalmente privados ou não.

O IPQV foi estimado considerando três dimensões analíticas e onze indicadores (consulte o Quadro 1). Destaca-se que a escolha das dimensões foi baseada na literatura sobre medidas multidimensionais de privação, como os trabalhos de Barros e Silva (2006), Kageyama e Hoffmann (2006), Alkire e Santos (2010), Kerstenetzky et al. (2011), Albuquerque e Cunha (2012) e Brites et al. (2013). Os indicadores foram selecionados com base na disponibilidade de dados na base do VIGITEL 2021. Um total de 27.094 pessoas participaram do inquérito mais recente realizado pelo Ministério da Saúde. A pesquisa foi conduzida durante os anos de 2021 e 2022, com residentes das 26 capitais estaduais do Brasil e do Distrito Federal.

Os passos para a aplicação do Método Alkire-Foster (AF) estão apresentados no Quadro 1. Nesse sentido, é necessário definir as linhas de corte a serem utilizadas (Quadro 2). A primeira linha de corte foi estabelecida considerando se o indivíduo é classificado como privado ou não em cada um dos indicadores. A situação de não privação é o oposto dessa classificação. A segunda linha de privação é determinada pela mensuração se o indivíduo é multidimensionalmente privado ou não. Esse critério

é estabelecido da seguinte maneira: se o indivíduo é privado em três ou mais indicadores, é considerado multidimensionalmente privado.

Dimensões	Indicadores	Privado Se...*
Perfil Socioeconômico	Escolaridade	Tem menos que 9 anos de estudo
	Trabalho	Não trabalha
	Posse de plano de saúde	Não possui plano de saúde
Fatores de Risco	Excesso de Peso (**)	Tem excesso de peso
	Depressão	Tem depressão
	Frequência do Álcool	Ingere álcool de 3 a 4 dias ou de 5 a 6 dias da semana; ou ainda, todos dias da semana inclusive sábado e domingo.
	Fumante	É fumante (diário ou esporádico)
	Exercício físico	Não realiza exercícios físicos.
Repercussões das Doenças Crônicas	Hipertensão	Tem hipertensão
	Obesidade (***)	Tem obesidade
	Pessoas com diabetes	Tem diabetes

Quadro 2 – Dimensões, indicadores e linhas de corte para estimação do IPQV

Fonte: Elaborado pelos autores com base na disponibilidade de dados do relatório VIGITEL 2021.

* Nota: O indivíduo é considerado privado se enquadrar-se na condição definida.

** Nota: O Excesso de Peso é conhecido pela literatura como Sobrepeso.

***Nota: A obesidade, de acordo com Singh (2013), pode ser considerada DCNT ou Fator de Risco.

Dentre os passos da aplicação do método AF, um deles é a definição da segunda linha de privação, que consiste em determinar um valor de corte para cada indicador, de modo a classificar cada observação como privada ou não privada. Por exemplo, no caso do indicador Escolaridade, um indivíduo é considerado privado se tiver menos de nove anos de estudo, ou seja, se não concluiu o ensino fundamental.

Para o indicador Trabalho, um indivíduo é considerado privado se estiver desempregado, ou seja, não está trabalhando. No caso do indicador Posse de Plano de Saúde, um indivíduo é considerado privado se não possui plano de saúde, ou seja, não tem acesso a assistência médica e hospitalar privada contínua, nem cobertura de custos assistenciais em saúde prestados por empresas privadas registradas na Agência Nacional de Saúde Suplementar (ANS).

Para o indicador Excesso de Peso, um indivíduo é considerado privado se estiver com sobrepeso, ou seja, se seu índice de massa corporal (IMC) estiver entre 25kg/m² e 29,9kg/m². Isso indica que há um acúmulo de gordura corporal superior ao ideal para uma vida saudável.

No caso do indicador Depressão, um indivíduo é considerado privado se apresentar um transtorno de humor caracterizado por episódios depressivos leves, moderados ou graves. Isso inclui sintomas como apatia, irritabilidade, perda de interesse, tristeza, alterações motoras, ideias agressivas, desolação e múltiplas queixas somáticas (insônia, fadiga, anorexia), que afetam o bem-estar emocional da pessoa (BALLONE, 2003).

No caso do indicador Frequência do Álcool, um indivíduo é considerado privado se consome álcool de 5 a 6 dias por semana ou todos os dias, incluindo sábado e domingo. Isso indica uma perturbação crônica de comportamento relacionada ao consumo excessivo de álcool, que afeta a saúde e o funcionamento social e econômico da pessoa.

No caso do indicador Fumante, um indivíduo é considerado privado se é fumante, ou seja, se inala a fumaça do cigarro de forma direta, seja diariamente ou não. Considera-se fumante uma pessoa que fumou mais de 100 cigarros ou 5 maços de cigarros em toda a sua vida e continua fumando atualmente (OPAS, 2015). Para o indicador Exercício Físico, um indivíduo é considerado privado se não pratica exercícios físicos, ou seja, se é sedentário.

No caso do indicador Hipertensão, um indivíduo é considerado privado se for hipertenso, ou seja, se apresentar valores de pressão arterial sistólica (PAS) iguais ou acima de 140 mmHg e/ou pressão arterial diastólica (PAD) iguais ou acima de 90 mmHg em medições em repouso. Esses valores são considerados como pressão arterial elevada para pessoas com 13 anos ou mais (BARROSO et al., 2020, p. 528).

Para o indicador Obesidade, entende-se tanto o sobrepeso quanto a obesidade, pois ambos se referem ao acúmulo de gordura em excesso no corpo do indivíduo (BRASIL - Ministério da Saúde, 2021), o que leva a considerá-lo como uma pessoa privada se for obesa, de acordo com a definição da OMS, que considera um IMC⁶

⁶ De acordo ao Ministério da Saúde (2021), existem escalas de classificação de IMC – Índice de Massa Corpórea, que é calculado com base na relação de cálculo $IMC = \text{peso}/\text{altura}^2$, usadas para determinar peso ideal e gradações que indicam magreza, sobrepeso e obesidade. Para o sobrepeso, o IMC fica entre 25 e 29,99 enquanto a obesidade seria diagnosticada com o IMC a partir de 30. A obesidade grau I, grau II e a grau III (chamada de obesidade mórbida), caracterizam-se por condições mais graves, que impactam de maneira mais relevante a QV do indivíduo, condicionando a doença a um índice de mortalidade mais elevado.

igual ou superior a 30 kg/m² como obesidade. A faixa de peso normal varia entre 18,5 kg/m² e 24,9 kg/m². Indivíduos com IMC entre 25 kg/m² e 29,9 kg/m² são diagnosticados com sobrepeso, podendo já apresentar prejuízos relacionados ao excesso de gordura. A obesidade é caracterizada pelo excesso de gordura corporal, que pode afetar negativamente a saúde (OMS,2018).

Para o indicador Pessoas com Diabetes, um indivíduo é considerado privado se for diabético, ou seja, se apresentar níveis de glicose em jejum iguais ou superiores a 126 mg/dl ou se tiver produção insuficiente ou má absorção de insulina.

Após realizar a soma das privações de cada indivíduo, foi estabelecida a segunda linha de privação, denominada de "k". Essa linha define o número de privações que um indivíduo precisa ter para ser considerado multidimensionalmente privado. O valor escolhido para a segunda linha de privação foi de 6,114, correspondendo a um terço do total de indicadores avaliados neste estudo.

Em seguida, foram realizadas as estimativas de H, que representa a proporção de indivíduos privados na população analisada, e o cálculo de A, que identifica a proporção de indicadores nos quais os indivíduos são privados. Para esse cálculo, é necessário dividir o número total de indivíduos da amostra que apresentam 6 ou mais privações pelo total de indicadores avaliados.

Por fim, foi estimada a incidência ajustada, obtida multiplicando a proporção de indivíduos privados (H) pela média das privações (A). Esse indicador, denominado IPQV, resume em um único índice a incidência e a intensidade da privação multidimensional. A importância da incidência ajustada está na sua capacidade de avaliar as políticas públicas voltadas para o combate às privações (GALLO; ROCHE, 2011).

Ressalta-se que o valor do IPQV, que varia de 0 a 1 (sendo mais próximo de 1 indicando maior privação), pode ser interpretado como a porcentagem de uma possível situação de extrema privação que a população em questão poderia enfrentar (GALLO; ROCHE, 2011).

O cálculo da incidência de privação (H) mostra a proporção de pessoas privadas de QV e privadas em relação ao total de indivíduos analisados. Para cada valor de k, é calculada a porcentagem de indivíduos privados em relação ao total de indivíduos, conhecida como Headcount, conforme a seguinte equação:

$$H = \frac{q}{n}$$

Em que:

- H = Headcount;
- q = número de pessoas multidimensionalmente privados;
- n = número total de pessoas.

Nesse caso, a incidência de privação seria de 60%, ou seja, 60% da amostra é considerada multidimensionalmente privada. Alkire e Santos (2009) salientam que Headcount não é sensível com o número de privações que o pobre enfrenta.

Assim, calcula-se o hiato de privação média: o hiato de privação média (A) mostra a média da fração de dimensões que os indivíduos são privados. Esse índice também é importante, pois consegue captar a intensidade da privação, visto que representa quantos indicadores em média os indivíduos privados de QV são privados. Segundo Alkire e Foster (2009), o hiato de privação média é calculado somando-se a proporção total de privações que cada pessoa sofre e dividindo pelo total de pessoas multidimensionalmente privados.

Segundo o exemplo, há três pessoas multidimensionalmente privadas de QV, privadas em dois indicadores cada. Assim,

$$A = \frac{\sum_{i=1}^h \frac{\text{privações}}{11}}{N^{\circ} \text{ total de pobres } (3 \geq k)}$$

Nesse caso, A= 0,67 quer dizer que em média os indivíduos são privados em 67% dos indicadores considerados.

Por fim, o cálculo da incidência ajustada: a incidência ajustada (M0) mostra a proporção de privações que a população privada de QV enfrenta em relação ao máximo das privações que toda a população poderia sofrer. Quando M0 = 1, significa que toda a população é privada em todos os indicadores/dimensões considerados (GALLO; ROCHE, 2011). O cálculo da incidência ajustada, em que se multiplica H por A, ou seja, M0= (H x A), e seu valor pode variar de um (quando a população se aproxima da extrema pobreza). Consiste na multiplicação do hiato de privação pelo *headcount*. Seguindo o exemplo anterior, M0 = 0,6 * 0,67 = 0,40*(10). Isso significa

que o total de indicadores em que as pessoas privadas de QV sofrem privação representa 40% do máximo possível na amostra em estudo. É importante salientar que como o hiato de privação leva em consideração a proporção de privações que cada pessoa sofre, se a pessoa 1 for privada em três indicadores em vez de dois, o índice aumentaria. Nesse sentido, o M0 obedece ao axioma de monotonicidade.

3 ANÁLISE E DISCUSSÃO DOS RESULTADOS

Participaram do Inquérito VIGITEL 2021, 27.093 pessoas, brasileiros de todas as capitais do país e do Distrito Federal, com 18 anos ou mais de idade, que responderam às entrevistas entre os meses de setembro de 2021 e fevereiro de 2022. O inquérito VIGITEL objetiva que se forneçam subsídios em prol de realizar o planejamento para políticas públicas da área da saúde, visando enfrentar adequadamente os avanços das Doenças Crônicas Não Transmissíveis (DCNT) (VIGITEL, 2021).

Para a composição do questionário VIGITEL 2021, as perguntas versaram sobre diversas características dos entrevistados, destacando-se as sociodemográficas, de padrão alimentar e de realização (ou não) de atividade física, associando-as à possível ocorrência de DCNTs. Inquiriu-se, ainda, sobre o consumo do tabagismo e das bebidas alcoólicas (e sua frequência), autoavaliação acerca do estado de saúde; questões sobre consultas/diagnósticos médicos anteriores em relação a hipertensão arterial, DM e depressão; exames preventivos – e sua realização, para detectar precocemente câncer de cólo de útero e mamografias para as mulheres; utilização de planos de saúde particulares e/ou de convênios médicos; e por fim, perguntas sobre as situações que enfrenta no trânsito (VIGITEL, 2021).

Com relação ao Tabagismo, os principais achados do relatório VIGITEL 2021 apontaram as frequências de maior destaque conforme a Tabela 1.

Tabela 1 – Frequência de tabagismo

TABAGISMO	Do total dos entrevistados	Homens	Mulheres
frequência de adultos fumantes	9,1%	11,8%	6,7%
escolaridade entre 0 e 8 anos	12,9%	15,7%	10,3%
fumantes passivos no domicílio	6,9%	7,6%	6,4%
fumantes passivos no ambiente de trabalho	5,4%	8,1%	3,2%

Fonte: Relatório ViGITEL (2021).

Dentre as capitais brasileiras com maior incidência de tabagismo (Tabela 2), cita-se Campo Grande, que tem 14,5% dos fumantes brasileiros. Ainda, 22,2% dos brasileiros homens tabagistas residem nessa capital.

Tabela 2 – Tabagismo – Frequências de tabagismo nas principais capitais dos estados brasileiros e Distrito Federal

Capitais/DF	Total	Masculino	Feminino
	%	%	%
Aracaju	4	6,1	2,3
Belém	5,8	6,9	4,9
Belo Horizonte	10,6	14,7	7,2
Boa Vista	8,6	12,1	5,4
Campo Grande	14,5	22,2	7,5
Cuiabá	7,4	10,9	4,2
Curitiba	11,3	14,9	8,1
Florianópolis	8,8	9	8,7
Fortaleza	7,2	9,3	5,4
Goiânia	10,4	14,4	6,8
João Pessoa	7,5	10,5	5
Macapá	5,4	7,5	3,4
Maceió	6,5	10,1	3,6
Manaus	7,1	11,7	2,9
Natal	7	8,9	5,5
Palmas	7	10,3	4,1
Porto Alegre	9,3	11,9	7,3
Porto Velho	11,1	13,3	8,6
Recife	9,8	12,5	7,6
Rio Branco	11,2	12,9	9,6
Rio de Janeiro	7,7	8,1	7,2
Salvador	6,1	10,4	2,6
São Luís	5,3	10	1,5
São Paulo	11,2	13	9,7
Teresina	6,4	12,1	1,6
Vitória	7,1	9,1	5,5
Distrito Federal	11,8	17,7	6,6

Fonte: Relatório ViGITEL (2021).

Entre as mulheres, o maior índice de tabagistas está em São Paulo (9,7%), e o menor em São Luís (1,5%). As menores frequências de adultos tabagistas se encontram em Aracaju, sendo que 6,1% dos brasileiros homens e tabagistas residem nessa capital (Tabela 2).

Sobre o excesso de peso e obesidade da população do país, o VIGITEL 2021 apresenta os seguintes dados (Tabela 3):

Tabela 3 – Frequência de sobrepeso e obesidade por idade e sexo

SOBREPESO	Do total dos entrevistados	Homens	Mulheres
Índice de adultos com excesso de peso	57,2%	59,9%	55%
população com idade entre 18 e 24 anos	35,7%	33,6%	31,7%
população com idade entre 25 e 34 anos	54,4%	63,4%	46,6%
população com idade entre 35 e 44 anos	62,4%	62,9%	61,9%
população com idade entre 45 e 54 anos	64,4%	67,2%	61,9%
população com idade entre 55 e 64 anos	64,1%	61,9%	65,7%
população com idade entre 65 anos ou mais	60,7%	60,7%	60,7%
OBESIDADE	Do total dos entrevistados	Homens	Mulheres
índice de adultos obesos	22,4%	22%	22,6%
população com idade entre 18 e 24 anos	12,2%	13,1%	11,2%
população com idade entre 25 e 34 anos	20,8%	25,5%	16,6%
população com idade entre 35 e 44 anos	25,5%	25,1%	25,7%
população com idade entre 45 e 54 anos	26,2%	24,3%	28,3%
população com idade entre 55 e 64 anos	26,2%	22,1%	29,3%
população com idade entre 65 anos ou mais	21,8%	16,8%	25,3%

Fonte: Relatório VIGITEL (2021).

A Tabela 4 mostra os percentuais de adultos (≥ 18 anos) com excesso de peso (IMC ≥ 25 kg/m²), classificados por sexo, e pelas capitais dos estados brasileiros e do Distrito Federal.

Assim, sobre as capitais brasileiras onde há maior incidência de adultos com excesso de peso, menciona-se Porto Velho, tem 64,4% dos respondentes com essa característica. E a menor está em São Luís, com 49,3% dos brasileiros com excesso de peso (Tabela 4):

Tabela 4 – Percentual* de adultos (≥ 18 anos) com excesso de peso (IMC ≥ 25 kg/m²), por sexo, segundo as capitais dos estados brasileiros e o Distrito Federal

Capitais/DF	Total	Masculino	Feminino
	%	%	%
Aracaju	57,3	60,9	54,4
Belém	61,3	61,6	61,0
Belo Horizonte	58,6	58,4	58,7
Boa Vista	56,4	58,7	54,2
Campo Grande	58,1	61,3	55,4
Cuiabá	57,0	58,5	55,7
Curitiba	55,3	62,8	48,9
Florianópolis	56,9	59,5	54,5
Fortaleza	59,2	61,2	57,5
Goiânia	56,3	62,1	51,2
João Pessoa	59,6	66,5	53,9
Macapá	58,7	62,3	55,5
Maceió	58,1	59,2	57,2
Manaus	63,5	65,2	61,8
Natal	59,1	60,9	57,5
Palmas	50,1	55,9	45,0
Porto Alegre	62,2	64,4	60,4
Porto Velho	64,4	67,5	61,0
Recife	56,7	58,1	55,6
Rio Branco	60,4	63,2	57,8
Rio de Janeiro	56,1	64,0	49,4
Salvador	53,2	50,8	55,1
São Luís	49,3	51,4	47,5
São Paulo	57,4	57,6	57,3
Teresina	52,5	60,0	46,4
Vitória	51,5	55,8	47,8
Distrito Federal	56,4	59,4	53,7

Fonte: Relatório ViGITEL (2021).

Porto Velho é a capital que concentra maior percentual (67,5%) entre os homens e Manaus a que concentra maior percentual (61,8%) entre as mulheres. As menores frequências entre os homens adultos com excesso de peso se encontram em Salvador 50,8% e entre as mulheres, em Palmas (45%), conforme a Tabela 4.

Já sobre a obesidade e as frequências por capital, destaca-se a Tabela 5:

Tabela 5 – Percentual* de adultos (≥ 18 anos) com obesidade ($IMC \geq 30$ kg/m²), por sexo, segundo as capitais dos estados brasileiros e o Distrito Federal

Capitais/DF	Total	Masculino	Feminino
	%	%	%
Aracaju	25,0	27,9	22,6
Belém	24,0	23,8	24,2
Belo Horizonte	21,4	20,7	22,0
Boa Vista	23,6	25,0	22,3
Campo Grande	20,2	19,4	21,0
Cuiabá	23,8	22,6	25,0
Curitiba	22,6	24,3	21,1
Florianópolis	20,2	21,4	19,2
Fortaleza	23,7	23,2	24,1
Goiânia	23,3	26,7	20,3
João Pessoa	23,7	22,0	25,1
Macapá	23,7	23,7	23,7
Maceió	24,6	23,7	25,4
Manaus	25,0	23,3	26,6
Natal	23,0	23,0	23,1
Palmas	19,5	23,3	16,1
Porto Alegre	22,6	22,9	22,4
Porto Velho	26,4	26,6	26,2
Recife	22,6	17,7	26,5
Rio Branco	24,2	25,0	23,4
Rio de Janeiro	21,5	23,1	20,1
Salvador	20,5	18,6	22,0
São Luís	18,0	18,6	17,6
São Paulo	22,5	20,6	24,2
Teresina	20,3	24,1	17,2
Vitória	17,9	19,2	26,8
Distrito Federal	22,6	23,0	22,2

Fonte: Relatório ViGITEL (2021).

Sobre as menores frequências de obesidade da população do país, concentram-se nas capitais Vitória e São Luís. Já os maiores índices estão Porto Velho, Aracaju e Manaus.

De acordo à alimentação saudável, e hábitos de saúde, o Relatório Vigitel traz informações importantes (Tabela 6). No consumo de hortaliças, vegetais e frutas, são as mulheres (26,4%) as maiores consumidoras desses itens, segundo a recomendação de saúde; já os homens têm consumo menor para esses alimentos (16,9%). Entre os maiores de 18 anos, o consumo de alimentos vegetais ainda é baixo (22,1%), gradualmente aumentando conforme avança a idade dos respondentes, ou

quando há um nível de escolaridade igual ou maior que 12 anos de estudo. Na Tabela 6, são mostradas as três maiores frequências e as três menores, de consumo de hortaliças, frutas e verduras por capital e por sexo.

Tabela 6 – Maiores e menores frequências de consumo de hortaliças, verduras e frutas por capital e sexo

Posição	Capital	Homens	Mulheres	Posição	Capital	Homens	Mulheres
1º	Curitiba	37,1%					
2º	Belo Horizonte	36,7%					
3º	Porto Alegre	35,9%					
1º	Florianópolis		52,1%				
1º	Belo Horizonte		51,4%				
3º	Distrito Federal		50,3%				
				25º	Salvador	20,2%	
				26º	Rio Branco	19,2%	
				27º	São Luís	14,9%	
				25º	Salvador		30,6%
				26º	Porto Velho		28,2%
				27º	Rio Branco		25,7%

Fonte: Relatório ViGITEL (2021).

Outro consumo que é preocupante, refere-se aos refrigerantes em cinco ou mais dias da semana (Tabela 7). Dos entrevistados, os homens (17,2%) são os que têm consumo igual ou maior que cinco dias por semana e as mulheres respondem por 11,3%, de um total nacional de 14% entre os respondentes do VIGITEL. Na Tabela 7, são mostradas as três maiores e três menores frequências de consumo de refrigerantes em cinco ou mais dias da semana, por capital e por sexo.

Tabela 7 – Maiores e menores frequências de consumo de refrigerantes por capital e sexo

Posição	Capital	Homens	Mulheres	Posição	Capital	Homens	Mulheres
1º	Campo Grande	28,8%					
2º	Porto Alegre	27,6%					
3º	Cuiabá	24,6%					
1º	Porto Alegre		24,2%				
2º	Cuiabá		16,2%				
3º	Curitiba		15,8%				
				25º	Teresina	7,1%	
				26º	Salvador	6,9%	
				27º	Natal	4,6%	
				25º	Salvador		3,8%
				26º	Natal		3,4%
				27º	Maceió		2,8%

Fonte: Relatório ViGITEL (2021).

O consumo de refrigerantes em mais de cinco ou mais dias na semana diminui com o passar da idade, segundo o relatório. Já outra questão do inquérito versava sobre o dia anterior à ligação e se a pessoa havia consumido naquela data, cinco ou mais grupos de alimentos ultraprocessados. Houve 18,2% das respostas afirmativas para esse fato, e os homens (21,7%), consomem mais do que as mulheres (15,2%) esses tipos de alimentos.

Atividade física e lazer são apresentados como maior frequência de prática no sexo masculino (43,1%), já a população feminina respondeu por 31,3%. Conforme a idade do respondente, percebeu-se uma redução da prática conforme o avançar da idade, e conforme o aumento do nível da escolaridade.

Na Tabela 8, demonstra-se as três maiores e três menores frequências dos adultos praticantes de atividade física, que utilizam pelo menos 150 minutos por semana para essa prática:

Tabela 8 – Maiores e menores frequências de prática de atividade física por capital e sexo

Posição	Capital	Homens	Mulheres	Posição	Capital	Homens	Mulheres
1º	Belém	50,0%					
2º	Recife	49,2%					
3º	São Luís Aracaju	49,0%					
1º	Vitória		35,9%				
2º	Palmas		36,6%				
3º	Natal		1%				
				25º	Cuiabá	44,5%	
				26º	São Paulo	41,7%	
				27º	Campo Grande	39,7%	
				25º	Porto Alegre		30,1%
				26º	São Paulo		28,7%
				27º	Rio Janeiro		24,2%

Fonte: Relatório ViGITEL (2021).

Novamente, constatou-se que 48,2% dos entrevistados com elevado nível de escolaridade não relatou nível de prática de atividade física em período suficiente. Já os sedentários em relação à atividade física compõem 15,8% do total dos respondentes.

O abuso no consumo de álcool e bebidas alcoólicas (no caso das mulheres, em quantidade igual ou superior a quatro doses; para os homens, igual ou acima de cinco doses), num mesmo período (ocasião), nos últimos 30 dias antes das respostas à pesquisa) foi alvo de respostas entre 18,3% dos respondentes, sendo maior em

homens (25,0%) do que em mulheres (12,7%). Na Tabela 9, demonstra-se as três maiores e as três menores frequências dos adultos que consomem bebidas alcoólicas:

Tabela 9 – Maiores e menores frequências de consumo de bebidas alcoólicas nos últimos trinta dias por capital e sexo

Posição	Capital	Homens	Mulheres	Posição	Capital	Homens	Mulheres
1º	Belo Horizonte	36,2%					
2º	Cuiabá	35,0%					
3º	Vitória	32,6%					
1º	Florianópolis		17,6%				
2º	Rio de Janeiro		16,6%				
3º	Distrito Federal		16,3%				
				25º	Maceió	19,0%	
				26º	Rio Branco	17,4%	
				27º	Porto Alegre	15,6%	
				25º	São Paulo		9,7%
				26º	Curitiba		8,6%
				27º	Maceió		8,4%

Fonte: Relatório VIGITEL (2021).

Nos últimos passos do método AF é calculada a proporção dos privados de QV (H), a intensidade da privação ou privação média (A) e a incidência ajustada pela intensidade (M0). O “M0” é obtido, como destacado na seção sobre o método, por meio da multiplicação da proporção de privados (H) pela privação média (A), $M0 = H \times A$ (essa medida permite a decomposição, por exemplo, por área rural e urbana, entre outras).

No Quadro 3 apresentam-se os resultados agregados, que correspondem aos passos 7 a 11 do método AF descritos anteriormente, no ano de 2021, para cada $k \geq 6$:

Estimativa	Significado	Conceito	Cálculo	Resultado	Interpretação
H	Cálculo do percentual de privados de QV	É a proporção média de privação entre os privados de QV	Divide-se o total de pessoas que são privados em k (3.124) ou mais indicadoras pelo total geral da amostra (8.876 - nº de observações úteis), obtendo-se assim uma proporção.	0,3621	Das 8.876 observações úteis, em média, 36,21% ou 3.214 dos privados de QV apresentam grau de privação.
A	Cálculo do hiato da privação média	O hiato médio de privação significa em média em quantos indicadores avaliados os indivíduos são privados.	Soma-se a proporção das privações de todas as pessoas privadas e divide-se pelo número total de pessoas privadas de QV, obtendo-se o número médio de privações que uma pessoa privada de QV sofre.	0,7523	Dos 11 indicadores avaliados, em média, os indivíduos são privados em 8 (75,23%) deles.
Mo	Cálculo da incidência ajustada	Cálculo e incidência privada ajustada significa qual é a proporção de privação que a amostra privada enfrenta em relação ao máximo de todas as privações que toda a amostra poderia sofrer.	Multiplica-se H por A, ou seja, $MO = H \times A$. Seu valor pode variar de um (quando a população se aproxima da extrema privação) a zero (quando se afasta da situação de privação).	0,2724	Isso significa que o total de indicadores em que os privados de QV sofrem privação representa 27,24% do máximo possível na amostra em estudo.

Quadro 3 – Resultados da pesquisa de privação multidimensional das grandes regiões brasileiras para $k \geq 6$ indicadores em 2001 e 2011

Fonte: dados da pesquisa (2023).

A primeira estimativa no Quadro 3 é o cálculo do percentual de privados em qualidade de vida. Essa estimativa representa a proporção média de privação entre os indivíduos considerados privados na amostra. Para calcular esse percentual, divide-se o número total de indivíduos que são privados em 3.124 indicadores ou mais pelo número total de observações úteis na amostra, que é de 8.876 pessoas. O

resultado desse cálculo é de 0,3621, o que significa que, em média, 36,21% dos privados de QV na amostra (aproximadamente 3.214 pessoas) apresentam um grau de privação.

Essas condições podem incluir acesso limitado a serviços de saúde de qualidade, dificuldades no acesso a alimentos saudáveis, ambientes inadequados para a prática de atividades físicas e menor acesso às informações de prevenção e cuidado com a saúde. Portanto, o percentual elevado de privação entre os indivíduos privados de QV na amostra pode indicar um risco aumentado para o desenvolvimento e o agravamento das DCNTs nessa população.

A segunda estimativa no Quadro 3 é o cálculo do hiato da privação média. Essa estimativa indica em quantos indicadores os indivíduos privados de QV são privados, em média. Para calcular o hiato médio de privação, somou-se a proporção de privações de todas as pessoas privadas de QV e dividiu-se pelo número total de pessoas privadas de QV na amostra. O resultado desse cálculo é de 0,7523, o que significa que, em média, os indivíduos privados de QV na amostra são privados em 8 dos 11 indicadores avaliados, o que representa aproximadamente 75,23% de privação.

Os indicadores de privação podem abranger aspectos como acesso a alimentos saudáveis, educação em saúde, ambientes propícios à prática de atividades físicas, entre outros. A privação nesses indicadores pode aumentar a exposição aos fatores de risco das DCNTs, como dieta inadequada, sedentarismo, tabagismo e consumo excessivo de álcool. Portanto, a alta incidência de privações entre os privados de QV na amostra sugere que eles podem enfrentar um fardo maior relacionado às DCNTs, devido a uma combinação de fatores de risco e dificuldades de acesso a cuidados de saúde preventivos e de tratamento.

A terceira estimativa no Quadro 3 é o cálculo da incidência ajustada. Esse cálculo representa a proporção de privação que a amostra privada enfrenta em relação ao máximo de todas as privações possíveis que toda a amostra poderia sofrer. Para calcular essa incidência ajustada, multiplica-se o percentual de pessoas privadas de QV (calculado anteriormente) pelo hiato da privação média. O resultado desse cálculo é de 0,2724, o que significa que o total de indicadores em que os privados de QV na amostra sofrem privação representa 27,24% do máximo possível de privação na amostra em estudo.

No contexto do fardo econômico das DCNTs no Brasil, essa estimativa pode ser interpretada como uma indicação da carga desproporcional dessas doenças sobre os indivíduos em situação de privação. Estudos têm demonstrado que as DCNTs impõem altos custos econômicos, tanto para os sistemas de saúde como para os indivíduos e suas famílias. Os custos diretos incluem tratamento médico, medicamentos e cuidados de longo prazo, enquanto os custos indiretos envolvem perda de produtividade, incapacidade laboral e impactos sociais.

Portanto, a alta incidência de privações entre os indivíduos privados de QV na amostra pode contribuir para agravar o fardo econômico das DCNTs, colocando essas pessoas em maior vulnerabilidade socioeconômica e limitando suas oportunidades de superar a privação.

O foco da análise está na demonstração de oportunidades para superar a privação de QV. Assim, neste estudo, permitiu-se uma análise mais abrangente das condições de vida das pessoas com hipertensão, obesas e diabéticas, possibilitando uma aproximação maior com a intensa e complexa problemática que envolve a vivência da doença crônica, a qual compromete muito além do que apenas a esfera biológica, e que acaba interferindo de diferentes formas no próprio estilo de vida das pessoas acometidas, do seu grupo familiar e social.

Uma pessoa privada em qualidade de vida pode sofrer múltiplas desvantagens ao mesmo tempo – por exemplo, além dos problemas de saúde ou subnutrição, pode sofrer com falta de água potável ou eletricidade, falta de trabalho ou trabalho de má qualidade ou ainda sofrer pelo baixo nível de escolaridade. Logo, concentrar-se apenas num fator, como a saúde, não é suficiente para captar a verdadeira realidade da privação.

Os achados indicam que quase um terço da amostra refere-se a uma população em situação de privação, ou seja, tem a sua qualidade de vida afetada em função da doença combinada com outros indicadores. A elevada incidência de privações entre os indivíduos privados de QV na amostra sugere que essas pessoas enfrentam dificuldades em várias dimensões de suas vidas. Isso pode incluir acesso limitado a serviços básicos, falta de oportunidades educacionais, restrições de emprego, dificuldades no acesso a cuidados de saúde e insegurança alimentar.

Essas privações podem aumentar o risco de desenvolvimento e agravamento das DCNTs. Indivíduos em situação de privação podem ter acesso limitado a

alimentos saudáveis, ambientes propícios à prática de atividades físicas e cuidados de saúde preventivos, contribuindo para a prevalência dessas doenças nessa população.

As descobertas desta pesquisa reforçam a multidimensionalidade do conceito de qualidade de vida. A utilização do método Alkire-Foster ofereceu *insights* valiosos sobre as múltiplas dimensões da privação em qualidade de vida que a população brasileira está submetida face a ocorrência das DCNTs e seu impacto econômico.

A qualidade de vida tem sido objeto de estudo ao longo do tempo, sendo um conceito intensamente marcado pela subjetividade, pois envolve componentes essenciais da condição humana, quer sejam físicos, psicológicos, sociais, culturais ou espirituais. Assim, o significado dessas descobertas corrobora para expandir a compreensão do impacto destas condições nas pessoas, bem como, a busca por formas que superem essa privação.

A estimativa de um índice de qualidade de vida pelo método Alkire-Foster pode somar-se ao conjunto de ações de vigilância das DCNTs, que possibilitam conhecer a distribuição, magnitude e tendência dessas doenças e de seus fatores de risco na população, identificando seus condicionantes sociais, econômicos e ambientais, com o objetivo de subsidiar o planejamento, execução e avaliação da prevenção e controle delas. Logo, no sentido de promover a qualidade de vida, prevenir e controlar as DCNTs, estes achados, com base em históricos passados e suas consequências, a partir de técnicas de *machine learning*, permitem realizar uma análise sobre esses eventos e criar prováveis resultados, auxiliando na criação de estratégias para rastrear as doenças.

As descobertas se relacionam com outros estudos à medida que trazem contribuições que estabelecem uma relação mútua entre dois ou mais elementos. Temas como erradicação da pobreza e da fome, saúde e bem-estar, educação, saneamento, energia acessível e limpa, trabalho decente e crescimento econômico são exemplos de campos de investigação que podem trazer contribuições para superar a privação apontada neste estudo.

Cabe mencionar alguns dos principais limitantes da análise aqui realizada. A adequação do método Alkire-Foster permitiu a reunião de uma série de variáveis influentes de pessoas avaliadas neste estudo. Entretanto, o modelo conceitual construído não representa em totalidade o fenômeno correlacionado à qualidade de

vida. Assim, são necessárias análises complementares, de forma a compor um estudo mais abrangente sobre a situação do impacto econômico destas doenças.

Para fins de acompanhamento da política de vigilância das doenças crônicas, caberiam estudos em perspectiva comparativa – entre instrumentos de medição de qualidade de vida, nacionais ou adaptados⁷.

Desta forma, seria possível discorrer sobre a amplitude dos efeitos que comprometem a qualidade de vida das pessoas com doenças crônicas.

4 CONCLUSÃO

O significado de qualidade de vida abrange em seu cerne, aspectos como bem-estar material, físico e emocional. As doenças crônicas afetam a qualidade de vida quando alteram significativamente as pessoas, limitando-as e privando-as de suas atividades cotidianas. A relação que se estabelece entre as doenças crônicas e a qualidade de vida traz um olhar interdisciplinar para o tema, pois a QV decorre de parâmetros objetivos – cujas referências são a satisfação das necessidades básicas e das necessidades criadas pelo grau de desenvolvimento econômico e social de determinada sociedade; e subjetivos – bem-estar, felicidade, amor, prazer, realização pessoal, que ao faltarem, podem desencadear as DCNTs.

Nos últimos anos, diferentes instrumentos foram construídos para medir qualidade de vida e as discussões que provocam. Ao refletir sobre a área da saúde e sua associação ao conceito de qualidade de vida e à avaliação econômica, novas fronteiras de estudo são descortinadas, surgindo os mais variados instrumentos criados para medi-la nessa referida concepção.

Os resultados dessa pesquisa indicam que existe uma alta proporção de privação de QV entre os indivíduos privados na amostra, o que revela a existência de desigualdades sociais significativas, e sugere que esses indivíduos enfrentam dificuldades em várias dimensões de suas vidas. Isso pode incluir acesso limitado a

⁷ Os mais frequentemente utilizados são: Sickness Impact Profile (SIP), Nottingham Health Profile (NHP), McMaster Health Index Questionnaire (MHIQ), Rand Health Insurance Study (Rand HIS), The Medical Outcomes Study 36-Item Short Form Health Survey (SF-36), Avaliação da Qualidade de Vida da Organização Mundial da Saúde (WHOQOL-100), entre outros (RODRIGUES NETO; FERREIRA, 2022)

serviços básicos, falta de oportunidades educacionais, restrições de emprego, dificuldades no acesso a cuidados de saúde e insegurança alimentar.

Essas privações podem aumentar o risco de desenvolvimento e agravamento das DCNTs, pois indivíduos em situação de privação podem ter acesso limitado a alimentos saudáveis, ambientes propícios à prática de atividades físicas e cuidados de saúde preventivos, contribuindo para a prevalência dessas doenças nessa população.

Destaca-se a importância de abordagens integradas que visam o combate à privação e a redução das desigualdades sociais. É necessário não apenas fornecer assistência direta aos indivíduos privados, mas também abordar as causas estruturais da privação, como falta de acesso a oportunidades educacionais e emprego decente. No contexto brasileiro, em que o fardo econômico das DCNTs é significativo, é fundamental investir em políticas públicas que visem a prevenção e o controle dessas doenças, especialmente entre a população em situação de privação.

Para isso, é necessário fortalecer os sistemas de saúde, garantindo o acesso universal a serviços de qualidade e promovendo ações para acesso à saúde de qualidade e prevenção de doenças. Outras políticas, que visem melhorar as condições socioeconômicas das populações pobres, como programas de transferência de renda, acesso à educação de qualidade e geração de empregos, podem contribuir para reduzir o fardo econômico das doenças crônicas não transmissíveis.

A integração de diferentes setores, como saúde, educação, trabalho e desenvolvimento social, é essencial para abordar de forma eficaz os determinantes sociais da saúde e reduzir as desigualdades existentes. É importante destacar a necessidade de monitorar e avaliar regularmente a eficácia das políticas implementadas, a fim de garantir que estejam alcançando os resultados desejados.

A sensibilização da sociedade também é fundamental para promover uma maior compreensão sobre as desigualdades sociais e o fardo das doenças crônicas não transmissíveis na população em situação de privação. Investir na pesquisa científica e na produção de evidências também desempenha um papel crucial na formulação e implementação de políticas eficazes e baseadas em dados sólidos.

No cenário global, a redução das desigualdades sociais e do fardo das doenças crônicas não transmissíveis é uma prioridade estabelecida pela Organização Mundial da Saúde (OMS) e outros órgãos internacionais. Portanto, é importante

considerar as melhores práticas e lições aprendidas em outros países para orientar as políticas e estratégias nacionais.

A crescente preocupação com questões relacionadas à qualidade de vida vem de um movimento dentro das ciências humanas e biológicas no sentido de valorizar parâmetros mais amplos que o controle de sintomas, a diminuição da mortalidade ou o aumento da expectativa de vida. Assim, qualidade de vida é abordada, por muitos autores, como sinônimo de saúde, e por outros como um conceito mais abrangente, em que as condições de saúde seriam um dos aspectos a serem considerados.

Por fim, a análise evidencia a existência de desigualdades sociais e o fardo econômico das DCNTs na população em situação de privação no Brasil. Para enfrentar esses desafios, são necessárias políticas integradas que abordem as causas estruturais da privação, melhorem as condições socioeconômicas, fortaleçam os sistemas de saúde e promovam ações de prevenção e controle das DCNTs. Somente por meio de abordagens abrangentes e sustentáveis poder-se-á reduzir as desigualdades sociais e promover uma sociedade mais saudável e justa.

REFERÊNCIAS CAPÍTULO 2

- ALBUQUERQUE, M. R.; CUNHA, M. S. Uma análise da pobreza sob o enfoque multidimensional no Paraná. **Revista de Economia**, v. 38, n. 3 (ano 36), p.45- 64, set.\dez.2012. Editora UFPR, 2012.
- ALKIRE, S. Choosing Dimensions: the capability approach and multidimensional poverty. **Munich Personal RePEc Archive (MPRA)**, n. 8862, 2008.
- ALKIRE, S.; FOSTER, J. Counting and multidimensional poverty measurement. **Journal of Public Economics**, v.95, n.7-8, p. 476-487, 2011.
- ALKIRE, S.; FOSTER, J. **Counting and multidimensional poverty measurement**. Oxford: OPHI. (Working Paper, n. 32), 2009.
- ALKIRE, S.; FOSTER, J. **Counting and multidimensional poverty**. Oxford: University of Oxford, 2007. (Oxford Poverty & Human Development Initiative Working Paper, n. 7).
- ALKIRE, S.; SANTOS, M. E. **Acute multidimensional poverty: a new index for developing countries**. Oxford: OPHI. (Working Paper, n. 38), 2010.
- ALKIRE, S.; SANTOS, M. E. Poverty and inequality measurement. In: DENEULIN, S.; SHAHANI, L. **An introduction to the human development and capability approach: freedom and agency**. Londres: Earthscan, p. 121-161, 2009.
- APÓSTOLO, J. et al. Conhecimento, atitudes e práticas dos enfermeiros sobre os instrumentos de avaliação da qualidade de vida. **Revista Latino-Americana De Enfermagem**, v. 26, n. E3017, Coimbra, Portugal, 2018.
- ARISTÓTELES. (350 A.C.). **Ética a Nicômaco**.
- BALLONE, G. J. (2003). Suicídio na adolescência. **PsiquWeb: psiquiatria geral**. Disponível em: <http://www.virtualpsy.org/infantil/suicidio.html>. Acesso em 13.dez.2022.
- BARROS, P.; SILVA, F. B. N. Origem e manutenção do comportamento agressivo na infância e adolescência. **Rev. Bras. Ter. Cogn.**, v. 2, n. 1, p. 55-66, jun. 2006, Rio de Janeiro.
- BARROSO, M. L.; MOURA, A. M. W. A.; PINTO, N. V. (2020). Correlation between general and abdominal obesity in diabetic and/or hypertensive active women. **Research, Society and Development**, v. 9, n. 7, p. e179973679.
- BLOOM, D. E.; CAFIERO, E. T.; JANÉ-LLOPIS, E.; ABRAHAMS-GESSEL, S.; BLOOM, L. R.; FATHIMA, S. The global economic burden of noncommunicable diseases, s. **The Global Economic Burden of Noncommunicable Diseases**. **World Economic Forum**, 2018.
- BRASIL (2021a). Ministério Da Saúde (MS). **Plano de ações estratégicas para o enfrentamento das doenças crônicas e agravos não transmissíveis no Brasil 2021-2030** [Recurso Eletrônico] / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Análise em Saúde e Vigilância de Doenças Não Transmissíveis. Brasília: Ministério da Saúde, 2021.
- BRASIL (2021b). Ministério da Saúde (MS). **Vigitel Brasil 2021: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico: estimativas sobre**

frequência e distribuição sociodemográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal em 2021. Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Análise em Saúde e Vigilância de Doenças Não Transmissíveis. Brasília: Ministério da Saúde, 2021.

BRITES, M.; FERREIRA, T. R. S.; MOURA, A. C.; LANZA, T.; MARIN, S. R. Pobreza Multidimensional nas grandes regiões brasileiras: uma aplicação do método Alkire Foster (AF) (2012). **I Seminário Jovens Pesquisadores 2013**.

COONS, S. J. Et Al. Recommendations on evidence needed to support measurement equivalence between electronic and paper-based patient-reported outcome (PRO) measures: ispor epro good research practices task force report. **Value in Health**, v. 10, n. s2, p. s121-s129, 2007.

DIENER, E.; OISHI, S.; TAY, L. Advances in subjective well-being research. **Nature Human Behaviour**, v. 2, n. 4, p. 253-260, 2018.

DURKHEIM, É. (1897). **O Suicídio**: Estudo de Sociologia. Editora Vozes.

FARQUHAR, M. Definitions of quality of life: a taxonomy. *Journal of Advanced Nursing*, Oxford, v.22, n.3, p.502-8, 1995.

FERES, J. C.; VILLATORO, P. **A viabilidade de se erradicar a pobreza: uma análise conceitual e metodológica**. Caderno de estudos. Desenvolvimento social em debate. 2013, p. 32.

FLECK, M. P. A.; LEAL, O. F.; LOUZADA, S.; XAVIER, M.; CHACHAMOVICH, E.; VIEIRA, G.; SANTOS, L.; PINZON, V. Desenvolvimento da versão em português do instrumento de avaliação de qualidade de vida da OMS (WHOQOL-100). **Brazilian Journal of Psychiatry**, v.21, n.1, p. 19–28, 1999.

GALLO, C. R.; ROCHE, J. M. **Las dimensiones de la pobreza em Venezuela y sus câmbios entre 1997 y 2010**: propuesta de uma medida multidimensional. Banco Central de Venezuela (BCV) 2011.

GÓMEZ, M.; SABEH, E. (2001). Calidad de vida: evolución del concepto y su influencia en la investigación y la práctica. **Salamanca: Instituto Universitario de Integración en la Comunidad**, Faculty of Psychology, University of Salamanca.

HUPPERT, F. A.; COOPER, C. L. The state of well-being science: concepts, measures, interventions, and policies. In: Huppert, F. A.; Baylis, N.; Keverne, B. (Eds.). **The Science Of Well-Being**. Oxford: Oxford University Press, 2014.

IBGE. Pesquisa Nacional por Amostra de Domicílios (PNAD). **Instituto Brasileiro de Geografia e Estatística**, 2021. Disponível Em: <https://www.ibge.gov.br/estatisticas/Sociais/Populacao/9127-Pesquisa-Nacional-Por-Amostra-De-Domicilios.html>.

JIA, H. et al. The impact of socioeconomic status on health-related quality of life among chinese urban residents: the mediating role of physical and psychological factors. **Quality Of Life Research**, v. 27, n. 9, p. 2297-2306, 2018.

KAGEYAMA, A.; HOFFMANN, R. Pobreza no Brasil: uma perspectiva multidimensional. **Economia e Sociedade**, v. 15, n. 1, p. 79-112, 2006.

KANT, I. (1785). **Fundamentação Da Metafísica Dos Costumes**.

KERSTENETZKY, C. L.; VECCHIO, R. D.; CARVALHO, M. M. de. Uma metodologia para estimação da pobreza multidimensional aplicada às regiões metropolitanas brasileiras – 2003 e 2008. **Center for Studies on line Quality and Development**, Niteroi, 2011.

KEYES, C. L. M. et al. (2002). The mental health continuum: from languishing to flourishing in life. **Journal of Health and Social Behavior**, v. 43, n. 2, p. 207-222, 2002.

LOCKE, J. (1689). **Ensaio sobre o entendimento humano**.

MINAYO, M. C. S.; HARTZ, Z. M. A.; BUSS, P. M. (2000). Qualidade de vida e saúde: um debate necessário. **Ciência & Saúde Coletiva**, v. 5, n. 1, p. 7–18.

OMS. Organização Mundial da Saúde. (2018). Relatório: *Saving lives, spending less: A strategic response to noncommunicable diseases*. **World Health Organization**. <https://iris.who.int/bitstream/handle/10665/272534/WHO-NMH-NVI-18.8-eng.pdf>

OMS. Organização Mundial da Saúde. (2002). World Health Survey (WHS). **World Health Organization**. Disponível em: <https://www.who.int/healthinfo/survey/whs/en/>. Acesso em: 19.jan.2022.

OMS. Organização Mundial da Saúde. (1994). **WHOQOL-BREF**: introduction, administration, scoring and generic version of the assessment: field trial version. Geneva: WHO, 1994.

OPAS – Organização Pan-Americana da Saúde. **Cuidados inovadores para condições crônicas**: organização e prestação de atenção de alta qualidade às doenças crônicas não transmissíveis nas Américas. Washington, DC: 2015.

PALOMINO, B.; LÓPEZ, G. Reflexões sobre a qualidade de vida e desenvolvimento. **Região e Sociedade**, v. 11, n. 17, p. 171-185, 1999.

PATRICK, D. L.; ERICKSON, P. **Health status and health policy**: quality of life in health care evaluation and resource allocation. Oxford: University Press, 2016.

PNUD. Programa das nações unidas para o desenvolvimento. **Índice de Desenvolvimento Humano** (IDH). 2020. Disponível em: <Http://Hdr.Undp.Org/En/Indicators/137506>. Acesso em: 15.mar.2023.

QUARTOS, C.; GARZÓN, M.O. A noção de qualidade de vida e sua medição. **Rev CES Salud Pública**, v. 4, n. 1, p. 36-46, 2013.

REVICKI, D. A. et al. Recommendations on health-related quality of life research to support labeling and promotional claims in the United States. **Quality of Life Research**, v. 15, n. 6, p. 1167-1177, 2006.

ROTH, G. A.; ABATE, D.; ABATE, K. H.; ABAY, S. M.; ABBAFATI, C.; ABBASI, N.; MURRAY, C. J. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the global burden of disease study 2017. **The Lancet**, v. 392, n. 10159, p. 1736-1788, 2020.

SEN, A. K. Poverty: an ordinal approach to measurement. **Econometrica**, v. 44: 219-231, 1976.

THOITS, P. A. Mechanism is linking social ties and support to physical and mental health. **Journal of Health and Social Behavior**, v. 52, n. 2, p. 145-161, 2011.

WEBER, M. (1904-1905). A Ética Protestante E O Espírito Do Capitalismo.

WORLD HEALTH ORGANIZATION. (2018). Noncommunicable diseases country profiles 2018. **World Health Organization**.

CAPÍTULO 3 – MACHINE LEARNING PARA ANÁLISES PREDITIVAS EM DOENÇAS CRÔNICAS NÃO TRANSMISSÍVEIS: UMA REVISÃO SISTEMÁTICA DE LITERATURA

RESUMO

As doenças crônicas não transmissíveis (DCNTs) são as principais causas de óbitos, perda de qualidade de vida (QV), limitação no trabalho e, por conseguinte geram expressivos impactos econômicos. Visando a redução da morbidade e mortalidade, sistemas baseados em inteligência artificial (IA) são desenvolvidos para auxiliar profissionais da saúde na tomada de decisões. Este estudo buscou identificar o conhecimento produzido sobre o uso de modelos de predição para DCNTs, verificando a elite do *corpus* textual, distribuída segundo a reputação de periódicos em nove países. Em relação à expressão *Prediction*, verificou-se a ligação à expressão *Risk-Factors*, demonstrando que a maior parte dos estudos foram conduzidos neste tipo de aglomerado. Encontrou-se a formulação de seis clusters, destacando-se três proeminentes. Foi realizado o mapeamento e análise dos estudos na área de Administração da Informação (ADI) com ênfase em Tecnologia e Saúde, das bases de dados SCOPUS e WEB OF SCIENCE. As análises foram realizadas através das três leis da bibliometria: Lotka, Bradford e Zipf. A análise de caráter descritiva, do *corpus* textual, revelou um total de quarenta artigos, dispostos nesta RSL, com auxílio dos *softwares* Histcite, Iramuteq e VOSviewer, organizados em mapeamentos e nas análises do estudo.

Palavras-chave: Modelos Preditivos, Administração da Informação, Tomada de Decisão, Tecnologia e Saúde.

ABSTRACT

Non-communicable chronic diseases (NCDs) are the main causes of death, loss of quality of life (QoL), limitation at work and economic impacts in Brazil and worldwide. Aiming at reducing morbidity and mortality, systems based on artificial intelligence (AI) are developed to assist health professionals in decision-making. This study sought to identify knowledge about the use of prediction models for NCDs, verifying the elite of the textual corpus, distributed according to the reputation of journals, in nine countries. With regard to the expression *Prediction*, there was a link to the expression *Risk-Factors*, demonstrating that most of the studies were conducted in this type of cluster. The formulation of six clusters was found, with three prominent ones standing out. Studies in the area of Information Management (ADI) with an emphasis on Technology and Health, from the SCOPUS and WEB OF SCIENCE databases, were mapped and analyzed. Analyzes were performed using the three laws of bibliometrics: Lotka, Bradford and Zipf. The descriptive analysis of the textual corpus revealed a total of forty articles, arranged in this RSL, with the help of the Histcite, Iramuteq and VOSviewer software, organized in mappings and in the study analyses.

Keywords: Predictive Models, Information Management, Decision Making, Technology and Health.

1 INTRODUÇÃO

Globalmente, a cada cinco mortes ocorridas, três são atribuídas às doenças crônicas não transmissíveis (DCNTs) (Hajat & Stein, 2018), principalmente aquelas ocorridas em razão das quatro DCNTs que mais acometem as populações, a saber, as doenças crônicas cardiovasculares, doenças pulmonares crônicas, diabetes e câncer (Wang et al., 2016; Wang & Wang, 2020; Davagdorj et al., 2021; Ferdousi et al., 2021; Wang et al., 2021). De acordo com a Organização Mundial da Saúde (OMS), as DCNTs envolvem fatores de risco, principalmente controláveis e modificáveis (Ferdousi et al., 2021).

Entre os países de baixa renda, esse número é crescente e mais elevado que em outros países, o que leva os governos e sociedade a preocuparem-se com o desenvolvimento da prevenção, buscando formas de gerenciar o avanço das DCNTs na população e nos sistemas de saúde (Wang et al., 2016). Nesta seara, até mesmo órgãos mundiais como as Organizações das Nações Unidas (ONU, *apud* PAHO¹, 2022) e Pan American Health Organization (PAHO², 2022) mantêm agendas para acompanhar a evolução das DCNTs e criar formas de monitorá-las eficientemente.

Ações de enfrentamento às DCNTs têm sido direcionadas ao estímulo de alimentação saudável, promoção da inclusão da atividade física nos hábitos salutaros, e prevenção do uso de tabaco e álcool como formas de diminuir o acometimento por essas doenças (Davagdorj et al., 2021; Wang et al., 2021). Outras ações são colocadas em prática, além das maneiras convencionais de prevenção, controle e acompanhamento das DCNTs na saúde pública de cada país.

No Brasil, as DCNTs são monitoradas anualmente pela Secretaria de Vigilância do Ministério da Saúde, através da pesquisa VIGITEL⁸, um inquérito telefônico decorrente de um acordo entre a OMS e governo brasileiro firmado em 2006. Apesar disso, há deficiências nesse método, decorrentes da falta de personalização dos pacientes, baixo retorno de respostas, da recusa em responder a esse tipo de abordagem e da dificuldade em conjugar as informações de forma aplicada, com um sistema de vigilância efetivo e representativo.

⁸ Realizado pelo Ministério da Saúde, o VIGITEL - Vigilância de Fatores de Risco e Proteção para Doenças Crônicas em Inquérito Telefônico. monitora a frequência e a distribuição dos principais fatores de risco e proteção associados à Doenças Crônicas Não Transmissíveis (DCNT) em todas as capitais dos 26 estados brasileiros e o Distrito Federal.

Soluções ciberfísicas utilizam as informações geradas no mundo físico, conjugando-as às tecnologias IoTs (*Internet of Things*), buscando uma comunicação perfeita entre as informações (Ferdousi et al., 2021). Sistemas integrados são criados, envolvendo diferentes modelos de predição de DCNTs (Davagdorj et al., 2021; O'Donovan et al., 2020; Davagdorj et al., 2020; Morgenstern et al., 2020; Carvalho et al., 2019), muitos deles disponibilizados aos usuários finais sem expor de que forma são realizados, e quais parâmetros são considerados na sua construção (Carvalho et al., 2019; Elshawi et al., 2019).

Novas tecnologias, como o *Machine Learning* (Aprendizado de Máquinas), possibilitam a realização de predição das DCNTs com maior qualidade para a gestão da saúde (Ichikawa et al., 2017). Como uma forma de Inteligência Artificial (IA), a *Machine Learning* é o meio que facilita a aprendizagem, pelos computadores, a respeito de dados que sejam recebidos automaticamente, transformando-os em dados processados, capazes de encontrar soluções ou respostas mais assertivas (Netto et al., 2021b).

Dessa forma, pergunta-se: qual é o estado da arte das pesquisas envolvendo modelos de predição de DCNTs? Propõe-se, como objetivo, identificar o estado da arte das pesquisas envolvendo modelos de predição de DCNTs. Evidencia-se que este estudo espera contribuir para o crescimento do conhecimento sobre novos e usuais métodos de predição para as DCNTs, mostrando um compêndio de estudos relacionados ao tema como estado da arte das pesquisas.

Destaca-se que, antes de realizar uma proposta inovadora em qualquer campo do conhecimento, é necessário estudar o que tem sido proposto no rol científico. E, para investigar esse conhecimento foi escolhido o método quantitativo de Revisão Sistemática de Literatura (RSL) que, segundo Chung et al. (2006), é a forma científica apropriada para que se resuma a literatura, com protocolos determinados para que, a partir de critérios, se determinem quais estudos deverão compor a revisão.

1.1 *Machine learning* para análises preditivas em DCNTs

A saúde pública está baseada em aspectos sociodemográficos, determinantes para estabelecer padrões gerais de saúde, assim como do bem-estar das populações.

Os dados coletados pelos sistemas de saúde sobre as condições populacionais utilizam sistemas diversificados, nem sempre integrados e, em alguns casos, com baixa efetividade no aproveitamento relevante das informações geradas (LUO et al., 2015).

O uso de tecnologias de IA na área da saúde tem sido evidenciado através das inúmeras soluções aplicáveis em análise preditiva, que utilizam dados e percepções de agentes ambientais para executar ações de diagnóstico e de prevenção de doenças, além de otimizar o desenvolvimento de diversas atividades relacionadas à gestão de saúde (DAVAGDORJ et al., 2021; DELPINO et al., 2022). Dentre as diferentes subáreas de IA, as que se relacionam ao *Machine Learning* (ML) envolvem o aprendizado de aplicações que processam, reconhecem e detectam padrões nos dados gerados pelos médicos, enfermeiros, atendentes, pesquisadores e hospitais (NETTO et al., 2021b; MORGENSTERN et al., 2020). *Insights* de análises preditivas auxiliam na tomada de decisão, minimizando as possibilidades de incertezas e podem ser capazes de encontrar padrões entre os diversos dados coletados. Com base em históricos passados e suas consequências, é possível realizar uma análise sobre esses eventos e criar prováveis resultados estimados (NETTO et al., 2021b). Também auxiliam no desenvolvimento das políticas públicas de saúde, principalmente em regiões em que há recursos limitados nesse setor (BARSASELLA et al., 2021).

A análise preditiva compreende uma série de aplicações que empregam algoritmos para entender o conjunto de dados que já existem, gerando padrões a partir deles para a predição (O'DONOVAN et al., 2020). A construção dessa análise pode envolver ambientes supervisionados (com preditores adicionados a respostas de interesse) ou não supervisionados (apenas com preditores), que podem aumentar a estimativa de risco para determinadas características relacionadas à saúde pública (SANTOS, 2018). Para o diagnóstico ou prognóstico, os modelos preditivos auxiliam na criação de estratégias para rastrear as doenças e podem ajudar a estabelecer intervenções, ou mesmo, colaborar para a construção de propostas de tratamentos profiláticos e prognósticos para as DCNTs (MORGENSTERN et al., 2020).

A modelagem preditiva tradicional em saúde e área médica populacional utiliza modelos de regressão estatísticos paramétricos, que consideram a distribuição específica e linear dos dados (MORGENSTERN et al., 2020; ASGARI et al., 2021). Entretanto, os métodos de aprendizagem que utilizam modelos preditivos na

atualidade derivam dos modelos de *machine learning* lineares, ou seja, de algoritmos simples, avançando para modelos de regressão linear, que permitem desfechos contínuos (variáveis dependentes) (TIWARI et al., 2022; MORGENSTERN et al., 2020; SANTOS et al., 2019).

Diferentes particularidades são aplicadas para os modelos de *machine learning*, visando flexibilizar sua utilização na previsão de riscos que estejam relacionados à saúde e medicina (ASGARI et al., 2021; FENG; JIAO, 2021). O fomento de aplicações de modelos de aprendizagem (ML), que permitam intervenção precoce, permite a melhor tomada de decisão e escolha de tratamentos, através dos resultados decorrentes dos modelos de aprendizagem de máquina (HAJAT; STEIN, 2018). Nesses modelos, supõe-se que os preditores possam variar e serem agrupados por determinadas variáveis, que de forma potencialmente dinâmica, possam produzir resultados que sirvam para desenvolver novos sistemas de saúde mais eficientes.

2 PROCEDIMENTOS METODOLÓGICOS

Para contribuição científica dos estudos pautados na área de Administração da Informação (ADI), com ênfase em Tecnologia e Saúde, pretendeu-se a realização desta pesquisa, visto que o desenvolvimento técnico da RSL permite a identificação, o mapeamento e, finalmente, a análise das pesquisas que possuem relevância frente a um problema ou tópico de pesquisa que está objetivado no estudo específico. Desta forma, se oportunizam maiores atributos em discussões, se confrontados aos estudos tradicionais que são populares como as revisões bibliográficas (TRANFIELD et al.2003; KITCHENHAM, 2004; BIOLCHINI et al.2007).

Sendo assim, as revisões sistemáticas literárias possuem certo rigor metodológico, pois o desenvolvimento de uma RSL ocorre através da elaboração de um protocolo de pesquisa, de maneira a impor viés de legitimidade ao estudo e ofertar uma boa base de conhecimento estruturado para as futuras tomadas de decisão. O passo a passo da RSL é disposto de três fases, as quais podem ser identificadas sucessivamente como: Planejamento da Revisão, Condução da Revisão e, para finalizar, Disseminação do Conhecimento (TRANFIELD et al., 2003).

2.1 Primeira Fase: Planejamento da Revisão

A primeira fase está baseada na seleção dos especialistas do estudo sobre o tema principal, sendo a fase que define o corpus de pesquisa. A definição do corpus é dada como um conjunto textual que é selecionado e organizado a fim de expressar algum sentido para a linguagem (BAUER; AARTS, 2000).

Segundo Bauer e Aarts (2000), o corpus da pesquisa é muito importante, pois é baseado nele que são extraídos os atributos desenvolvidos de maneira quantitativa, bem como as reproduções de vieses qualitativo a partir do conteúdo analisado. A primeira etapa concebe a elaboração do protocolo de pesquisa, com objetivo de dar sentido às etapas realizadas ao longo da RSL. O protocolo de pesquisa é composto de: elementos como o problema de pesquisa, a população e amostra, ele também define qual estratégia foi adotada para a inclusão e exclusão dos estudos identificados na Revisão Sistemática Literária (TRANFIELD et al., 2003).

O corpus textual deste estudo foi determinado através da execução de quatro etapas, sendo estas delineadas no que recomendam Almeida e Goulart (2017), com o intuito principal da minimização da seleção dos artigos. As etapas são compreendidas em aplicações de filtros aos artigos. Estes filtros são aplicados quanto aos idiomas, e quanto à reputação dos periódicos, com a finalidade de assegurar a qualidade dos estudos e para finalizar a análise de conteúdo, que é realizada por dois revisores por período.

2.2 Segunda Fase: Condução da Revisão

A segunda fase é composta de uma busca abrangente e justa (TRANFIELD et al., 2003), em que esta é realizada através do auxílio de operadores de busca, os quais direcionam a pesquisa através da identificação de palavras-chave, sendo essas relacionadas aos eixos temáticos que são definidos na hora da busca. A estratégia de busca consistiu em pesquisar publicações científicas que têm relevância na área temática, com o auxílio das palavras-chave “*non-communicable diseases*” e “*prediction models*”, sendo estas direcionadas ao problema de pesquisa deste estudo.

O período de levantamento de dados (os artigos) foi dos últimos 10 anos (2012 a 2022), em que o ponto de partida no ano de 2012 condiz como início comum para

as bases de dados SCOPUS (*Elsevier*TM) e WEB OF SCIENCE (*Clarivate Analytics*TM), as quais são utilizadas nesta Revisão Sistemática de Literatura (RSL).

Justifica-se a escolha das bases de dados SCOPUS (*Elsevier*TM) e WEB OF SCIENCE (*Clarivate Analytics*TM), devido ao fato delas abrangerem diversas bases de periódicos como as dispostas a seguir: *Arts & Humanities Citation Index, Bentham Science, Cambridge University Press, Cambridge University Press, Conference Proceedings Citation Index (Science), Conference Proceedings Citation Index, Current Contents Connect, Derwent Innovations IndexSM, Elsevier, Emerald, Emerging Sources Citation Index, IEEE Xplore, InderScience Publishers, KCI, Nature Publishing Group, Oxford University Press, Oxford University Press, Russian Science Citation Index, SAGE, SciELO Citation Index, Science Citation Index Expanded, Social Science & Humanities, Social Sciences Citation, Springer, Taylor & Francis Group, Wiley-Blackwell, Wolters Kluwer* (SCOPUS, 2017; Clarivate Analytics, 2018).

A Scopus e a Web of Science, são bases de dados multidisciplinares que cobrem a literatura acadêmica de quase todas as disciplinas. A Scopus é o maior o maior banco de dados de resumos e citações da literatura com revisão por pares e a Web of Science dá acesso a múltiplos bancos de dados que, quando combinados, totalizam mais de 33 mil periódicos com foco em ciências da vida, ciências biomédicas, engenharia, ciências sociais, artes e humanidades, com cobertura de 1899 em diante.

Sendo assim, toma-se a decisão de optar somente por artigos científicos devido ao fato destes serem capazes de demonstrar as fronteiras do conhecimento. Após uma gama abrangente de testes, os termos de busca selecionados para esse estudo foram: (“*non-communicable*” AND “*diseases*” AND “*prediction*” AND “*models*”). Como escopo desse estudo foram selecionados os títulos, as palavras-chave e o resumo dos artigos. A busca desta pesquisa ocorreu no mês de outubro do ano de 2022. Por fim, têm-se as seguintes *Strings* dessa pesquisa na base de dados SCOPUS (*Elsevier*TM) e WEB OF SCIENCE (*Clarivate Analytics*TM), conforme Quadro 4:

BASES	STRINGS DE BUSCA
SCOPUS (Elsevier™)	TITLE-ABS-KEY ((non-communicable AND diseases AND prediction AND models)) AND (LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2015) OR LIMIT-TO (PUBYEAR, 2014) OR LIMIT-TO (PUBYEAR, 2013) OR LIMIT-TO (PUBYEAR, 2012)) AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English"))
WOS - WEB OF SCIENCE (Clarivate Analytics™)	TOPICO ((non-communicable AND diseases AND prediction AND models)) AND (LIMIT-TO (PUBYEAR, 2022) OR LIMIT-TO (PUBYEAR, 2021) OR LIMIT-TO (PUBYEAR, 2020) OR LIMIT-TO (PUBYEAR, 2019) OR LIMIT-TO (PUBYEAR, 2018) OR LIMIT-TO (PUBYEAR, 2017) OR LIMIT-TO (PUBYEAR, 2016) OR LIMIT-TO (PUBYEAR, 2015) OR LIMIT-TO (PUBYEAR, 2014) OR LIMIT-TO (PUBYEAR, 2013) OR LIMIT-TO (PUBYEAR, 2012)) AND (LIMIT-TO (DOCTYPE, "ar")) AND (LIMIT-TO (LANGUAGE, "English"))

Quadro 4 – Strings de busca

Fonte: elaborado pelo pesquisador (2023).

A primeira etapa caracteriza-se pela realização de buscas em bases de dados. Os critérios adotados para essa primeira etapa apontam: artigo como o tipo de documento, sendo os artigos limitados ao período de 2012 a 2022, bem como ao idioma inglês. Ao final desta etapa foram encontrados 112 artigos na plataforma SCOPUS (Elsevier™) 65 artigos na plataforma WOS - WEB OF SCIENCE (Clarivate Analytics™), totalizando 177 artigos.

A segunda etapa caracterizou-se pela exclusão de 18 estudos duplicados através do *software Histcite*. Sendo assim, ao finalizar esta etapa restaram 159 artigos.

A terceira etapa caracteriza-se em verificar a aderência dos estudos, no que diz respeito a cada área contextual, focando assim na pergunta de pesquisa. Para este fim, a seleção dedicou-se a leitura dos resumos, palavras-chaves no corpus dos textos, do método – aqui em especial, se o tipo de algoritmo se utilizava do *machine learning* – e, para quais fatores de risco e doenças destinavam-se. Dessa forma, foram excluídos 78 artigos.

Finalmente, a quarta etapa caracteriza-se na reputação dos periódicos como critério de seleção. Todas as etapas são mostradas na Figura 1.

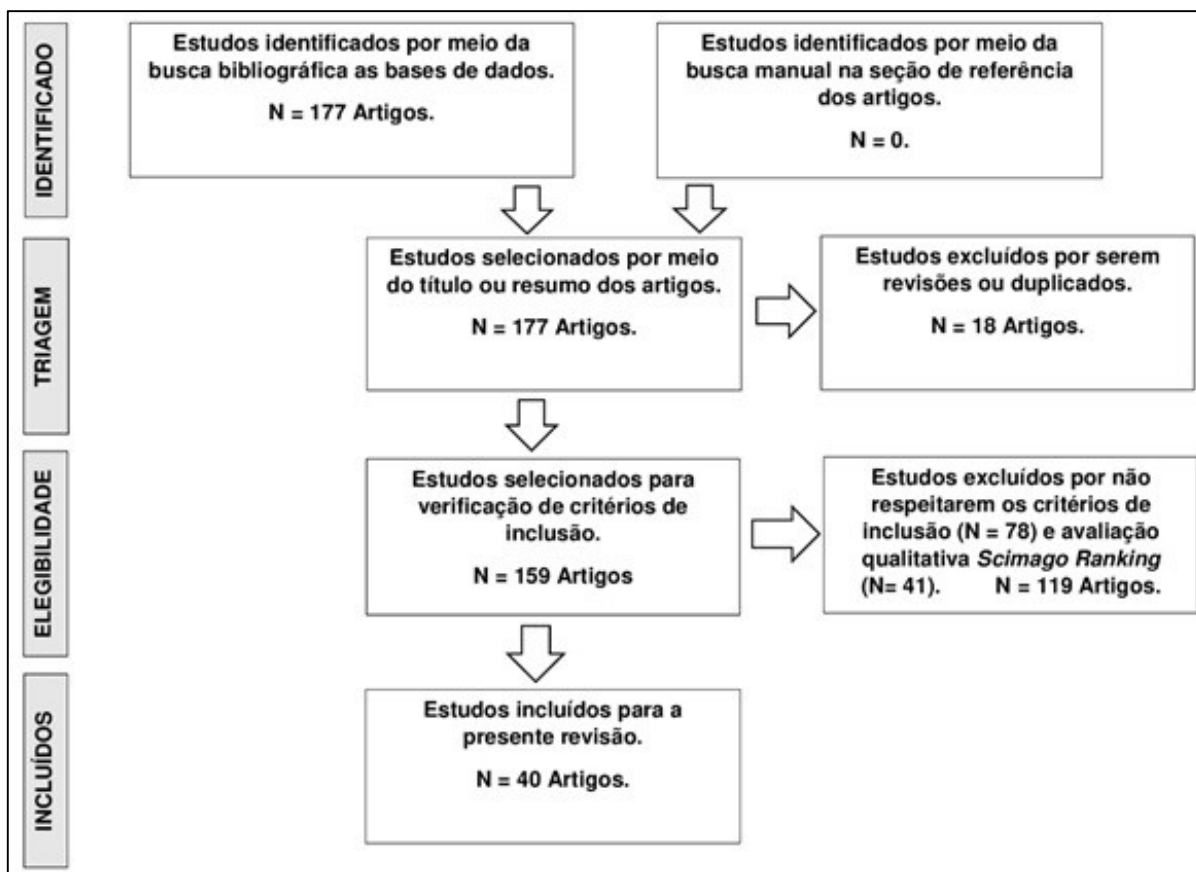


Figura 1 – Condução da Revisão Sistemática de Literatura

Fonte: elaborado pelo autor (2023).

Dessa maneira, na quarta etapa, foram inseridos no corpus textual do estudo somente artigos publicados em periódicos que compõem o primeiro quartil de citação, de acordo com o *Scimago Journal Ranking* (SJR, 2021). Assim, nesta etapa, foram excluídos 41 artigos, restando uma amostra final do corpus textual da pesquisa de 40 artigos.

O levantamento dos artigos nas bases foi realizado na data de 12 de outubro de 2022, a seleção dos artigos deste estudo realizou-se no período de 28 de outubro a 27 novembro de 2022 e contou com a presença simultânea de dois especialistas na área de ciências sociais e aplicadas. O *download* dos dados junto às bases foi realizado no dia 12 de outubro de 2022. O *corpus* da pesquisa foi disposto em planilha eletrônica (Excel), destacando os principais elementos de cada artigo individualmente, que contribuíram para a análise deste estudo. Foram extraídos dos artigos indicadores como de citação e conteúdo.

Os indicadores de citação foram codificados em sete elementos: i) ano em que o artigo foi publicado; ii) periódico vinculado; iii) título do artigo; iv) número de citações

do artigo; v) nome dos autores no artigo; vi) quantidade de autores e; vii) país de atuação dos autores do estudo. Os indicadores de conteúdo contemplavam os elementos analisados nos estudos, reunindo um total de doze características: i) palavras-chave citadas no artigo; ii) objetivo do artigo; iii) contribuição do artigo; iv) temática relacionada as áreas contextuais: cultura organizacional, conhecimento, gestão e tecnologia; v) abordagem quantitativa, qualitativa ou mista; vi) tipo de classificação metodológica do estudo; vii) procedimento de coleta de dados realizado; viii) uso de dados primários ou secundários; ix) variáveis ou categorias de análise utilizadas; x) principais resultados do estudo; xi) limitações da pesquisa e; xiii) sugestões para futuras pesquisas.

2.3 Terceira Fase: Disseminação do Conhecimento

A terceira fase da RSL consiste na disseminação dos resultados obtidos nesse estudo. Tranfield et al. (2003) apontam que os resultados tendem a divagar conhecimento. Todavia, o rigor imposto na elaboração da revisão sistemática dispõe que o conhecimento é disseminado em duas subfases.

2.3.1 Primeira subfase: Análise Detalhada

A primeira subfase consiste em fornecer uma análise de caráter detalhado sobre as características dos artigos que compuseram o corpus textual dessa pesquisa, sendo essa análise disposta de nove itens, conforme metodologia sugerida por Chung et al. (2006), que são caracterizados por: a) distribuição anual do corpus da pesquisa; b) análise do vínculo institucional dos autores e coautores; c) análise da composição da autoria dos artigos; d) elite de pesquisa; e) relacionamento entre os autores dos artigos; f) estatística da quantidade de citações recebidas pelos autores; g) relacionamento entre os autores mais cocitados; h) palavras mais relevantes encontradas no título, resumo e palavras-chave dos artigos do corpus; i) e relacionamento entre as palavras-chave nos estudos que compõe o corpus.

O *software HistCite™* foi o escolhido para a realização das análises de letra (a) até letra (d). O *software* tem por função gerar mapas cronológicos de coleções

bibliográficas que resultam das buscas por assunto, autoria e publicações institucionais. Segundo Bornmann & Marx (2012), os arquivos exportados foram criados de modo que todas as referências citadas para cada documento de origem pudessem ser capturadas. O *HistCite*[™] busca em suas funções criar historiografias cronológicas, destacando os trabalhos mais citados.

Os autores Bornmann & Marx (2012) ainda retomam que podem ser feitas listagens por classificações de autor, revista, instituição ou vocabulário. As análises de letra **(e)** até a letra **(g)**, contaram com ajuda do *software* VOSviewer. Jeyasekar & Saravanan (2015) delineiam que essa ferramenta é especificamente projetada para construir e visualizar mapas bibliométricos, tendo atenção voltada para a representação dos gráficos de mapas.

Por fim, as análises de letra **(h)** e **(i)**, foram projetadas com o auxílio de dois *softwares*: o VOSviewer⁹ e o IRAMUTEQ¹⁰ (*Interface de R pour les Analyses Multidimensionnelles de Textes et de Questionnaires*). Camargo & Justo (2013) caracterizam que esse *software* é disposto no *software* R¹¹, que permite uma variabilidade de análises do *corpus*.

2.3.2 Segunda subfase: Análise Temática

A segunda subfase da etapa de disseminação do conhecimento prepara a análise mais aprofundada do *corpus*, a análise temática. Esta parte foi elaborada através da filtragem de trabalhos que apresentaram os modelos de *machine learning* utilizados nos estudos para identificar fatores de riscos ou doenças crônicas que representavam temáticas de pesquisas similares encontradas nos 40 artigos.

Nesta etapa também é elaborada a análise das potenciais variáveis, que podem ser usadas para elaborar futuras pesquisas sobre a temática do estudo. A criação dos clusters¹² e categorias amplia descrição minuciosa das contribuições, ressaltando partes relevantes que estão no *corpus* com base na literatura, sendo isso confirmado por Tranfield et al. (2003).

⁹ Permite a criação de mapas bibliométricos, como: de autorias, de títulos, resumos, palavras-chave, periódicos e artigos. (VAN ECK; WALTMAN, 2010).

¹⁰ Permite identificar a Leide Zipf e Lotka (VAN ECK; WALTMAN, 2010).

¹¹ Este *software* é ligado ao pacote estatístico R, serve para análise de conteúdo, lexicometria e análise de discurso, utilizado para fazer árvore da similitude, nuvem de palavras e análise estatística de palavras (VAN ECK; WALTMAN, 2010).

¹² Os clusters foram criados utilizando-se o *software* VOSviewer para ter uma análise dos grupos de cocitação. Para isso, foram utilizadas palavras relacionadas com o título e abstract dos artigos.

3 APRESENTAÇÃO E ANÁLISE DOS RESULTADOS

Esta seção objetiva apresentar a análise do corpus textual da RSL. As análises aqui realizadas estão subsidiadas nas três leis clássicas da bibliometria: a) Lei de Lotka (1926), que tem por objetivo avaliar o grau de produtividade dos autores; b) Lei de Bradford (1953), tem por finalidade mensurar a produtividade que dispõe os periódicos; c) Lei de Zipf (1949), com o intuito de mensurar a frequência de ocorrência e coocorrência de determinadas palavras que estão dispostas no texto.

A análise descritiva do *corpus* textual possui indicadores que estão relacionados aos indicadores de citação e de conteúdo, os quais foram reportados no percurso metodológico deste estudo. Sendo assim, os itens aqui avaliados foram a quantidade de artigos publicados ao longo tempo, a distribuição geográfica de cada autor e coautor, a composição de autoria nos trabalhos pesquisados e, finalmente, a estimativa da elite de pesquisa que compõe o corpus textual. São evidenciadas as análises que se referem ao acoplamento bibliográfico dos autores, sendo estes baseados na quantidade de citações que são retornadas.

O objetivo dessa análise é buscar identificar e, ao mesmo tempo, aprofundar o conhecimento sobre um determinado grupo de pesquisadores ao longo do tempo. Considerando que estes contribuem para a produção científica, em aglomerados, ou seja, em cooperação mútua por meio de artigos que são elaborados e assinados em conjunto. Além disso, essa análise se propõe a acompanhar o desenvolver das áreas do conhecimento e os padrões de publicação e autoria (HOU et al., 2018).

A reputação dos periódicos também foi avaliada, por ser um fator importante para a RSL, tendo como objetivo conhecer a qualidade dos periódicos que estão no estudo. Este dado é feito a partir dos indicadores de citação, buscando fornecer ao leitor do estudo uma situação de ajuda na apresentação da RSL. Outro item importante que se destaca é a análise das palavras, que consiste em verificar os índices do coocorrência, sendo uma análise que pode ser feita através da construção da nuvem de palavras, bem como pela análise da similitude entre as palavras. Para este fim, foram utilizados os artigos que compuseram o corpus textual. A partir daí, trata-se da distribuição anual do corpus da pesquisa, com fins de analisar os índices de produtividade e mais.

O *corpus* textual deste estudo corresponde ao total amostral de 40 artigos, dispostos em 27 periódicos ou *Journals*. Logo, contribuíram para essa amostra um total de aproximadamente 239 autores e coautores, sendo estes registrados nas bases de coleta de dados SCOPUS (*Elsevier™*) e WEB OF SCIENCE (*Clarivate Analytics™*).

A Figura 2 caracteriza a ocorrência de publicações ao longo dos anos, a quantidade e a tendência. Destaca-se, com base nos critérios de inclusão, que não foram identificadas publicações nos anos de 2013, 2016 e 2019.

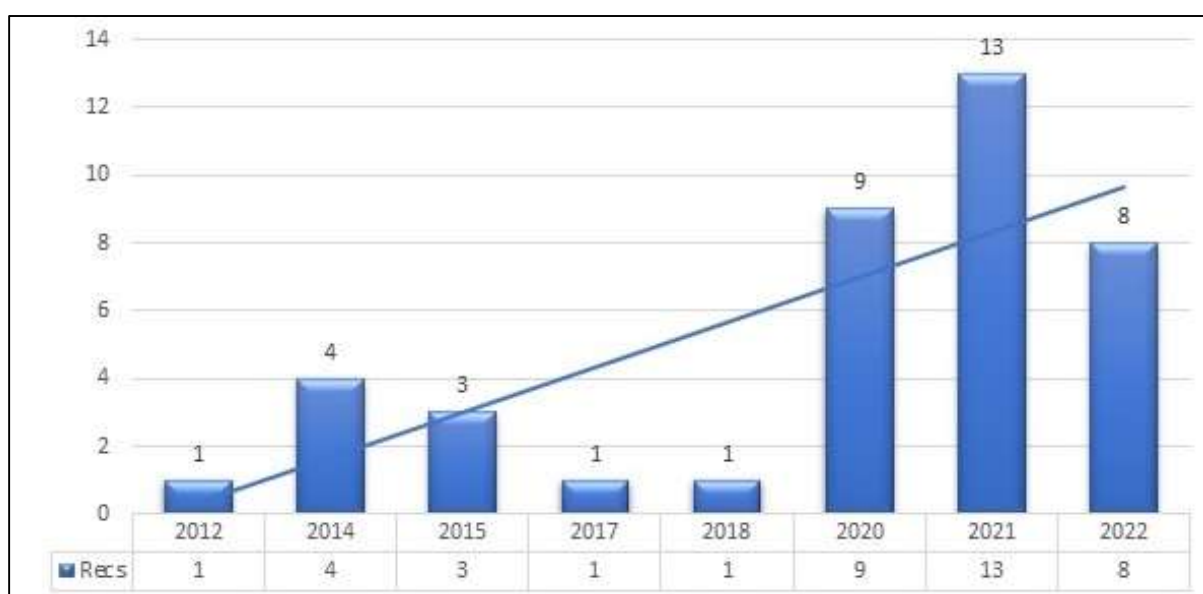


Figura 2 – Distribuição anual do corpus textual em número de artigos

Fonte: elaborado pelo autor (2023).

Em 2012, o estudo de Webber et al. (2012) intitulado: *High rates of obesity and non-communicable diseases predicted across L*, pioneiro, já elencava as DCNTs como sendo uma grande preocupação de saúde pública para toda a América Latina. O estudo deixa claro que um fator de risco modificável chave para DCNTs é o sobrepeso e a obesidade, destacando a necessidade de políticas para reduzir as taxas de prevalência e melhorar os níveis crescentes de DCNTs. Desta forma, na busca de um prognóstico, o estudo fez uma análise de regressão transversal para projetar o IMC e as tendências de doenças relacionadas até 2050. Testaram até que ponto as intervenções que diminuem o índice de massa corporal (IMC) têm efeito sobre o número de casos de incidência evitados para cada doença. Sem intervenção, as

tendências da obesidade continuariam a aumentar em grande parte da América Latina (WEBBER et al., 2012).

O Brasil participa, mais recentemente, apenas com dois estudos. Em 2021, destaca-se o artigo de Pinto Filho et al. (2021) intitulado *Prognostic value of electrocardiographic abnormalities in adults from the Brazilian longitudinal study of adults' health*. O autor afirma que as doenças cardiovasculares (DCVs) são doenças não transmissíveis altamente evitáveis. O estudo comprova que o Eletrocardiograma é uma ferramenta potencial para a estratificação de risco em relação às DCV, ou seja, avaliou-se o papel do ECG na predição de todas as causas e mortalidade cardiovascular.

Em 2022, o estudo de Costa e Gouveia (2022), intitulado *The Body Adiposity Index is not applicable to the Brazilian adult population*, reforça que a obesidade é uma doença grave que onera os sistemas de saúde pública em todo o mundo, sendo fator de risco para o desenvolvimento de diversas doenças crônicas não transmissíveis, que estão relacionadas à quantidade e distribuição da gordura corporal. Afirma que a avaliação da composição corporal por meio de técnicas simples e de baixo custo pode auxiliar na detecção precoce do excesso de gordura, permitindo a prevenção e o tratamento tanto da obesidade quanto de doenças associadas. Assim, identificar e propor índices antropométricos válidos para esse fim pode ser um grande aliado dos programas de saúde.

Desta forma, o referido estudo verifica a validade do Índice de Adiposidade Corporal (BAI) em relação à Absorciometria de Raios-X de Dupla Energia (DXA) para estimar o percentual de gordura corporal em adultos brasileiros, bem como propor um novo modelo matemático para estimar o percentual isento de gordura massa dessa população. A conclusão é que na amostra estudada, o BAI apresentou baixa validade para estimar a gordura corporal, enquanto o novo modelo proposto mostrou-se uma boa opção para avaliar a composição corporal de adultos brasileiros (COSTA; GOUVEIA, 2022).

Como pode ser observado pela linha de tendência na Figura 2, foi apenas nos últimos três anos (2020-2022) que ocorreu a maior expansão de publicações nesta área de pesquisa, com destaque para os anos de 2020, 2021, que apresentaram 09 e 13 artigos publicados, respectivamente. Esse período de um terço de década é

responsável por cerca de 76% do total de artigos que compõem o corpus da pesquisa, reforçando, assim, a contemporaneidade do tema.

A Figura 3 apresenta a distribuição anual do corpus textual. Observa-se que entre os anos de 2022 e 2012 o percentual de publicações oscilou, totalizando 3%, equivalente a um artigo publicado na área por ano, fazendo parte do corpus desse estudo. Logo, percebe-se um elevado crescimento do ano de 2020 (09 artigos) em relação a 2021 (13 artigos), equivalendo-se a um acréscimo de 23% para 33%, acompanhados de um decréscimo no ano seguinte 2022 (08 artigos) equivalente a 20% da distribuição anual.

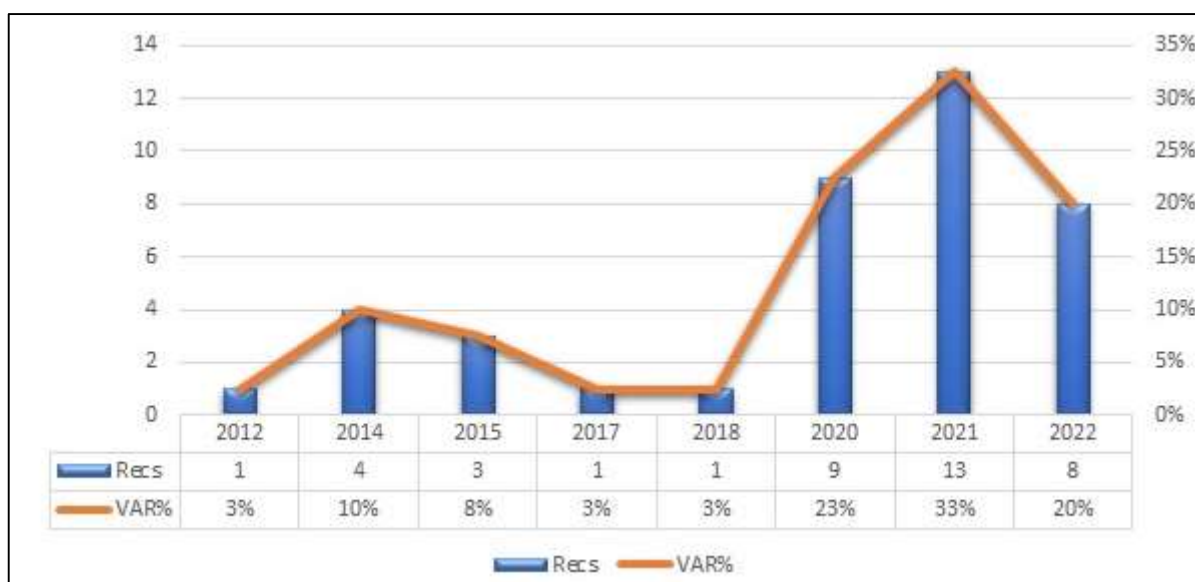


Figura 3 – Distribuição Anual do Corpus Textual em Percentual

Fonte: elaborado pelo autor (2023).

Entretanto, a Figura 3 aponta novamente uma queda no ano de 2022 (08), em que o percentual ficou em 20%, o que leva a concluir que nesse ano a temática passou a ser pouco abordada pelos pesquisadores. Tal demonstração de decréscimo nas publicações no ano de 2022 relaciona-se aos eventos decorrentes da pandemia, que dificultaram a progressão de novas pesquisas em campo em razão dos *lockdowns* e das atividades em *home-office*. A queda também se deu em virtude de cortes orçamentários por parte do governo brasileiro na liberação dos recursos públicos que se direcionavam às pesquisas (AGÊNCIA BRASIL – EBC, 2023).

Cabe ressaltar que os anos de produção científica mais elevados nessa área de modelos preditivos foram os anos de 2020 (09 artigos), 2021 (13 artigos),

totalizando respectivamente 56% do corpus textual da pesquisa, sendo 22 artigos publicados.

Devido ao fato de que essa pesquisa foi elaborada no mês de outubro de 2022, neste referido ano, o gráfico ainda se encontra no total de 20%. Todavia, a expectativa para esse ano é de que esse percentual ainda possa sofrer alterações.

Quanto à distribuição geográfica de autores, foi realizada uma análise referente aos artigos publicados (Figura 4). No que tange ao vínculo institucional, os cerca de 239 autores e coautores estão distribuídos entre 35 países. Ao observar a Figura 4, é possível perceber a liderança da República Popular da China, possuindo 16 autores e coautores que declararam possuir vínculo com instituições chinesas. Este montante corresponde a 4,18% do total de autores e coautores, que são responsáveis pela produção de 10 artigos, equivalente a 25% do corpo textual, evidenciando forte concentração desse país em pesquisas voltadas para esta área.



Figura 4 – Distribuição dos Autores do Corpus Textual por País

Fonte: Elaborado pelos autores a partir da ferramenta MapChart (2022).

AN – Article's Number (número de artigos).

Entre os países que mais publicam, estão os Estados Unidos em segunda colocação, em um total de 07 autores e coautores. A Austrália e o Reino Unido compartilham a terceira colocação, com 05 autores vinculados por país, sendo cada

país responsável pela produção de 5 artigos, equivalente a 12,5% do corpo textual. Em quarto lugar, aparecem a Índia e a África do Sul com 03 autores e coautores vinculados por país.

O Brasil surge em quinto lugar, juntamente com o Canadá, Coreia do Sul, Indonésia, Arábia Saudita, Tailândia e Vietnã, todos com 02 autores e coautores por país. Além disso, outros 22 (vinte e dois) países ao redor do globo são responsáveis pela publicação de ao menos 1 (um) artigo do corpus textual, entres eles: Áustria, Alemanha, Barein, Egito, Etiópia, Finlândia, França, Hong Kong, Irã, Irlanda, Itália, Japão, Ilhas Mauricio, Noruega, Omã, Qatar, Eslovênia, Espanha, Suécia, Suíça, Taiwan e Tanzânia.

Outrossim, torna-se importante observar que o somatório das quantidades de artigos distribuídos por país é de 69 e dos artigos que compõem o corpus textual é de 40. Esta diferença de 29 artigos para mais é explicada pela colaboração entre os países, sendo que os artigos são contados duas ou três vezes, dependendo do número de países que colaboram em cada artigo. A relação de colaboração entre autores de diferentes países é explicada na Figura 5, que evidencia a existência ou não de colaboração entre os países por meio da formação de clusters calculados a partir das relações de coautoria.

Também foi analisada a composição de autoria dos estudos que compõem o corpus com o intuito de compreender, em termos temporais, a intensidade com que os autores têm trabalhado em parcerias de pesquisas ou de colaborações com outros coautores. Na Figura 5, tem-se o perfil das coautorias para os artigos presentes no corpus textual.

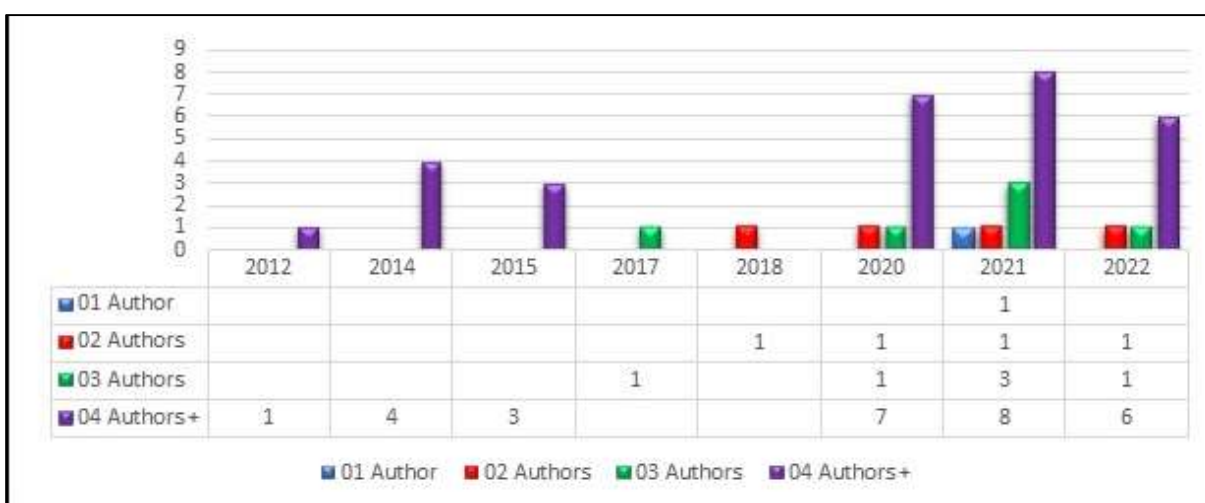


Figura 5 – Decomposição de autoria do corpus textual

Fonte: elaborado pelo pesquisador (2023).

No ano de 2014, destaca-se ainda a presença de uma publicação de artigo com 8 (oito) autores *Title: Alarming predictions for obesity and non-communicable diseases in the Middle East*. O estudo teve como objetivo modelar tendências de obesidade e futuras doenças relacionadas à obesidade para nove países no Oriente Médio, tal fato pode justificar o volumoso quantitativo de autores na obra.

A propósito, no que diz respeito à composição de autoria, observa-se, na Figura 5, a presença de intensidade de parcerias em termos de pesquisa. E, à medida que os anos vão avançando, essas parcerias se intensificam, como demonstra o período de 2020 a 2022. Vale enfatizar que a elevação na parceria entre autores e coautores denota acentuada cooperação de pesquisa e aponta a inserção dessa temática nos campos de interesse crescente por parte da comunidade acadêmica.

Em 2018, surge a primeira parceria em termos de pesquisa com 2 autores. Somente em 2021 é publicado estudo de autoria única.

Como complemento da análise anterior, estimou-se ainda a elite de pesquisa do Corpus textual com base no trabalho desenvolvido por Price (1976). Para o autor, o número de autores mais produtivos é equivalente a raiz quadrada da quantidade de autores e coautores responsáveis pelos artigos que compõem o corpus da pesquisa. Assim, ao extrair a raiz quadrada dos 239 autores e coautores, o valor obtido foi de 15,34, ou seja, equivalente a 15 autores, caracterizados como a elite desta pesquisa, tal como pode ser visualizado na Figura 6.

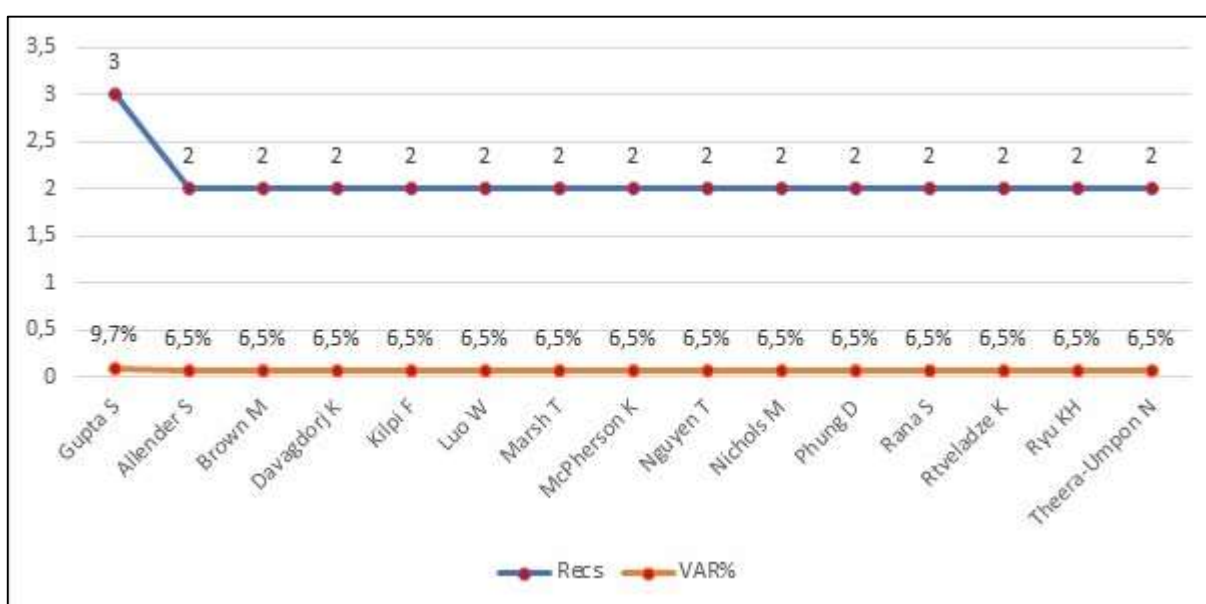


Figura 6 – Representação da Elite de Pesquisa do Corpus Textual

Fonte: elaborado pelo pesquisador (2023).

Por fim, foi realizada uma análise bibliométrica das cocitações dos autores, baseada nas ligações das citações. Segundo Small (1973), esta análise permite identificar a frequência com que dois autores são citados simultaneamente por algum elemento da literatura mais recente.

Neste compasso, Wallin (2005) destaca que quanto maior for o quantitativo de pesquisadores citando os mesmos dois artigos, maior será a probabilidade de que essa dupla citação não seja uma ocorrência do acaso. Tal fato evidencia um tipo de assunto relacionado a publicações citadas dentro das áreas de investigação pretendidas.

Observa-se na Figura 7 a rede de cocitação de autores, que foi extraída dos 239 autores, dizendo respeito as referências citadas nos 40 artigos do corpus analisados. Esta figura foi elaborada com o auxílio do *software* VOSviewer, em que o tipo de análise escolhido foi “*co-citation*”, a unidade de análise escolhida foi “*cited authors*” e o método de contagem utilizado foi “*full counting*”, que atribui o mesmo peso para cada link entre os autores co-citados. Além disso, definiu-se um número mínimo de 04 citações para cada autor, conforme mapeamento do próprio *software*. Esta foi a configuração que possibilitou a melhor visualização da rede, resultando em 30 nós, que representam um total de 2.435 laços relacionais entre os autores citados de maneira simultânea sendo que foram selecionados 52 autores.

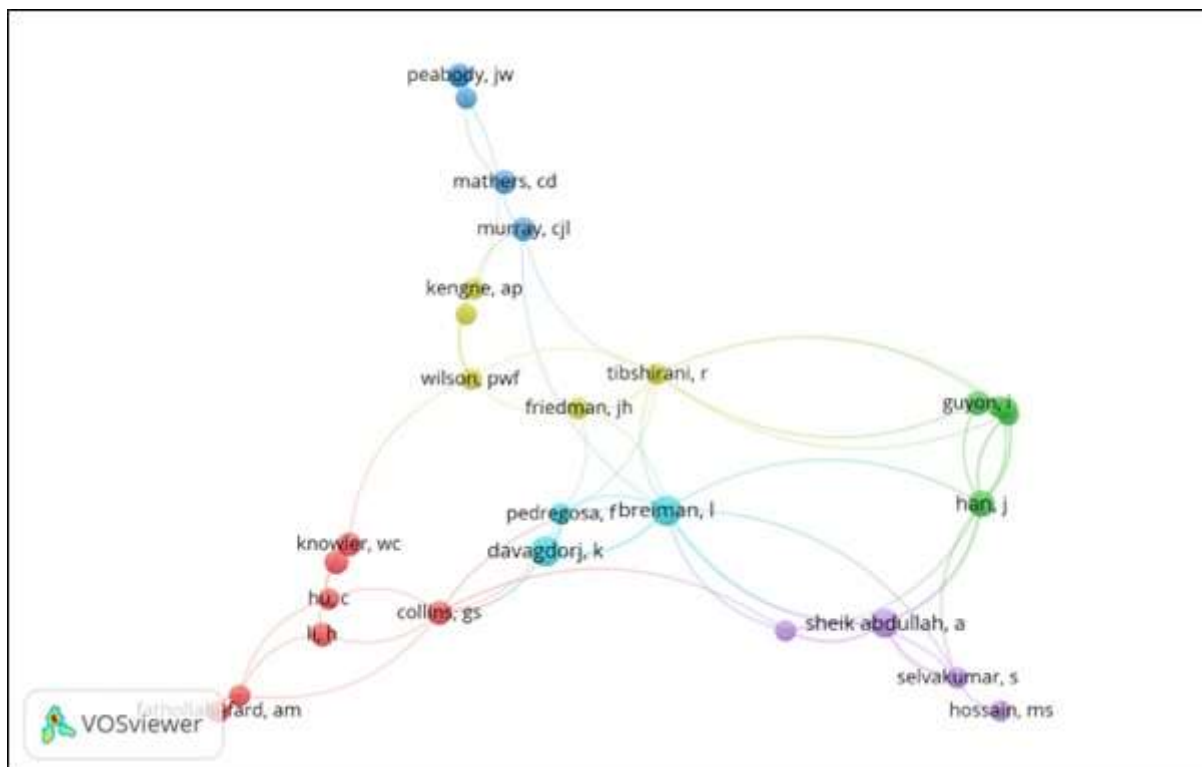


Figura 7 – Rede de cocitação de autores com 'nós' calculados segundo a quantidade de citações
 Fonte: elaborado pelo pesquisador (2023).

Na Figura 7, é possível visualizar a formulação de 6 clusters nas seguintes cores: vermelho amaranço, azul capri, lilás, verde, abacate e azul grisáceo, em que as linhas estabelecem a interação de cocitação entre os autores e a espessura dos “nós” é proporcional ao quantitativo de citação de cada autor. Pode-se perceber dois clusters proeminentes, contendo autores que são citados em conjunto mais vezes. O Primeiro, de cor azul capri, traz como principais autores Breiman (8 citações) e Davagdorj (8 citações), ambos pertencentes ao cluster 6. No Cluster 5, cor lilás, destaca-se o segundo autor mais citado Sheik Abudallah (7 citações). E, no cluster 2, na cor verde, aparece como terceiro autor mais citado Han com (6 citações).

Para a Análise Acoplamento Bibliográfico, com auxílio do software VOSviewer, os parâmetros definidos foram “*bibliographic coupling*”, para o tipo de análise; “*documents*” para a unidade de análise e “*full counting*” para o método de contagem, atribuindo o mesmo peso para cada link entre os documentos. Os dados compreenderam 82 documentos, **36 meet threshold**, com mínimo 4 citações, extraídos do corpus textual (40 artigos), sendo essa, a configuração que possibilitou a melhor visualização da rede, e resultou em 4 Clusters com 09 itens acoplados, descritos na Figura 8.

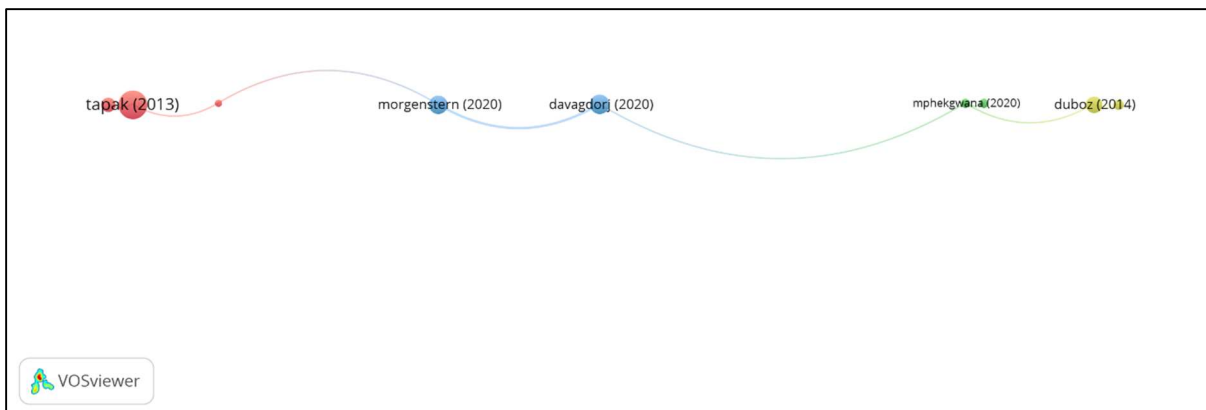


Figura 8 – Rede de documentos acoplados com 'nós' calculados segundo a quantidade de citações

Fonte: elaborado pelo pesquisador (2023).

Na análise do acoplamento foi encontrado um total de 4 clusters dispostos no corpus do estudo. O Cluster 1 é representado pelo maior número de citações (72), seguido do cluster 3 (48), cluster 4 (26) e, finalmente, cluster 2 (12). Em relação às conexões, dois artigos apresentaram a maior conectividade, os estudos de Davagdorj et al. (2020) e Morgenstern et al. (2020), ambos com 4 laços cada um, respondendo por 40% das conexões gerais e um total de 20 ligações.

Em termos de citações, o destaque foi para o trabalho criado por Tapak et al. (2013), nó 3, com um total de 54 citações. Em seguida, com o segundo maior de citações observa-se o trabalho de Davagdorj et al. (2020) com um total de 25 citações. Por fim, Morgenstern et al. (2020) com um total de 23 citações. O trabalho de Tapak et al. (2013) representa 75% das citações recebidas no Cluster 1 (C1) e 34,17% das citações em relação a todos os clusters.

No Cluster 2 o trabalho de Mphekgwana et al. (2020) com 6 citações clusters equivale a 15% das conexões gerais. O Cluster 3 é representado pelo segundo maior número de citações pelo estudo de Davagdorj et al. (2020), que possui 25 citações, juntamente com Morgenstern et al. (2020), com 23 citações, totalizando 48 citações, equivalente a 30,37% das citações em relação a todos os clusters.

O Cluster 4 foi destacado no artigo de Duboz et al. (2014) com o total de 18 citações equivale a 69,23% do cluster, seguido de Rosini et al. (2013) com 8 citações, conforme Tabela 1.

Tabela 10 – Documentos acoplados com 'nós' calculados segundo a quantidade de citações

CLUSTER	AUTOR COCITADO	LIGAÇÃO	CITAÇÕES	PERIÓDICO	TÍTULO
1	Adavi et al.(2016)	1	14	MEDICAL JOURNAL OF THE ISLAMIC REP	Neural Artificial Networks versus Bivariate Logistic Regression Pre
1	Sheik Abdul-lah (2021)	2	4	SOFT COMPUTING	Assessment and Evaluation of CHD Risks Factors Using Weighed Ranked
1	Tapak (2013)	2	54	HEALTHCARE INFORMATICS RESEARCH	Real-Data Comparison of Data Mining Methods in Prediction of Diabetes
2	Bali (2019)	1	6	INTERNATIONAL JOURNAL OF RECENT	Diabot: A Predictive Medical Chatbot Using Ensemble Learning
2	Mphekgwana (2020)	3	6	INTERNATIONAL JOURNAL OF ENVIROMENT	Hypertension Prevalence and Determinants Among Black South Africa
3	Davagdorj (2020)	4	25	INTERNATIONAL JOURNAL OF ENVIROMENT	Xgboost-based Framework for Smoking-induced Noncommunicable Disease
3	Morgenstern (2020)	4	23	BMJ OPEN	Predicting Population Health with Machine Learning: A Scoping Review
4	Duboz (2014)	2	18	JORNAL OF HUMAN HYPERTENSION	Hypertension Prevalence, Awareness, Treatment and Control in Dakar
4	Rosini (2013)	1	8	CLINICAL OF BIOCHEMISTRY	Simultaneous Prediction of Hyperglycemia Dyslipidemia in School

Fonte: Dados da Pesquisa (2023).

A Tabela 10 elenca os principais periódicos com os maiores quantitativos de artigos publicados para a temática abordada pela RSL. A avaliação da produtividade de pesquisadores e periódicos pode ser realizada por meio de indicadores de citação que oferecem um conjunto de informações sintetizadas sobre o funcionamento da comunidade científica, considerando parâmetros levantados *ex-post*. Além disso, considera também o binômio autoria-citação, que permite a descrição da dinâmica e da densidade das contribuições acerca uma temática ou especialidade. Adiante, a Tabela 2 mostra alguns indicadores bibliométricos relacionados à produtividade e à reputação dos 27 periódicos, dos quais os artigos selecionados foram publicados.

No que diz respeito à produtividade, a Lei de Bradford permite estimar o grau de relevância de periódicos que operam em áreas do conhecimento específicas. Os periódicos que contemplam uma maior publicação de artigos acerca de um assunto tendem a estabelecer um núcleo supostamente de qualidade superior e de maior relevância nesta determinada área do conhecimento (Tabela 11).

Tabela 11 – Indicadores bibliométricos dos periódicos do corpus textual

JOURNALS	SJR QUARTILE	NUMBER OF ARTICLES	SJR INDEX 2021	H INDEX
PLOS ONE	Q1	6	0,85	367
INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH	Q1	4	0,81	138
INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS	Q1	3	1,14	114
BMC PUBLIC HEALTH	Q1	2	1,16	159
BMJ OPEN	Q1	2	0,98	121
IEEE ACCESS	Q1	2	0,93	158
ACTA DIABETOLOGICA	Q1	1	1,05	73
MICROBIOME	Q1	1	4,39	95
JOURNAL OF CLINICAL EPIDEMIOLOGY	Q1	1	3,11	223
HEART	Q1	1	2,15	189
IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS	Q1	1	1,8	137
JOURNAL OF CLINICAL ENDOCRINOLOGY & METABOLISM	Q1	1	1,75	363
JOURNAL OF MEDICAL INTERNET RESEARCH	Q1	1	1,74	158
COMPUTERS ENVIRONMENT AND URBAN SYSTEMS	Q1	1	1,7	97
JOURNAL OF EPIDEMIOLOGY AND COMMUNITY HEALTH	Q1	1	1,67	178
DIABETES AND METABOLIC SYNDROME: CLINICAL RESEARCH AND REVIEWS	Q1	1	1,59	56
COMPUTERS IN BIOLOGY AND MEDICINE	Q1	1	1,31	102
FRONTIERS IN PUBLIC HEALTH	Q1	1	1,03	64
FRONTIERS IN NUTRITION	Q1	1	1,02	44
BMC HEALTH SERVICES RESEARCH	Q1	1	1	122
PUBLIC HEALTH	Q1	1	0,97	80
JOURNAL OF DIABETES INVESTIGATION	Q1	1	0,93	55
PUBLIC HEALTH NUTRITION	Q1	1	0,9	146
VASCULAR HEALTH AND RISK MANAGEMENT	Q1	1	0,9	73
BMC MEDICAL INFORMATICS AND DECISION MAKING	Q1	1	0,83	79
MATHEMATICS	Q1	1	0,74	36
MULTIMEDIA TOOLS AND APPLICATIONS	Q1	1	0,72	80

Fonte: Dados da Pesquisa (2023).

Em relação a lei de Bradford, ou lei da dispersão, esta incide sobre um conjunto de publicações (periódicos) para determinar o núcleo dos mais produtivos em uma determinada temática. Sua representação é realizada por zonas e sua distribuição é apresentada em três zonas, sendo todas representadas por 33% das publicações (Bradford, 1934 apud ALABI, 1979).

Dessa forma, através da medição da produtividade das revistas, se estabelece o núcleo, além das áreas de dispersão sobre determinado assunto em um mesmo conjunto de periódicos. Assim, de acordo com a Tabela 2, os 27 periódicos foram posicionados em ordem decrescente de produtividade e distribuídos em três zonas.

Destaca-se que todos os periódicos fazem partes do primeiro quartil (Q1), segundo a classificação da plataforma *Scimago Journal & Country Rank* (SJR). Portanto, esta primeira zona composta por 9 periódicos (33% das publicações), proporcionalmente, é a zona mais produtiva, respondendo 22 artigos ou 55% do corpus da pesquisa.

A segunda zona, formada por 09 periódicos, responde por 22,50% do total dos periódicos. Logo, esta segunda zona composta por 9 artigos, proporcionalmente, é a zona menos produtiva, respondendo por 22,50 % do corpus da pesquisa. Por sua vez, a terceira zona, formada por 09 periódicos, responde por 22,50% do total dos

periódicos. Logo, essa terceira zona composta por 9 artigos, proporcionalmente, também é a zona menos produtiva, respondendo por 22,50% do corpus da pesquisa. Por último, o periódico *PLOS ONE Journal* se destaca pela publicação de 6 artigos.

Em relação à reputação dos periódicos, possível visualizar que os 40 artigos do corpus textual foram publicados em periódicos situados no primeiro quartil de citações, o que mostra a visibilidade e prestígio destes. Esta mesma figura ainda traz os Índice H e o *Journal Citation Report (JCR) impact factor*¹³. Destaca-se, que apesar da inserção dos periódicos na área de Administração da Informação (ADI) com ênfase em Tecnologia e Saúde definida para este estudo, os valores desses índices englobam todas as áreas do conhecimento em que os periódicos operam. Inclusive, que os referidos índices foram analisados de forma isolada, sem comparação com as respectivas zonas.

Ao analisar o Índice H, o *PLOS ONE Journal*, que se encontra na primeira zona, acaba se sobressaindo em relação aos demais, pois apresenta um índice H¹⁴ INDEX 367, o que permite afirmar que esse periódico teve, pelo menos, 367 de seus artigos citados 367 vezes, refletindo a sua elevada reputação.

A Tabela 11 apresenta os valores do *JCR impact factor*, que divulga a frequência média de citações recebidas pelos artigos publicados nos periódicos, no qual utilizou-se o ano de 2021 (último estudo realizado) como referência para o cálculo. Na presente pesquisa, ao considerar este índice, o destaque foi do *MICROBIOME Journal*, que apresentou um JCR igual a 4,39.

Ao observar a Tabela 12, é possível perceber a distribuição dos periódicos por país com as suas respectivas áreas de inserções.

¹³ SJR impact factor. Definição. Citações a um periódico no ano JCR para itens publicados nos dois anos anteriores, dividido pelo número total de itens citáveis (artigos e resenhas) publicados no periódico nos dois anos anteriores.

¹⁴ Índice H O índice h é o número de artigos (h) de um país que receberam pelo menos h citações. Ele quantifica tanto a produtividade científica do país quanto o impacto científico e também é aplicável a cientistas, periódicos, etc.

Tabela 12 – Distribuição dos *Journals* por País

%	AMOUNT	COUNTRY	JOURNAL	JOURNALS INSERTION AREA
37,04%	10	United Kingdom	BMC HEALTH SERVICES RESEARCH	Health Policy
		United Kingdom	BMC MEDICAL INFORMATICS AND DECISION MAKING	Health Policy
		United Kingdom	BMC PUBLIC HEALTH	Public Health, Environmental and Occupational
		United Kingdom	BMJ OPEN	Medicine - Miscellaneous
		United Kingdom	COMPUTERS ENVIRONMENT AND URBAN SYSTEMS	Ecological Modeling
		United Kingdom	COMPUTERS IN BIOLOGY AND MEDICINE	Computer Science Applications
		United Kingdom	HEART	Cardiology and Cardiovascular Medicine
		United Kingdom	JOURNAL OF EPIDEMIOLOGY AND COMMUNITY HEALTH	Epidemiology
		United Kingdom	MATHEMATICS	Mathematics - Miscellaneous
		United Kingdom	PUBLIC HEALTH NUTRITION	Medicine - Miscellaneous
18,52%	5	United States	IEEE ACCESS	Computer Science
		United States	IEEE JOURNAL OF BIOMEDICAL AND HEALTH INFORMATICS	Biotechnology
		United States	JOURNAL OF CLINICAL ENDOCRINOLOGY & METABOLISM	Biochemistry
		United States	JOURNAL OF CLINICAL EPIDEMIOLOGY	Epidemiology
		United States	PLOS ONE	Multidisciplinary
14,81%	4	Switzerland	FRONTIERS IN NUTRITION	Endocrinology, Diabetes and Metabolism
		Switzerland	FRONTIERS IN PUBLIC HEALTH	Public Health, Environmental and Occupational
		Switzerland	INTERNATIONAL JOURNAL OF ENVIRONMENTAL RESEARCH AND PUBLIC HEALTH	Health, Toxicology and Mutagenesis
		Switzerland	MICROBIOME	Microbiology
11,11%	3	Netherlands	DIABETES AND METABOLIC SYNDROME: CLINICAL RESEARCH AND REVIEWS	Endocrinology, Diabetes and Metabolism
		Netherlands	MULTIMEDIA TOOLS AND APPLICATIONS	Media Technology
		Netherlands	PUBLIC HEALTH	Medicine - Miscellaneous
3,70%	1	Ireland	INTERNATIONAL JOURNAL OF MEDICAL INFORMATICS	Health Informatics
3,70%	1	Italy	ACTA DIABETOLOGICA	Endocrinology
3,70%	1	Canada	JOURNAL OF MEDICAL INTERNET RESEARCH	Health Informatics
3,70%	1	Australia	JOURNAL OF DIABETES INVESTIGATION	Medicine - Miscellaneous
3,70%	1	New Zeland	VASCULAR HEALTH AND RISK MANAGEMENT	Medicine - Miscellaneous

Fonte: elaborado pelo pesquisador (2023).

Assim, no que diz respeito a este tipo distribuição, verificou-se que os 27 periódicos estão distribuídos em 09 países, sendo em primeiro lugar, o Reino Unido, com 10 *Journals* (37,04% do total dos periódicos). Na segunda posição, destacam-se os Estados Unidos, com 05 *Journals* (52% do total dos periódicos), e segue-se a Suíça, com 04 *Journals* (14,81% do total dos periódicos). Na quarta posição, está a Holanda com 03 *Journals* (11,11% do total dos periódicos), e seguem-se na sequência posicional, a Austrália, Canadá, Irlanda, Itália e Nova Zelândia, cada um deles com 01 *Journal* por país (3,70% do total dos periódicos).

A Nuvem de Palavras (Figura 9), refere-se à organização e ao agrupamento gráfico das palavras em função da sua frequência ou ocorrência, permitindo a identificação de seu conteúdo lexical, assim percebendo quais as expressões mais evocadas na pesquisa.

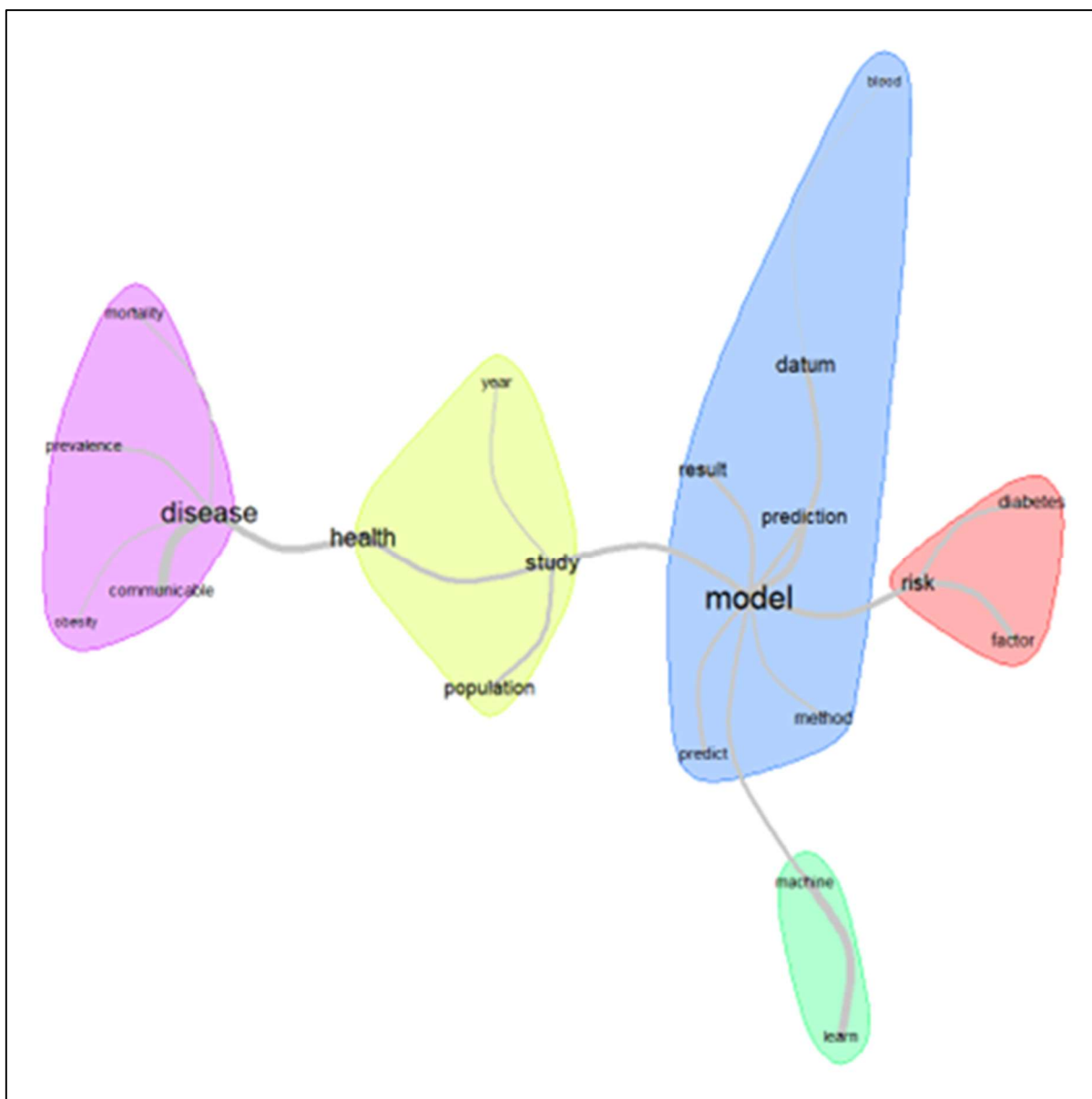


Figura 10 – Análise de Similitude. Extraído do software IRAMUTEQ

Fonte: elaborado pelo pesquisador (2023).

Conforme observado na árvore de co-ocorrência, os resultados indicaram que entre os pares de associação se observa uma forte relação entre: **model – machine – risk – study**. O elemento **model** aparece como elemento de maior centralidade e ainda apresenta relações importantes com inúmeros elementos que o complementam e lhe dão sustentação, interligando na formação de várias teias de desdobramentos interessantes e interconectados com a questão da predição e dos métodos.

Os **principais pares** de associação surgem entre as seguintes associações de palavras: (*model-risk*), (*model-machine*), (*model-study*) e (*health-disease*). A primeira grande conexão visualizada ocorre entre o binômio (*model – risk*), que se conecta com

factor - diabetes, o que remete ao interesse em mapear os fatores que levam ao surgimento de DCNTs, neste caso o diabetes. Segundo o Atlas do Diabetes (2021, p. 37), entre os 10 principais países ou territórios em número de adultos (20 a 79 anos) com diabetes, em 2021 o Brasil figura em 6º lugar com 15,7 milhões de pessoas com a doença, com previsão de chegar em 2045 a 23,2 milhões de pessoas com diabetes, ou seja, 47,47% a mais em relação a 2021.

A segunda grande conexão visualizada ocorre entre o binômio (*model – machine*), o que demonstra o interesse por estudos em modelos de aprendizado de máquina, ou seja, modelos preditivos que visem auxiliar na previsão de ocorrências e comportamentos, dando base para a tomada de decisões antecipadas. A terceira grande conexão visualizada ocorre entre o binômio (*model – study*), que se conecta com *population* e *health*, demonstrando a inclinação por estudos voltados a investigação da saúde populacional.

A quarta grande conexão visualizada ocorre entre o binômio (*health – disease*), que se conecta com (*communicable – prevalence – obesity – mortality*) demonstrado relevância por estudos que correlacionem essas variáveis. Em resumo, o que pode-se observar é que a partir da centralidade da palavra *model*, além de *risk*, existem um conjunto de aspectos importantes observados nas palavras (*study, health, population, disease, communicable, prevalence, obesity e mortality*), que ao se envolverem nesse processo nos leva ao conceito ampliado de saúde, onde a doença não é apenas a ausência de doença, mas um completo bem-estar físico, mental e social o que envolve inúmeras práticas de cuidado de si, com o corpo e diante do eminente envelhecimento.

A Figura 11 evidencia que a expressão *Prediction* (primeiro nó à esquerda) está ligada a *Risk-Factors*, demonstrando que a maior parte destes estudos foram conduzidos neste tipo de aglomerado e, que eles estão ligados a estudos populacionais que circundam a prevalência do diabetes mellitus a correlação com o índice de massa corpórea.

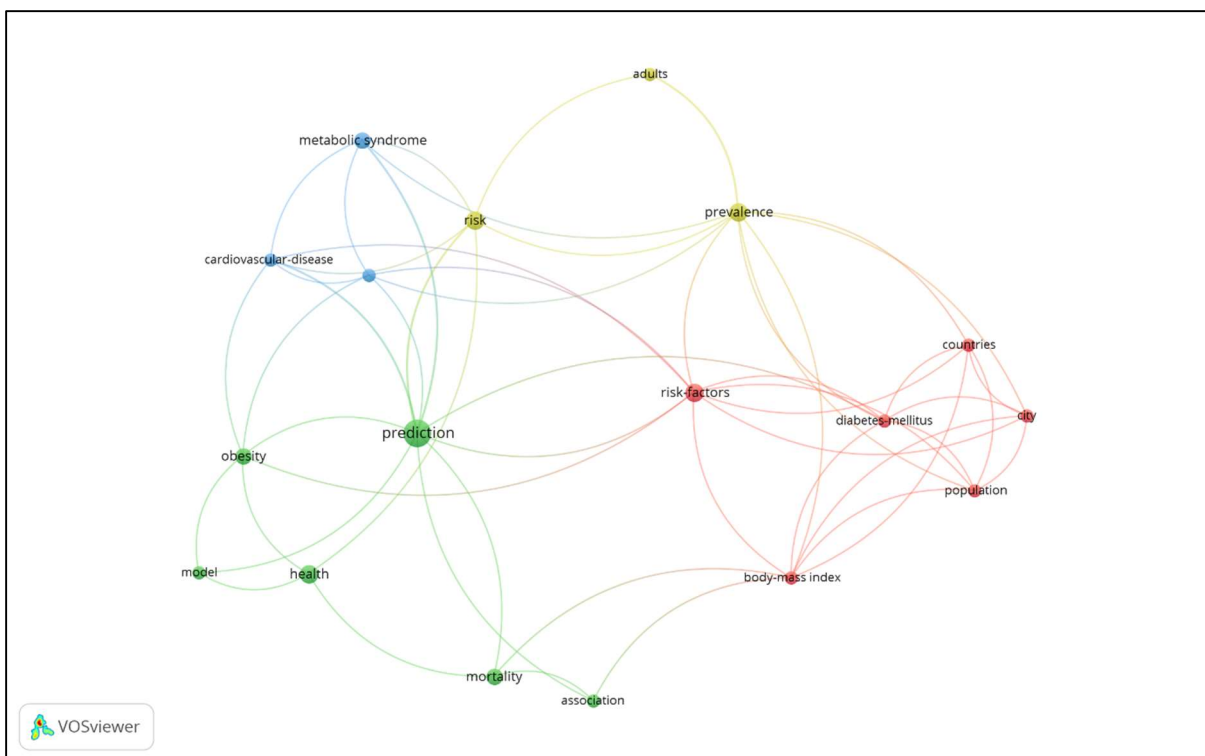


Figura 11– Rede de agrupamento de palavras

Fonte: elaborado pelo pesquisador (2023).

Para esta análise, com auxílio do software VOSviewer, os parâmetros definidos foram “*co-occurrence*”, para o tipo de análise; “*Keywords Plus*” para a unidade de análise e “*full counting*” para o método de contagem, atribuindo o mesmo peso para cada *link* entre os documentos. Os dados compreenderam 162 *Keywords*, 23 *meet threshold* com mínimo 2 ocorrências (*number of Keywords to be selected = 5*), extraídos do *corpus* textual (40 artigos), sendo a configuração que possibilitou a melhor visualização da rede, resultando em 12 (palavras-chaves selecionadas) itens acoplados, descritos na Figura 11.

Por fim, ao analisar os artigos incluídos na RSL, levantaram-se as principais sugestões dos autores/pesquisadores para pesquisas futuras, a partir do estudo de seus temas. Com essas expectativas, elaborou-se a proposição de uma agenda de pesquisa com propostas de temas a serem explorados dentro da área de Sistemas de Informação, conforme o Quadro 5 que segue:

Nome do Artigo e Weblink do Artigo	Autor	Proposta
<p>Modelling and prediction of global non-communicable diseases</p> <p>https://link.springer.com/article/10.1186/s12889-020-08890-4</p>	<p>Wang Y, Wang JF (2020)</p>	<p>Explorar as DCNTs causadas por indicadores físicos anormais (pressão arterial, lipídios no sangue, IMC, etc.) resultantes de hábitos alimentares e hábitos de vida adquiridos, pois são a parte principal da incidência da doença.</p>
<p>Xgboost-based framework for smoking-induced noncommunicable disease pr</p> <p>https://www.mdpi.com/1660-4601/17/18/6513</p>	<p>Davagdorj K, Pham VH, Theera- Umpon N, Ryu KH (2020)</p>	<p>Ampliar o estudo no sentido de melhorar a interpretabilidade do modelo de aprendizado profundo de máquina global e local dos modelos caixa-preta e do efeito causal no cenário dos modelos preditivos.</p>
<p>Predicting population health with machine learning: A scoping review</p> <p>https://bmjopen.bmj.com/content/10/10/e037860</p>	<p>Morgenstern JD, Buajitti E, O'Neill M, Piggott T, Goel V, et al. (2020)</p>	<p>Futuros projetos de aprendizado de máquina poderiam incorporar conjuntos de dados maiores e mais recursos não tradicionais. Uma maior utilização de recursos como o HealthMap, as redes sociais, os padrões de pesquisa na Internet, a detecção remota e os relatórios da OMS permitiriam mais trabalho em regiões sem fontes formais de dados e enriqueceriam a investigação noutras. Outra perspectiva largamente inexplorada é a utilização da aprendizagem automática e de dados de alta dimensão para incorporar representações mais ricas dos determinantes sociais da saúde. Além disso, avaliar o impacto dos modelos de previsão nas decisões tomadas em saúde populacional e na prática da saúde pública.</p>
<p>Establishment of noninvasive diabetes risk prediction model based on t</p> <p>https://www.sciencedirect.com/science/article/pii/S1386505621000551</p>	<p>Li J, Chen Q, Hu X, Yuan P, Cui L, et al. (2021)</p>	<p>Incorporar o modelo de aprendizagem profunda para segmentação de imagens de linguagem TDAS*, que são usadas para realizar diagnósticos das imagens da língua. Além disso, coletar mais dados para explorar um melhor algoritmo de aprendizado de máquina e uma estratégia de integração de treinamento, a fim de obter um modelo preditivo mais preciso de previsão de risco de diabetes. (* Tongue Diagnosis Analysis System (TDAS, do inglês Sistema de análise de diagnóstico de língua)</p>

<p>Associations between habitual diet, metabolic disease, and the gut microbiota using latent Dirichlet allocation</p> <p>https://microbiomejournal.biomedcentral.com/articles/10.1186/s40168-020-00969-9</p>	<p>Breuninger TA, Wawro N, Breuninger J, Reitmeier S, Clavel T, et al. (2021)</p>	<p>Investigar a relação longitudinal entre a ingestão alimentar habitual e a estrutura da microbiota intestinal no desenvolvimento de uma série de doenças não transmissíveis.</p>
<p>Exploratory study on classification of diabetes mellitus through a combined Random Forest Classifier</p> <p>https://bmcmmedinformdecismak.biomedcentral.com/articles/10.1186/s12911-021-01471-4#Sec25</p>	<p>Wang XC, Zhai MM, Ren ZP, Ren H, Li MC, et al. (2021)</p>	<p>Explorar múltiplos métodos de redução de dimensionalidade de recursos para detectar a correlação de variáveis a fim de identificar indivíduos com alto risco para diabetes permitindo estabelecer estratégias de prevenção da doença. Além disso, incluir nos estudos dados de antecedentes familiares sobre fatores de risco para hiperlipidemia.</p>
<p>Predicting length of stay and mortality among hospitalized patients with type 2 diabetes mellitus and hypertension</p> <p>https://www.sciencedirect.com/science/article/pii/S1386505621001957#s0070</p>	<p>Barsasella D, Gupta S, Malwade S, Aminin, Susanti Y, et al (2022)</p>	<p>Pesquisas futuras são necessárias para uma descrição mais detalhada do LoS* hospitalar, detalhes da alta do paciente e motivo da mortalidade.</p> <p>(*) previsão de tempo de internamento (LOS, do inglês length of stay)</p>
<p>Synthetic Patient Data Generation and Evaluation in Disease Prediction</p> <p>https://ieeexplore.ieee.org/document/9851514</p>	<p>Rodriguez-Almeida AJ, Fabelo H, Ortega S, Deniz A, Balea Fernandez FJ, et al (2022)</p>	<p>Ajustar profundamente as Redes Adversariais Gerativas Tabulares Condicionais (CTGANs) deste estudo. Desenvolver ferramentas para a interpretabilidade deste algoritmo esclarecendo o processo de geração de dados sintéticos para investigadores e médicos de IA. E, validar clinicamente este estudo no campo clínico, acelerando o desenvolvimento de algoritmos baseados em IA que poderiam auxiliar durante a prática clínica.</p> <p>Redes Adversariais Gerativas Tabulares Condicionais (GTGANs do inglês, Conditional Tabular Generative Adversarial Networks)</p>

Quadro 5 – Proposta de agenda de pesquisa sobre modelos preditivos doenças crônicas não transmissíveis

Fonte: Elaborado pelo autor.

4 CONSIDERAÇÕES FINAIS

Segundo o objetivo proposto neste estudo, identificar e sintetizar as evidências relevantes disponíveis sobre estudos de predição em DCNTs, através da revisão sistemática da literatura; buscou-se auxílio na RSL, e a partir desse ponto foram mapeados e analisados os estudos na área de “Administração da Informação (ADI) com ênfase em Tecnologia e Saúde das bases de dados SCOPUS (*Elsevier*TM) e WEB OF SCIENCE (*Clarivate Analytics*TM). Considera-se, tendo em vista as análises realizadas, que a pesquisa é assertiva, pois verifica e demonstra o panorama atual das pesquisas científicas sobre modelos preditivos para doenças crônicas.

Conforme as análises organizadas no estudo, realizaram-se com base nas três leis da bibliometria, a saber: Lei de Lotka, que avalia a produtividade dos autores; Lei de Bradford, que mensura o grau de produção dos periódicos; e, por fim, a Lei de Zipf, que mensura a frequência de ocorrência e coocorrência das palavras expostas pelos textos.

A Lei de Lotka presume que existe uma pequena elite de pesquisadores mais prolíficos, enquanto a grande maioria pouco publica. Assim, verificou-se que a elite do corpus textual contou com 15 autores, em que um deles possui 03 artigos, sendo que o restante 02 artigos. Dos 40 artigos do corpus textual, 31 deles (77,50%) foram produzidos por essa elite de pesquisa.

Em relação a Lei de Bradford, que se refere à distribuição da reputação dos periódicos, têm-se que a primeira zona contempla nove periódicos que seriam os mais produtivos, enquanto as demais zonas (segunda e terceira) mostraram-se iguais em número de artigos e produtividade, resultando em nove periódicos por zona, num total de vinte e sete periódicos que foram contemplados.

No que se refere a Lei de Zipf, utilizou-se a Nuvem de Palavras e o Gráfico de Similitude, resultando na conclusão de que nos artigos que fazem parte do corpus de análise a preocupação com tecnologias que auxiliem na previsão de ocorrências e comportamentos, dando base para a tomada de decisões antecipadas. No que diz respeito à expressão *Prediction*, verificou-se que está ligada a expressão *Risk-Factors*, demonstrando que a maior parte destes estudos foram conduzidos neste tipo de aglomerado.

Sobre a distribuição por países, foi possível verificar que os 27 periódicos estavam distribuídos em 09 países, destacando-se em ordem de publicações: 1º lugar, Reino Unido: 10 *Journals* (37,04% do total dos periódicos), 2º lugar, Estados Unidos: 05 *Journals* (18,52% do total dos periódicos), 3º lugar, Suíça: 04 *Journals* (14,81% do total dos periódicos), 4º lugar, Holanda: 03 *Journals* (11,11% do total dos periódicos), 5º lugar, Austrália, Canadá, Irlanda, Itália e Nova Zelândia: cada um deles com 01 *Journal* por país (3,70% do total dos periódicos).

Foi possível verificar a formulação de seis clusters. Cabe ressaltar, que as publicações verificadas no ano 2022 não representam a totalidade do período, uma vez que a pesquisa foi realizada no mês de outubro do referido ano, e muitas bases não indexaram o restante das publicações.

Sobre as limitações da pesquisa que podem ter afetado essa RSL, estão o viés de publicação (e outros similares como viés de linguagem), riscos de viés nos estudos primários (limitação metodológica dos estudos primários), além de dificuldades em combinar estudos que podem ter diferenças nas populações, intervenções, comparadores e definição dos desfechos (heterogeneidade clínica). Assim, é necessário explorar criticamente a evidência apresentada pela metanálise.

As informações que foram apresentadas podem servir de base científica para o desenvolvimento de pesquisas futuras. Assim, como sugestões para estudos futuros em modelagem preditiva, apontam-se possibilidades de desenvolvimento, como por exemplo, a proposição de uma agenda de pesquisa com propostas de temas a serem explorados dentro da área de Sistemas de Informação, e ainda, desenvolver outros estudos no campo da saúde

REFERÊNCIAS CAPÍTULO 3

AGÊNCIA BRASIL – EBC – Empresa Brasileira de Comunicação. **Brasil teve queda de 7,4% na produção científica entre 2022 e 2021.** (2023). Disponível em: <https://agenciabrasil.ebc.com.br/geral/noticia/2023-07/brasil-teve-queda-de-74-na-producao-cientifica-entre-2022-e-2021>. Acesso em 28.jul.2023.

ALABI, G. Bradford's law and its application. **International Library Review**, v. 11, n. 1, p. 151-158, 1979.

ALMEIDA, C. P. B.; GOULART, B. N. G. Como minimizar vieses em revisões sistemáticas de estudos observacionais. **Revista CEFAC**, São Paulo, v. 19, n. 4, p. 551-555, 2017.

ASGARI, S.; KHALILI, D.; ZAYERI, F.; AZIZI, F.; HADAEGH, F. Dynamic prediction models improved the risk classification of type 2D. **Journal of clinical epidemiology**. v.140, p.33-43, 2021.

ATLAS DO DIABETES (2021). 10th Edition. **International Diabetes Federation**. Disponível em: https://diabetesatlas.org/idfawp/resource-files/2021/07/IDF_Atlas_10th_Edition_2021.pdf. Acesso em 15.dez.2022.

BARSASELLA, D., GUPTA, S., MALWADE, S., AMININ, SUSANTI, Y., TIRMADI, B., SYED-ABDUL, S. (2021). Predicting length of stay and mortality among hospitalized patients with type 2 diabetes mellitus and hypertension. **International Journal of Medical Informatics**, v. 154, n.104569. 2021.

BAUER, M. W.; AARTS, B. Corpus construction: a principle for qualitative data collection. **Qualitative Researching With Text, Image, and Sound: A Practical Handbook**, p. 19-37, 2000.

BIOLCHINI, J. C. A.; MIAN, P. G.; NATALI, A. C. C.; CONTE, T. U.; TRAVASSOS, G. H. Scientific Research Ontology To Support Systematic Review In Software Engineering. **Advanced engineering informatics**, v. 21, n. 2, p. 133-151, 2007.

BORNMANN, L.; MARX, W. Histicite analysis of papers constituting the h index research front. **Journal of Informetrics**, v. 6, n. 2, p. 285-288, 2012.

BREIMAN, L. Random Forests. **Machine Learning**, v. 45, n. 5, p. 32 (2001).

BREUNINGER, T. A.; et al. Associations between habitual diet, metabolic disease, and the gut microbiota using latent Dirichlet allocation. **Microbiome**. v. 16, n. 9 (1), p. 61, mar, 2021.

CAMARGO, B. G.; JUSTO, A M. IRAMUTEQ: Um software gratuito para análise de dados textuais. **Temas Psicol.** [Online]. v. 21, n. 2, p. 513-518. 2013.

CARVALHO, D. V.; PEREIRA, M. E.; CARDOSO, J. S. Interpretabilidade do aprendizado de máquina: uma pesquisa sobre métodos e métricas. **Electronics**, v. 8, n. 8, p. 832, jul. 2019.

CHUNG, K. C.; BURNS, P. B.; KIM, H. M. Clinical Perspective: A Practical Guide to Meta Analysis. **The Journal of Hand Surgery**. v. 31A, n. 10, p.1671, 2006.

CLARIVATE ANALYTICS. Web of science databases, 2018. Cother Hajat, Emma Stein. The global burden of multiple chronic conditions: a narrative review. **Preventive Medicine Reports**, v. 12, p. 284-293, 2018.

COSTA, O.; GOUVEIA, L. B. Uma proposta para um sistema inteligente de previsão do risco de doenças crônicas. “no prelo”. **XIX Congresso Brasileiro de Informática em Saúde – CBIS**. Campinas/SP, Brasil. 29/11 A 02/12 De 2022 - Campinas/SP – Brasil, 2022.

DAVAGDORJ, K. ; LEE, J.S. ; PHAM, V.H. ; RYU, K.H. (2020). A comparative analysis of machine learning methods for class imbalance in a smoking cessation intervention. **Applied Sciences**, v.10, n.9, p.3307, 2020.

DAVAGDORJ, K.; BAE, V. W.; PHAM, N. H.; THEERA-UMPON, N.; RYU, K. (2021). Explainable artificial intelligence based framework for non-communicable diseases prediction, **IEEE Access**, v. 9, p. 123672-123688, 2021.

DELPINO, F.M.; COSTA, A.K.; FARIAS, S.R.; CHIAVEGATTO FILHO, A.D.P.; ARCÊNCIO, R.A.; NUNES, B.P. Machine learning for predicting chronic diseases: a systematic review, **Public Health**, v. 205, p.14-25, 2022.

DUBOZ, P.; BOËTSCH, G.; GUEYE, L.; MACIA, E. (2014). Hypertension prevalence, awareness, treatment, and control in Dakar (Senegal). **Journal of Human Hypertension**, v. 28, n. 8, p.489–493.

ELSHAWI, M.H.R.; AL-MALLAH, M.H.; SAKR, S. (2019). On the interpretability of machine learning-based model for predicting hypertension. **BMC Med Inform Decis Mak**, v.19, n.1, p.146, dez, 2019.

FENG, C.; JIAO, J. (2021). Predicting and mapping neighborhood-scale health outcomes: a machine learning approach. **Computers, Environment and Urban Systems. Environ. Urban Syst.**, v.85, 101562, jan.2021.

FERDOUSI, R.; HOSSAIN, M.; EL SADDIK, A. Early-Stage risk prediction of non-communicable disease using machine learning in health cps. **IEEE Access**, v. 9, p. 96823-96837, 2021.

HAJAT, C; STEIN, E. (2018). The global burden of multiple chronic conditions: a narrative review. **Preventive Medicine Reports**, v. 12, p. 284-293, 2018.

HOU, J.; YANG, X.; CHEN, C. Emerging trends and new developments in information science: a document co-citation analysis (2009–2016). **Scientometrics** , v. 115, p. 869–892, 2018.

ICHIKAWA, D.; SAITO, T.; OYAMA H. Impact of predicting health-guidance candidates using massive health check-up data: a data-driven analysis. **International Journal of Medical Informatics**, v.106, p.32-36, out. 2017.

JEYASEKAR, J. J.; SARAVANAN, P. (2015). Impact of collaboration on indian forensic science research: a scientometric mapping from 1975 to 2012. **Journal of Scientometric Research**, v. 4, n. 3, p.135-142.

KITCHENHAM, B. **Procedures for performing systematic reviews**. Keele, UK, Keele University, 2004.

LI, M.; PORTER, A. L.; WANG, Z. L. Evolutionary trend analysis of nanogenerator research based on a novel perspective of phased bibliographic coupling. **Nano Energy**, v. 34, p. 93-102, 2017.

LUO, W.; NGUYEN, T.; NICHOLS, M.; TRAN, T.; RANA, S.; GUPTA, S.; PHUNG, D.; VENKATESH, S.; ALLENDER, S.; DOWD, J. B. Is demography destiny? application of machine learning techniques to accurately predict population health outcomes from a minimal demographic dataset. **Plos One**, v. 10, n. 5, p. e0125602, 2015.

MORGENSTERN, J.D.; BUAJITTI, E.; O'NEILL, M. et al. Predicting population health with machine learning: a scoping review. **BMJ Open**, v.10, e037860, 2020.

MPHEKGWANA, P.M.; MALEMA, N.; MONYEKI, K.D.; MOTHIBA, T.M.; MAKGAHLELA, M.; KGATLA, N.; MAKGATO, I.; SODI, T. Hypertension prevalence and determinants among black south african adults in semi-urban and rural areas. **Int J Environ Res Public Health**, v. 17, n. 20, p. 7463, out. 2020.

NETTO, R.; FABRE, S.; FONTANA, T.A.; LIVRAMENTO, V.; PILLA, L.L.; BEHJAT, L.; GUNTZEL, J.L. Algorithm selection framework for legalization using deep convolutional neural networks and transfer learning. **IEEE Transactions on computer-aided design of integrated circuits and systems**, IEEE, 2021b.

O'DONOVAN, M.; SEZGIN, D.; KABIR, Z.; LIEW, A.; O'CAOIMH, R. Assessing global frailty scores: development of a global burden of disease-frailty index (GBD-FI). **Int J Environ Res Public Health**, v. 17, n. 16, p. 5695.

PAHO¹ (2022). **Chefes de estado se comprometem com Pacto Global de Doenças Não Transmissíveis para salvar 50 milhões de vidas até 2030**. Disponível em : <https://www.paho.org/pt/noticias/21-9-2022-chefes-estado-se-comprometem-com-pacto-global-doencas-nao-transmissiveis-para>. Acesso em: 17.mar.2023.

PAHO¹ (2022). **Chefes de estado se comprometem com Pacto Global de Doenças Não Transmissíveis para salvar 50 milhões de vidas até 2030**. Disponível em : <https://www.paho.org/pt/noticias/21-9-2022-chefes-estado-se-comprometem-com-pacto-global-doencas-nao-transmissiveis-para>. Acesso em: 17.mar.2023.

PAHO² (2022). **Quinquennial report 2018-2022 of the director of the pan american sanitary bureau**. Disponível em: https://www.paho.org/sites/default/files/od366-e-quinquennial-report-director-paho-2018-2022-rep1_0.pdf. Acesso em: 10.abr.2023.

PINTO-FILHO, M. M., BRANT, L. C., DOS REIS, R. P., GIATTI, L., DUNCAN, B. B., LOTUFO, P. A., DA FONSECA, M. J. M., MILL, J. G., DE ALMEIDA, M. D. C. C., MACFARLANE, P., BARRETO, S. M., & RIBEIRO, A. L. P. (2021). Prognostic value of electrocardiographic abnormalities in adults from the Brazilian longitudinal study of adults' health. **Heart (British Cardiac Society)**, v. 107, n. 19, p. 1560–1566.

PRICE, D. S. A General theory of bibliometric and other cumulative advantage processes. **Journal of the American Society for Information Science**, v. 27, n. 5, p. 292–306, 1976.

RODRIGUEZ-ALMEIDA, A.J.; FABELO, H.; ORTEGA, S.; DENIZ, A. BALEA-FERNANDEZ, A. J.; et al. Synthetic Patient Data Generation and Evaluation in Disease Prediction Using Small and Imbalanced Datasets. **IEEE J Biomed Health Inform**, v. PP, aug.,2022.

ROSINI, N., MACHADO, M. J., WEBSTER, I. Z., MOURA, S. A., CAVALCANTE, L.DAS., & DA SILVA, E. L. (2013). Simultaneous prediction of hyperglycemia and dyslipidemia in school children in Santa Catarina State, Brazil based on waist circumference measurement. **Clinical biochemistry**, v. 46, n.18, p. 1837–1841.

SANTOS, H. G. Comparação da performance de algoritmos de machine learning para a análise preditiva em saúde pública e medicina. 2018. **Tese** (Doutorado Em Epidemiologia) – Faculdade De Saúde Pública, Universidade De São Paulo, São Paulo, 2018.

SANTOS, H.G.; NASCIMENTO, C.F.; IZBICKI, R.; DUARTE, Y.A.O.; CHIAVEGATTO FILHO, A.D.P. Machine learning para análises preditivas em saúde: exemplo de aplicação para prever óbito em idosos de São Paulo, Brasil. **Cadernos de Saúde Pública**, 2019, v. 35, n. 7, p.e00050818.

SCOPUS. **Scopus content coverage guide**. Elsevier, 2017. Disponível em: https://www.elsevier.com/___data/assets/pdf_file/0007/69451/0597-Scopus-Content-Coverage-Guide-US-LETTER-v4-HI-singles-no-ticks.pdf. Acesso em 12.dez.2022.

SMALL, H. (1973). Co-Citation in the scientific literature: a new measure of the relationship between two documents. **Journal of the American Society for Information Science**, v.24, n.4, p. 265-269.

TAPAK, L.; MAHJUB, H.; HAMIDI, O.; POOROLAJAL, J. (2013). Real-data comparison of data mining methods in prediction of diabetes in Iran. **Healthcare informatics research**, v. 19, n. 3, p. 177–185.

TIWARI, A., CHUGH, A., SHARM, A. Ensemble framework for cardiovascular disease prediction. **Computers in Biology and Medicine**, v. 146, 105624, 2022.

TRANFIELD, D.; DENYER, D.; SMART, P. Towards a methodology for developing evidence-informed management knowledge by means of systematic review. **British Journal of Management**, v. 14, n. 3, p.207-222, 2003.

VAN ECK, N. J.; WALTMAN, L. Pesquisa de software: VOSviewer, um programa de computador para mapeamento bibliométrico. *Cienciometria*, 84, 523-538, 2010.

WALLIN, J. A. Bibliometric methods: pitfalls and possibilities. **Basic & Clinical Pharmacology & Toxicology**, v. 97, n. 5, p. 261-275, 2005.

WANG et al. (J Infect Dis 2016; 213:1380–7), **The Journal of Infectious Diseases**, v. 214, n. 10, p. 1613, nov, 2016.

WANG, X. C.; ZHAI, M. M.; REN, Z. P.; REN, H.; LI, M.C. et al. Exploratory study on classification of diabetes mellitus through a combined random forest classifier. **BMC MEDICAL INFORMATICS AND DECISION MAKING**, v. 20, n.21, p.1: art. 105, mar, 2021.

WANG, Y.; WANG, J. Modelling and prediction of global non-communicable diseases. **BMC Public Health**, v. 20, n. 1, p. 1-13, 2020.

WEBBER, L.; KILPI, F; MARSH, T.; RTVELADZE, K.; BROWN, M.; et al. (2012) High Rates of Obesity and Non-Communicable Diseases Predicted across Latin America. **PLOS ONE**, v. 7, n. 8, p. e39589.

CAPÍTULO 4 – MODELAGEM E PREVISÃO DE OBESIDADE E DIABETES: UM ESTUDO A PARTIR DO INQUÉRITO TELEFÔNICO VIGITEL 2021

RESUMO

O Diabetes Mellitus tipo 2 (DM2) e a Obesidade (OBY), ambas doenças crônicas não transmissíveis (DCNTs), impactam a utilização e aumento dos custos de cuidados de saúde. A previsão, através de inteligência artificial (IA), para resultados clínicos futuros pode auxiliar na eficiência e prontidão da saúde. Como objetivo, esse estudo busca avaliar o desempenho de modelos de redes neurais artificiais com capacidade de predição de obesidade e diabetes no Brasil a partir de variáveis colhidas no inquérito VIGITEL 2021. Para tanto, com base no Relatório VIGITEL, definiram-se treze variáveis de entrada e avaliou-se o desempenho de modelos de predição de IA para Diabetes (DIA) e OBY no Brasil, segundo dados de 24.217 mil respondentes do inquérito. Como modelo de saída, realizaram-se três testes, considerando apenas obesidade, diabetes e ambos. Utilizou-se um modelo de previsão baseado em algoritmo de aprendizado de máquina para Perceptron Multicamadas (MLP) para predição de risco para DIA e OBY. Com classificação (ótima), o modelo de regressão linear (MRL) que melhor previu OBY foi (MLP 13-2-1); com classificação (boa) MRL que melhor previu OBY e DIA foi (MLP 13-2-2); com classificação (fraco), o MRL que pior previu Diabetes (MLP 13-2-1). Conclui-se que as previsões com uso de IA podem ajudar no desenvolvimento de políticas e estratégias de saúde que reduzam a carga de doenças em ambientes com recursos limitados.

Palavras-chave: Sistemas de Apoio à Decisão, Redes Neurais Artificiais, Tecnologia e Saúde, DCNTs.

ABSTRACT

Type 2 Diabetes Mellitus (DM2) and Obesity (OBY), both non-communicable chronic diseases (NCDs), impact the use and increase of health care costs. Artificial intelligence (AI) prediction of future clinical outcomes can aid in healthcare efficiency and readiness. Based on VIGITEL, thirteen input variables were defined and the performance of AI prediction models for Diabetes (DIA) and OBY in Brazil was evaluated, according to data from 24,217 thousand survey respondents. As an output model, we performed three tests considering only obesity, diabetes, and both. A prediction model based on machine learning algorithm for Multilayer Perceptron (MLP) was used for risk prediction for DIA and OBY. Results: Rated (optimal) - linear regression model (MRL) that best predicted OBY was (MLP 13-2-1); with (good), MRL that best predicted OBY and DIA was (MLP 13-2-2); with (weak), the MRL that worst predicted Diabetes (MLP 13-2-1). It is concluded that predictions using AI can help in the development of health policies and strategies that reduce the burden of disease in resource-limited settings.

Keywords: Decision Support Systems, Artificial Neural Networks, Technology and Health, NCDs.

1 INTRODUÇÃO

Historicamente, entre as doenças crônicas não transmissíveis (DCNTs), o *diabetes mellitus* (DM) representa a quarta maior causa de mortalidade entre a população global. Da década de 1980 até o ano de 2014, houve um incremento de quatro vezes na prevalência do diabetes, passando de aproximadamente 100 milhões de pessoas para 420 milhões atingidas pela doença, principalmente na população adulta mundial (aumento relevante de 4,7% para 8,5% na taxa padronizada por idade) (OGURTSOVA et al., 2018).

Como condição crônica que atinge severamente as funções do pâncreas, os *Diabetes Mellitus* tipo 1 (DM1) e tipo 2 (DM2) trazem prejuízos que resultam no aumento da taxa de glicose do sangue, causando complicações como a hipertensão, a insuficiência renal, a necessidade de amputar membros inferiores em alguns casos, a possibilidade de cegueira, de acidentes vasculares cerebrais e de danos aos nervos (AMPOFO et al., 2020).

Segundo a Organização Mundial da Saúde – OMS, a obesidade, que se define como índice de massa corporal (IMC), igual ou maior a 30,0 kg/m², associa-se como um fator de risco para o diabetes, incluindo-se ainda fatores de malefícios como estilo de vida não saudável e hábitos como o tabagismo e o alcoolismo. Da mesma forma, sua associação pode ser vista em conjunto às doenças cardiovasculares, ao câncer e ao acidente vascular encefálico. Adiciona-se ao conjunto de fatores para obesidade e diabetes, os hábitos alimentares ricos em alta densidade energética, gorduras e açúcares, além da sedentariedade.

Países desenvolvidos, com renda mais elevada que a maioria global têm observado o crescimento da obesidade, principalmente nas últimas três décadas. Da mesma forma, nos países com baixa e média renda, a incidência da obesidade triplicou, em uma linha contínua de crescimento.

A relação que se verifica entra a obesidade e DM é compulsória ao futuro da humanidade, como uma tendência difícil de ser freada. Diferentes estudos mostram que a DM muda em prevalência e dinâmica, nas diferentes regiões, segundo Ampfo et al. (2020)¹⁵. Outros fatores como idade e urbanização não foram contabilizados em

¹⁵ Whiting et al. (2011); IDF Diabetes Atlas (2015); Ogurtsova et al. (2018); Guariguata et al. (2014); Cho et al. (2018); Lin et al. (2018); Saeedi et al. (2019); Finkelstein et al. (2012); McPherson et al. (2019).

alguns desses estudos. Assim, a tendência dinâmica subjacente da prevalência de DM associada à obesidade necessita de maiores estudos, visto que através desses modelos, não foi possível agrupar as transições globais das doenças, sendo necessários novos estudos para realizar percepções ou previsões expressivas e prospectivas.

Estudo da World Health Organization (2013), destaca que apesar dos inúmeros fatores de risco para o DM, riscos modificáveis, como a obesidade, devem ser considerados, seguindo a meta global que visa à redução de mortalidade prematura por DM e obesidade (de forma simultânea), conforme preceituado pelo Plano de Ação Global para a prevenção e controle de DCNTs até 2025 (WHO, 2013). Como contorno de contribuição para as lacunas vistas em diferentes pesquisas sobre o DM e os fatores que se associam a ele como riscos agravantes, a abordagem sobre o aprendizado das máquinas (*Machine-learning* – ML) apresenta-se como uma maneira de mapear de forma auto-organizada e agrupada, porém de uso ainda escasso, na identificação de padrões da prevalência do DM em associação à obesidade em 183 países (AMPOFO et al., 2020).

Uma rede neural é um método de inteligência artificial que ensina computadores a processar dados de uma forma inspirada pelo cérebro humano. É um tipo de processo de *machine learning*, chamado aprendizado profundo, que usa nós ou neurônios interconectados em uma estrutura em camadas, semelhante ao cérebro humano (DAVENPORT; KALAKOTA, 2019).

Assim, este estudo busca avaliar o desempenho de modelos de redes neurais artificiais com capacidade de predição de obesidade e diabetes no Brasil a partir de variáveis colhidas no Relatório da Vigilância de Fatores de Risco e Proteção para Doenças Crônicas em Inquérito Telefônico VIGITEL 2021.

O VIGITEL é realizado anualmente através de pesquisa telefônica, em forma de inquérito, e abrange todos as capitais dos Estados brasileiros, incluindo-se o Distrito Federal. Nele, estima-se com que frequência e como se distribuem os principais fatores de riscos e proteção para cada DCNT (BRASIL, 2021b). Um fato relevante na pesquisa VIGITEL - um dos relatórios mais amplos sobre saúde no país - é que o estudo consiste em coletar os dados e descrever a evolução anual dos

índices da amostra populacional, como: obesidade, hipertensão, diabetes e mais (BRASIL, 2021b).

2 REFERENCIAL TEÓRICO

2.1 Relação entre obesidade e diabetes

Definem-se a obesidade e o sobrepeso como a acumulação de tecido adiposo em excesso, estendendo-se a níveis prejudiciais à saúde corpórea, psicossocial, e que afeta o bem-estar pessoal. Considerada uma ameaça à condição de salubridade a nível global, preocupa as autoridades governamentais. Sua prevalência impacta significativamente os custos econômicos da saúde, principalmente pelos riscos sociais que apresenta e pelas altas taxas de morbidade e mortalidade em decorrência da doença (AL-GOBLAN et al., 2014).

A origem da obesidade está ligada, principalmente, a um estilo de vida não saudável e sedentário, além do consumo de alimentos superindustrializados, que substituem a alimentação natural, e a reposição *in natura* de necessidades alimentares. Mudanças sociais e econômicas, que Hruby & Hu (2015) chamam de “obesogênicas”, incluem menor nível de esforço (ao se movimentar, locomover), uso de recursos tecnológicos avançados (no trabalho e em casa – os mecanismos automatizadores como a “Alexa”), e compõem o aumento de chance de desenvolvimento da obesidade (HRUBY; HU, 2015). Suas consequências podem ser minimizadas com a manutenção de hábitos de vida saudáveis e intensificação da atenção à disposição corpórea para manter o equilíbrio e a condição física (OLIVEIRA et al., 2011).

Importantes fatores que representam risco para a obesidade estão relacionados a fatores genéticos e características étnicas e raciais, além do ambiente social, econômico e cultural em que as pessoas vivem. Considera-se que os comportamentos de cada indivíduo em resposta a esses fatores e condições representam agentes que podem ser decisivos na prevenção da obesidade (HRUBY; HU, 2015).

A obesidade e o sobrepeso são representados por um IMC elevado do indivíduo, maior ou igual a 30 kg/m². Entretanto, o excesso de gordura corporal (que também é encontrada em 1/3 da população considerada com peso normal) em relação à massa magra, pode ser subestimada, principalmente quando se concentra em forma de gordura visceral, aumentando o risco de consequências, como a arteriosclerose e mortalidade (LEITNER et al., 2017).

Conhecida como relevante fator de risco para o diabetes mellitus de tipo 2, a obesidade pode alcançar 80 a 90% dos indivíduos que têm DM, associando-se à elevação de índice da massa corporal. Outras consequências do excesso de gordura se refletem em outras doenças como a hipertensão, e taxas elevadas de triglicérides e de colesterol (ESCOBAR, 2009).

Para Araújo et al. (2022), considera-se que a obesidade, e efetivamente a gordura visceral acumulada no organismo, compreendem fatores de elevada gravidade para o risco cardiovascular, assim como para os distúrbios na homeostase glicose-insulina. Então, ocorrem diferentes alterações fisiopatológicas, que provocam um funcionamento do fígado que extrai menor quantidade de insulina e produz maior quantidade de glicose, diminuindo a captação da glicose pelo tecido muscular. Como respostas a essas ocorrências, pode haver intolerância à glicose em diferentes graus em pessoas que têm DM2, dificultando o controle glicêmico adequado, que leva a maiores níveis de hemoglobina glicosilada (HbA1c) (SILVEIRA; VIEIRA; SOUZA, 2018). Os pacientes acometidos por DM2 têm dificuldades para manter o controle glicêmico, o que representa fator de risco para o aumento de complicações microvasculares (SANTOS et al., 2015).

As consequências e complicações do DM2 têm frequências variadas, a depender da população em estudo. Estudos mostram que as pessoas com DM2 estão propensas a taxas de mortalidade por doenças cardíacas com relação a não diabéticos, em um grau de duas a quatro vezes maior. No caso de doença vascular periférica (DVP), existe a propensão quatro vezes maior, assim como para os acidentes vasculares cerebrais (AVC). Fatores etiológicos nessa relação referem-se à hiperglicemia, dislipidemia, hipertensão arterial sistêmica e tabagismo (QUEIROZ et al., 2011; SCHEFFEL, et al. 2004).

Autores defendem que a possibilidade de DM2 decorrer dos elevados níveis de obesidade nos pacientes brasileiros é bastante provável (QUEIROZ et al., 2011;

SCHEFFEL, et al. 2004). Há crescimento do número de obesos no país e, de acordo com o Relatório VIGITEL (2021), considera-se que aproximadamente seis de cada dez indivíduos brasileiros estão acima do peso, representando 59,9% entre os homens e 55% das mulheres entrevistadas.

Para utilizar essas informações com efetividade, visando ações de predição e encaminhamento de ações corretivas e preditivas, é utilizada a inteligência artificial.

2.2 Uso de inteligência artificial

No ramo da tecnologia, diversos métodos vêm sendo desenvolvidos a fim de auxiliar no combate à epidemia. Dentre essas técnicas, a inteligência artificial (IA) vem sendo aplicada e vista no ramo da saúde como solução inovadora, justamente por incorporar os benefícios da tecnologia e da psicologia, atendendo diferentes necessidades de saúde por meio de várias mídias (STEPHENS et al., 2019).

A rede neural cria um sistema adaptativo que os computadores usam para aprender com os erros e se aprimorar continuamente. As redes neurais artificiais tentam solucionar problemas complicados, como resumir documentos ou reconhecer rostos com grande precisão (DAVENPORT; KALAKOTA, 2019).

O cérebro humano é a inspiração por trás da arquitetura da rede neural. As células do cérebro humano, chamadas de neurônios, formam uma rede altamente interconectada e enviam sinais elétricos entre si para ajudar os seres humanos a processarem informações. Da mesma forma, uma rede neural artificial é feita de neurônios artificiais que trabalham juntos para resolver um problema. Os neurônios artificiais são módulos de *software*, chamados de “nós”, e as redes neurais artificiais são programas ou algoritmos de *software* que, em seu núcleo, usam sistemas de computação para solucionar cálculos matemáticos (DAVENPORT; KALAKOTA, 2019).

As redes neurais podem ajudar os computadores a tomarem decisões inteligentes com assistência humana limitada. Isso ocorre porque elas aprendem e modelam relacionamentos entre os dados de entrada e de saída complexos e não lineares (DAVENPORT; KALAKOTA, 2019).

Os modelos de rede neural têm vantagem sobre os métodos estatísticos por serem livres de distribuição e não serem necessários conhecimentos prévios sobre as

distribuições estatísticas das classes nas fontes de dados, a fim de aplicar esses métodos para classificação (FLECK et al., 2016). Na área da saúde, as Redes Neurais Artificiais (RNA) são interessantes como métodos para modelar tarefas repetitivas e extensas, ou nos casos em que muitos detalhes são considerados nas investigações (diagnósticos, monitoramento de pacientes, telemedicina, dentre outros) (HAYKIN, 2008).

Belciug e Gorunescu (2014) desenvolveram uma nova técnica de aprendizado baseada em RNA bayesiana para o diagnóstico de diabetes, com base em um banco de dados de relatórios médicos. Em paralelo, desenvolveram também algoritmos para diagnosticar câncer de mama, câncer de pulmão e infarto agudo do miocárdio. Após análise estatística, o método proposto mostrou superar as técnicas convencionais de diagnóstico em quase todos os aspectos.

Em relação ao diagnóstico de complicações do DM, Pratt et al. (2016) utilizaram RNA para construir um sistema para diagnóstico da retinopatia diabética, através de imagens de fundo digital, e classificar com precisão a gravidade da doença. O sistema é capaz de verificar a presença de microaneurismas, exsudato e hemorragia na retina, além de realizar a classificação.

Tais sistemas adaptativos de IA, baseados em RNA, buscam treinar as redes neurais para identificar, por exemplo, as características do diabético e enviar imagens em tempo real aos médicos, que podem classificar e avaliar com precisão a condição clínica do paciente (PRATT et al., 2016). Por essa razão, as redes neurais se mostram úteis para a área da saúde nas questões diagnósticas e por manter bancos de dados atualizados e contínuos sobre os pacientes, que podem ser acessados por qualquer profissional que participe da rede

3 METODOLOGIA

3.1 Características da pesquisa

A presente pesquisa possui natureza empírica e de abordagem quantitativa. Considerando que a discussão sobre a temática da *Ciência de Dados em Saúde* é

recente no país, dada a escassez de pesquisas relacionadas ao fenômeno, o estudo caracteriza-se como exploratório, no que tange à evolução do tema. A pesquisa exploratória tem por característica a descoberta sobre determinado assunto tornando o problema familiar ao pesquisador; serve para aprimorar ideias e descobrir intuições (GIL, 2017).

Além disso, por visar compreender a relação existente entre os fatores de risco e proteção das DCNTs e o estilo de vida dos indivíduos, a pesquisa caracteriza-se também como explicativa. Segundo Gil (2017), os estudos explicativos preocupam-se em identificar os fatores determinantes ou que contribuam para que ocorra um determinado fenômeno. Para tal, necessitam de uma análise estatística mais aprofundada como a análise multivariada (RICHARDSON et al., 2012).

3.2 Modelos de predição e avaliação de desempenho para obesidade e diabetes

Por meio deste estudo, a intenção é validar um modelo preditivo de Obesidade e Diabetes do tipo 2, através da utilização de aprendizagem de máquina por meio de diferentes variáveis, que estejam disponíveis em meios de fácil acesso, como os resultados dos bancos de dados VIGITEL.

A rede neural proposta neste trabalho (Figura 12) foi desenvolvida implementando-se as técnicas de redes neurais Feedforward e Backpropagation, com função sigmóide comum para uso na aprendizagem da rede. Além disso, a arquitetura da RNA utilizada foi de duas camadas (oculta e saída), conforme ilustra a Figura 12. Para a modelagem, foram utilizadas redes neurais artificiais, como modelo *Multi Layer Perceptron (MLP)* com duas camadas.

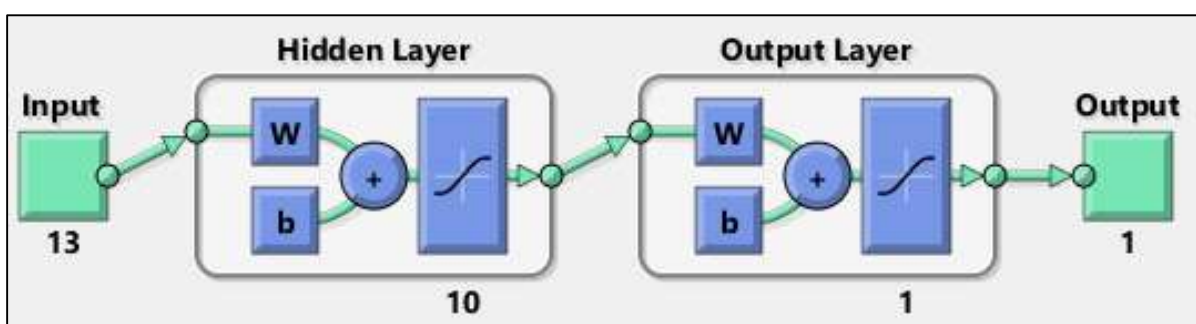


Figura 12 – Estrutura do Modelo Multi Layer Perceptron (MLP)
 Fonte: elaborado pelos autores a partir do modelo de Haykin (2008).

A rede neural *perceptron* de múltiplas camadas, também conhecida como MLP, é um sistema de neurônios ligados por conexões sinápticas divididos em camadas de entradas, neurônios escondidos e neurônios de saída. A MLP foi concebida para resolver problemas mais complexos, pela sua arquitetura de neurônios escondidos. (BEAGLE; JACKSON, 1994).

Para analisar o desempenho dos modelos testados, foram realizadas a sequência de testes com os dados modelados (dados de saída) e os dados reais do banco de dados VIGITEL. A análise dos dados será realizada através da regressão linear ($y = a + bx$), em que x representa os dados reais e y representa os valores estimados pelo modelo.

Para a avaliação, serão utilizados os seguintes índices: erro médio, índice de concordância, coeficiente de correlação e índice de desempenhos. Segundo Valença (2005), esses são os índices mais utilizados para avaliação de desempenho de redes neurais.

O *erro médio* (ME) foi calculado através da equação 01:

$$ME = \frac{\sum_{i=1}^n 1 (E_i - O_i)}{n} \quad (01)$$

Sendo E_i é o valor estimado, O_i , é o valor real e n número observações.

O coeficiente de correlação (c) será estimado pela equação (02):

$$r = \frac{\sum_{i=1}^n (E_i - E) \cdot (O_i - O)}{\sqrt{[\sum (E_i - E)^2] \cdot [\sum (O_i - O)^2]}} \quad (02)$$

Sendo E_i é o valor estimado, O_i , é o valor real e n número observações, E a média dos valores estimados e O à média dos valores observados.

O índice de concordância (c) fornece o grau de exatidão entre as variáveis envolvidas, uma vez que está relacionada a diferença entre os valores estimados em relação aos valores observados, sendo seu campo de variação de 0 (nenhuma concordância) a 1 concordância perfeita (WILLMOUT, 1981).

O índice de concordância foi calculado pela expressão (03):

$$C = 1 - \frac{\sum(E_i - O_i)^2}{\sum(|E_i - E|) + (|O_i - O|)^2} \quad (03)$$

Sendo E_i é o valor estimado, O_i , é o valor real e n número observações.

Conhecendo-se esses indicadores foi determinado o índice de desempenho ID, que segundo Camargo e Sentelhas (1997), pode ser calculado pela equação (04):

$$ID = r \cdot c \quad (04)$$

Onde r é o coeficiente de correlação e c o índice de concordância.

O índice ID tem a finalidade de avaliar o desempenho do método proposto, considerando as seguintes classes de interpretação (Quadro 6), de acordo com Costa (2004).

Classes	Valores de ID	Desempenho
1	> 0.85	Ótimo
2	0.76 a 0.85	Muito Bom
3	0.66 a 0.75	Bom
4	0.61 a 0.65	Regular
5	0.51 a 0.60	Fraco
6	0.41 a 0.50	Muito Fraco
7	< 0.41	Péssimo

Quadro 6 – Classificação do índice de desempenho

Fonte: Costa (2004).

Com base nos dados obtidos na análise, foi realizada a classificação de cada modelo proposto (sendo utilizados três modelos, que estão na Tabela 14: um teste de algoritmo apenas para Obesidade, um segundo para Obesidade e Diabetes e finalmente, um terceiro apenas para Diabetes), realizando a validação para estimação de indivíduos com obesidade ou diabetes ou obesidade e diabetes.

3.3 Dados Secundários – banco de dados

A técnica de coleta escolhida para a pesquisa é coleta de dados em banco de dados (GIL, 2017), também chamada de Pesquisa em base de dados/informações, que visa a análise de informações da base VIGITEL (Vigilância de Fatores de Risco e Proteção para Doenças Crônicas por Inquérito Telefônico) elaborada pela Coordenação Geral de Doenças e Agravos Não Transmissíveis, Departamento de Análise de Situação de Saúde e Secretaria de Vigilância em Saúde do Ministério da Saúde do Brasil.

O pré-processamento de dados é uma etapa crucial em que o conjunto de dados é preparado para treinamento antes de construir os modelos de classificação. Existem várias etapas importantes no pré-processamento de dados, como limpeza de dados, dimensionamento e seleção de recursos. Foram utilizados os dados da VIGITEL do ano de 2021, com inicialmente 27.093 respondentes, no qual, ao verificar as inconsistências e resposta faltantes, buscou-se modelar matematicamente um conjunto de dados de 24.217 mil respondentes do inquérito que estavam completos.

Como forma de compreender as possíveis variáveis, foram utilizadas como base as recomendações fornecidas pela OMS, chamada *Package of Essential Noncommunicable (PEN) Disease Interventions*, ou Pacote de Intervenções Essenciais em Doenças Não Transmissíveis. No PEN da OMS, realizado em relação aos cuidados primários que ocorrem em cenários de baixos recursos, se visa conjugar ações em um conjunto de intervenções prioritárias, que conjuguem custo e efetividade à um atendimento de qualidade aceitável. Seu intuito é reforçar e fortalecer o sistema de saúde, contribuindo para a formação de pilares sólidos na atenção primária, mesmo com recursos limitados, direcionados ao tratamento das DCNTs.

No Quadro 7, pode-se observar as variáveis utilizadas no trabalho, com base nas variáveis utilizadas pelo Relatório VIGITEL 2021:

Código	Descrição	Valor
Idade	idade (em ano)	anos
Sexo	Sexo	1 - masculino 2 - feminino
Peso	peso (em kg)	quilos
Altura	Altura em cm	centímetros
Álcool	Consumo de álcool	0 - não 1 - sim
Exercício	Pratica exercício físico	0 - não 1 - sim
Fumante	Fumante	0 - não 1 - sim
Flvreco	Consumo recomendado de frutas e hortaliças.	0 - não 1 - sim
score_sf_2cat	Consumo de cinco ou mais grupos de alimentos não ou minimamente processados protetores para doenças crônicas.	0 - não 1 - sim
score_upp_2cat	Consumo de cinco ou mais grupos de alimentos ultraprocessados.	0 - não 1 - sim
Inativo	Inatividade física	0 - não 1 - sim
Saruim	Avaliação ruim da condição de saúde	0 - não 1 - sim
depressão	Depressão	0 - não 1 - sim
Obesid	Obesidade	0 - não 1 - sim
Hart	Hipertensão arterial	0 - não 1 - sim
Diab	Diabetes	0 - não 1 - sim

Quadro 7 – Variáveis utilizadas no estudo

Fonte: dados da pesquisa (2023).

No modelo descrito no PEN, destacam-se os cinco tipos de intervenções sugeridas para detecção de doenças cardiovasculares, diabetes, doenças respiratórias crônicas, diagnóstico precoce do câncer e componentes de estilo de vida saudável.

4 RESULTADOS

Conforme exposto anteriormente, foram utilizadas variáveis disponíveis no inquérito do VIGITEL do ano de 2021, totalizando 24.217 respondentes válidos, considerando a normalização e refinamento do banco de dados. A amplitude da idade dos respondentes foi de 89 anos, considerando a faixa etária de 18 a 107 anos. Outros dados do perfil do respondente podem ser visualizados na Tabela 13:

VARIÁVEL	CLASSIFICAÇÃO	FREQUÊNCIA	
		ABSOLUTA	PERCENTUAL
SEXO	Masculino	8.803	36%
	Feminino	15.414	64%
CONSUMO DE ÁLCOOL	Não	6.107	25%
	Sim	18.110	75%
PRÁTICA DE EXERCÍCIO FÍSICO	Não	10.762	44%
	Sim	13.455	56%
FUMANTE	Não	22.551	93%
	Sim	1.666	7%
OBESO	Não	18.862	78%
	Sim	5.355	22%
DIABETE	Não	20.663	85%
	Sim	3.554	15%
HIPERTENSÃO	Não	14.743	61%
	Sim	9.474	39%
DEPRESSÃO	Não	21.093	87%
	Sim	3.124	13%

Tabela 13 – Diagrama de frequência da amostra VIGITEL ano 2021

Fonte: dados da pesquisa (2023).

Conforme a Tabela 13, pode-se perceber que a maioria dos respondentes são do sexo feminino, totalizando 64% da amostra e 36% do sexo masculino. Quando observada a informação do consumo de álcool, nota-se que 25% responderam não consomem bebidas alcólicas e 75% dos respondentes afirmam que consomem regularmente ou esporadicamente. Lembrando que, conforme a OPAS (2015), entre os principais fatores de risco para as doenças crônicas não transmissíveis (DCNTs) estão: a alimentação não saudável e o sedentarismo, que acarretam sobrepeso e obesidade, assim como o tabagismo e o uso abusivo de álcool.

Entre os respondentes, 22% são classificados como obesos, sendo que entre eles, 1.093 também possuem diabetes, duas das principais DCNTs, representando 20,4% do total de obesos. Este índice foi superior ao exposto no trabalho desenvolvido por Malta et al. (2013), que apresentaram uma amostra com 11,8% de obesos com diabetes em estudos de fatores associados à diabetes.

Ainda em relação aos obesos, aproximadamente 47,12% dos respondentes não realizam nenhum tipo de exercício físico e 87,4% apresentam hábitos alimentares ruim (alimentos ultraprocessados e industrializados). Considerando os dados observados e as variáveis disponíveis, foram gerados modelos de redes neurais

artificiais para predição das duas principais DCNTs que apresentam dados no VIGITEL 2021.

Os modelos foram testados utilizando a rede MLP com duas camadas internas. Para o processo de treinamento, foram selecionadas 70% do tamanho da amostra inicial e 30% para o processo de validação do modelo. Foram utilizadas 13 variáveis de entrada, sendo elas: idade, sexo, peso, altura, uso de álcool, praticante de exercício físico, fumante, uso frequente de legumes e verduras, consumo de alimentos naturais, consumo de alimentos processados, inatividade física e condição saudável de saúde. Como modelo de saída, foram realizados 3 testes considerando apenas obesidade, diabetes e ambos.

Na Tabela 14 é possível verificar a avaliação do desempenho dos três modelos testados. É possível observar que o modelo de redes neurais artificial para a predição de obesidade teve o melhor desempenho, apresentando um índice de desempenho (ID) de 0,977, classificado como ótimo. O coeficiente de correlação foi de 98,5%.

Tabela 14 - Treinamento e validação dos RNA's obtidos para estimar indivíduos com diabetes ou obesidade ou diabetes e obesidade.

MODELO	SAÍDA	ME	R	ID	CLASSIFICAÇÃO
MLP 13-2-1	Obesidade	-0,0013	0,985	0,977	Ótimo
MLP 13-2-2	Obesidade e Diabetes	0,037	0,829	0,758	Bom
MLP 13-2-1	Diabetes	0,050	0,594	0,558	Fraco

Fonte: elaborado pelo autor (2023).

O modelo de redes com duas saídas, ou seja, para a predição de indivíduos com obesidade e diabetes, apresentou um desempenho classificado como bom, obtendo um ID de 0,758. Esse resultado demonstra que o modelo deve ser utilizado com cautela e deve-se explorar as variáveis de entrada para otimizar os resultados. Já em relação ao modelo de predição de indivíduos com diabetes, a rede neural artificial demonstrou um desempenho fraco. Esse comportamento pode estar interferindo no modelo anterior, visto que a variável diabetes foi introduzida como saída juntamente com a obesidade, reduzindo a precisão do resultado do modelo.

No gráfico 1 é possível analisar a dispersão da estimativa de indivíduos com obesidade

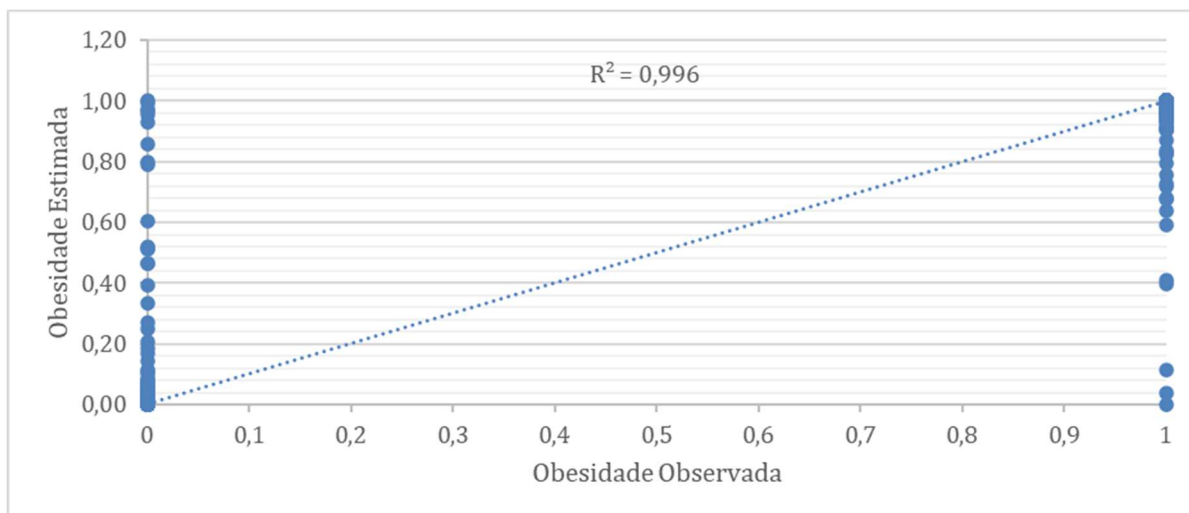


Gráfico 1- Dispersão entre os dados de obesidade observada e estimada

Fonte: resultados da pesquisa (2023).

No Gráfico 1 é possível verificar que os valores apresentam dispersão entre os valores observados e estimados. Isso é possível de perceber ao verificar os pontos de distância do valor real, sendo 0 para não obesos e 1 para obesos.

Já no Gráfico 2, é possível verificar a dispersão entre os indivíduos com diabetes, considerando os dados estimados e observados.

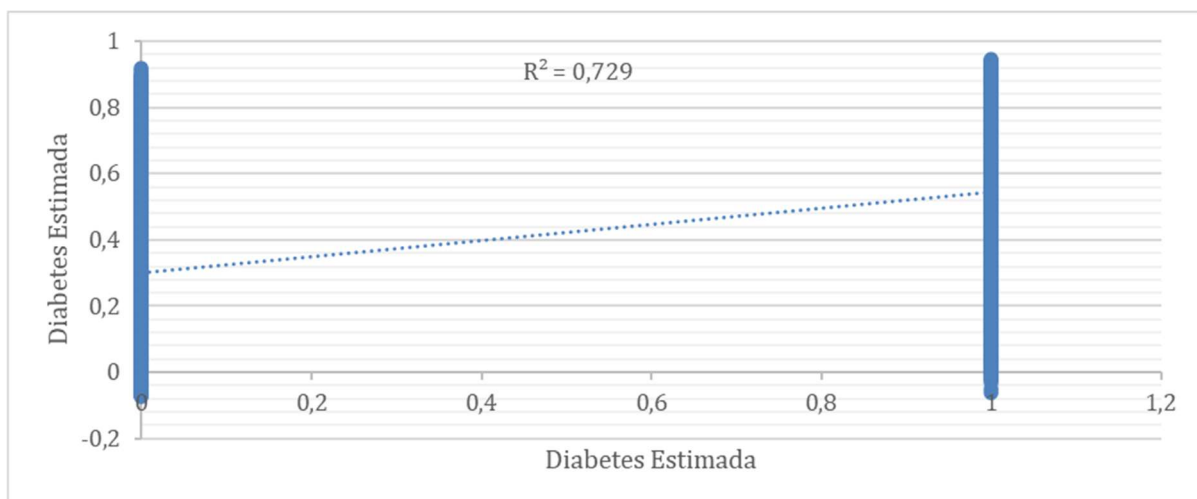


Gráfico 2- Dispersão entre os dados diabetes observada e estimada

Fonte: resultados da pesquisa (2023).

A dispersão para a estimativa de diabetes apresenta maior erro médio 0,050 (conforme Tabela 14), podendo ser evidenciado no gráfico ao estimar valores, considerando 0 para indivíduos sem diabetes e 1 para indivíduos com diabetes.

5 CONSIDERAÇÕES FINAIS

Como o presente trabalho, foi possível observar como os fatores de risco de doenças crônicas não transmissíveis estão presentes nos indivíduos, em especial, ao perceber os hábitos alimentares e físicos em pessoas obesas, corroborando com estudos e orientações disponibilizadas à sociedade. Ainda, ao realizar a modelagem com redes neurais artificiais utilizando o modelo *multi-perceptron* com as 13 variáveis de entradas, a predição de indivíduos obesos apresentou o melhor desempenho e pode ser aplicado em estudos e ações sobre a obesidade.

Já os modelos que incluíram a diabetes como variável de saída, não apresentou resultados satisfatórios, sendo o modelo de predição de obesos com diabetes o que apresentou desempenho bom, mas seu uso se restringe por apresentar um erro amostral maior. Para estudos posteriores, especialmente para a predição de indivíduos com diabetes, sugere-se avaliar novas variáveis de entrada para otimizar os resultados dos modelos, bem como a validação de novos modelos de *machine learning*.

De modo suplementar, o estudo baseado em algoritmos preditivos apresentou um modelo básico de vigilância digital, construído a partir de variáveis dentro de dados que já existem hoje, demonstrando que modelos artificiais conseguem prever com grande eficiência essa questão de obesidade e com potencialidades reais para diabetes.

REFERÊNCIAS CAPÍTULO 4

- AL-GOBLAN, A. S.; AL-ALFI, M. A.; KHAN, M. Z. Mechanism linking diabetes mellitus and obesity. **Diabetes, metabolic syndrome and obesity: targets and therapy**, v. 7, p. 587–591, 2014.
- AMPOFO, A.G.; BOATENG, E. B. Beyond 2020: modelling obesity and diabetes prevalence. **Diabetes Research and Clinical Practice**, v.167, 2020.
- ARAÚJO, G. B. et al. Relação entre sobrepeso e obesidade e o desenvolvimento ou agravamento de doenças crônicas não transmissíveis em adultos. **Research, Society and Development**, v. 11, n. 2, e50311225917, 2022.
- BELCIUG, S.; GORUNESCU, F. Error-correction learning for artificial neural networks using the Bayesian paradigm. Application to automated medical diagnosis. **Journal of Biomedical Informatics**, v. 52, n. 1, p. 329–37.
- BRASIL (2021b). Ministério da Saúde (MS). **Vigitel Brasil 2021: vigilância de fatores de risco e proteção para doenças crônicas por inquérito telefônico: estimativas sobre frequência e distribuição sociodemográfica de fatores de risco e proteção para doenças crônicas nas capitais dos 26 estados brasileiros e no Distrito Federal em 2021**. Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Análise em Saúde e Vigilância de Doenças Não Transmissíveis. Brasília: Ministério da Saúde, 2021.
- COSTA, S.V. Desenvolvimento e calibração de um mini tanque evaporimétrico. 2004. 80f. **Dissertação** (Mestrado em Engenharia Agrícola) Universidade Federal de Santa Maria, RS, 2004.
- DAVENPORT, T.; KALAKOTA, R. The potential for artificial intelligence in healthcare. **Future Healthcare Journal**, v. 6, n. 2, p. 94–98, 2019.
- ESCOBAR, F. A. Relação entre Obesidade e Diabetes Mellitus Tipo II em Adultos. **Cadernos UniFOA**. Volta Redonda, ano IV, n. 11, dezembro 2009. Disponível em: <http://www.unifoa.edu.br/cadernos/edicao/11/69.pdf>. Acesso em: 06.mar.2023.
- FLECK, L. et al. Redes Neurais Artificiais: princípios básicos. **Revista Eletrônica Científica Inovação e Tecnologia, Medianeira**, Paraná, v. 1, n. 13, p.47-57, jun. 2016.
- GIL, A.C. **Como Elaborar Projetos De Pesquisa**. 6.ed. São Paulo: Atlas, 2017.
- HAYKIN, S. **Neural networks and Learning Machines**, (3 ed.). Prentice Hall., 2008.
- HRUBY, A.; HU, F. B. (2015). The epidemiology of obesity: a big picture. **Pharmacoeconomics**, v. 33, n. 7, p. 673-89, jul, 2015.
- LEITNER, D. R.; FRÜHBECK, G.; YUMUK, V.; SCHINDLER, K.; MICIC, D.; WOODWARD, E.; TOPLAK, H. Obesity and type 2 diabetes: two diseases with a need for combined treatment strategies - EASO can lead the way. **Obesity Facts**, v. 10, n. 5, p. 483–492, 2017.
- MALTA, D. C.; ANDRADE, S. S. C. A.; OLIVEIRA, T. P.; MOURA, L.; PRADO, R. R.; SOUZA, M. F. M. (2019). Probabilidade de morte prematura por doenças crônicas não transmissíveis, Brasil e regiões: projeções para 2025. **Revista Brasileira de Epidemiologia**, v. 22, p. e190030.

OGURTSOVA, J.D.; ROCHA FERNANDES, Y.; HUANG, U.; LINNENKAMP, L.; GUARIGUATA, N.H.; CHO, D.; CAVAN, J.E.; SHAW, L.E.; MAKAROFF, I.D.F. Diabetes Atlas: global estimates for the prevalence of diabetes for 2015 and 2040. **Diabetes Research and Clinical Practice**, v.128, p. 40-50, 2018.

OLIVEIRA, E. N.; AGUIAR, R. C.; ALMEIDA, M. T. O.; ELOIA, S. C.; LIRA, T. Q. Benefícios da atividade física para saúde mental. **Saúde Coletiva**, v. 8, n. 50, p. 126-130, 2011.

OPAS – Organização Pan-Americana da Saúde. **Cuidados inovadores para condições crônicas**: organização e prestação de atenção de alta qualidade às doenças crônicas não transmissíveis nas Américas. Washington, DC: 2015.

PRATT, H., COENENB, F., BROADBENTC, D. M., HARDINGA, S. P., & ZHENG, Y. Convolutional Neural Networks for Diabetic Retinopathy. **Procedia Computer Science**, n.90, p. 200-205, 2016.

QUEIROZ, P.C.; AGUIAR, D.C.; PINHEIRO, R.P.; MORAES, C.C.; PIMENTEL, I.R.S.; FERRAZ, C.L.H.; FERRAZ, T.M.B.L. Prevalência das complicações micro e macrovasculares e de seus fatores de risco em pacientes com diabetes mellitus e síndrome metabólica. **Revista Brasileira De Clínica Médica**, V. 9, n. 4, p. 254-8, São Paulo, jul-ago 2011.

RICHARDSON, R. J. **Pesquisa Social: Métodos E Técnicas**. 3ed. São Paulo: Atlas, 2012.

SANTOS, A. L.; CECÍLIO, H. P. M.; TESTON, E. F.; ARRUDA, G. O.; PETERNELLA, F. M. N.; MARCON, S. S. Microvascular complications in type 2 diabetes and associated factors: a telephone survey of self-reported morbidity. **Ciência & Saúde Coletiva**, v. 20, n. 3, p. 761–770, 2015.

SCHEFFEL, R. S.; BORTOLANZA, D.; WEBER, C. S.; COSTA, L. A.; CANANI, L. H.; SANTOS, K. G.; CRISPIM, D.; ROISENBERG, I.; LISBÔA, H. R. K.; TRES, G. S.; TSCHIEDEL, B.; GROSS, J. L. Prevalência de complicações micro e macrovasculares e de seus fatores de risco em pacientes com diabetes melito do tipo 2 em atendimento ambulatorial. **Revista Da Associação Médica Brasileira**, v. 50, n. 3, p. 263–267, Porto Alegre, 2004.

SILVEIRA, E. A.; VIEIRA, L. L.; SOUZA, J. D. Elevada prevalência de obesidade abdominal em idosos e associação com diabetes, hipertensão e doenças respiratórias. **Ciência & Saúde Coletiva**, 23(3), 903–912, 2018.

STEPHENS, T N. et al. Feasibility of pediatric obesity and prediabetes treatment support through Tess, the AI behavioral coaching chatbot. **Translational Behavioral Medicine**, [s.l.], v.9, n.3, p.440-447, <http://dx.doi.org/10.1093/tbm/ibz043>, 2019.

WILLMOUT, C.J. On The Validation Of Models. **Physical Geography**, V.2, P.184-194, 1981.

WORLD HEALTH ORGANIZATION. **World report on ageing and health**. World Health Organization, 2013.

CAPÍTULO 5 – CONSIDERAÇÕES FINAIS DO ESTUDO

Dentre os dados armazenados no mundo inteiro, cerca de 30% são gerados pelo setor de assistência médica e de saúde, e boa parte deles tem um valor clínico, financeiro e operacional relevante para a indústria da saúde.

Diariamente, através dos atendimentos das Unidades Básicas de Saúde (UBS), no Brasil, geram-se um grande volume de informações, que podem ser aplicadas na construção de estratégias para apoiar ações de cuidado híbrido junto à população. Apesar dos avanços ocorridos no campo da vigilância das DCNTs, são muitas as lacunas que precisam ser preenchidas. Exemplificando, o Brasil pouco utiliza esses dados, que só podem ser aproveitados por meio de uma melhor interpretação e entendimento do comportamento dessas informações.

Nesse sentido, a utilização da Administração da Informação, através da implementação de tecnologias baseadas em IA, busca fornecer novos métodos para avaliar o desempenho e eficácia do tratamento de saúde juntos aos pacientes. A análise de dados por meio da IA pode fornecer *feedback* contínuo aos profissionais de saúde para que eles consigam revisar seus tratamentos prescritos. Inclusive, estas análises podem ajudar a identificar padrões que levam a uma melhor compreensão da saúde populacional. O trabalho de entendimento de dados promete transformar a compreensão da saúde, fornecendo novos *insights* sobre o desenvolvimento de terapias digitais, além da prestação de serviços de cuidados de saúde para comunidade e pessoal.

A aplicação desse novo paradigma na área da Administração requer processos que combinem a habilidade de adquirir, organizar e explorar os dados, e realizar a modelagem estatística, aplicando os algoritmos de *machine learning* para alcançar a predição e montar técnicas para visualização desses mesmos dados, fomentando uma adequada tomada de decisão.

Nesse contexto, a revisão sistemática de literatura sobre modelos preditivos para DCNTs, eliminou vieses e serviu para avaliar criticamente estudos primários. Nela, apresentou-se um arcabouço de pesquisas que foram feitas tratando deste tema, ou seja, mostrando ser este um tema relevante e de expressiva tendência de crescimento. Além disso, reuniram-se elementos teóricos que trouxeram a importância da utilização da IA no campo da saúde, além de apresentar estudos que já estão em

desenvolvimento, como potenciais trabalhos que precisam ser otimizados, apontando futuros caminhos para a Administração da Informação na Saúde.

A utilização do *Machine Learning* em áreas como a saúde consiste em uma tendência revolucionária, pois se constitui como exemplo das aplicações de Tecnologias da Informação e Comunicação (TICs) em atividades voltada à medicina e enfermagem, em um processo que atualmente é conceituado como saúde digital.

O modelo de redes neurais multicamadas, demonstrou ótimo desempenho para a estimação de indivíduos com obesidade e potencialidade nos estudos para predição de indivíduos com diabetes, considerando as 13 variáveis de entradas disponíveis no estudo. De modo suplementar, o estudo baseado em algoritmos preditivos apresentou um modelo básico de vigilância digital, construído a partir de variáveis dentro de dados que já existem hoje, demonstrando que modelos artificiais conseguem prever com grande eficiência essa questão de obesidade.

O rápido desenvolvimento de tecnologias que envolvem processos de IA tem acompanhado a chamada revolução 4.0, e se processa em cenário de amplificação de horizontes e aplicações possíveis, de forma vasta e sem retorno. Mesmo assim, a IA já cerca as inúmeras atividades cotidianas.

Portanto, já que a tecnologia do *Machine Learning* é uma grande aliada da medicina preventiva, a partir deste estudo é possível construir um produto, e esse produto pode ser utilizado para fazer previsões no mapeamento básico na área de saúde, permitindo identificar pessoas obesas ou diabéticas, e melhor, por meio de um simples questionário digital, gerar prognósticos da doença, que poderão ser confirmados ou não pelos profissionais de saúde na forma de diagnóstico.

Assim, de agora em diante, ao invés de realizar uma coleta de dados através de um inquérito telefônico (VIGITEL – BRASIL, 2021b), seria possível aplicar o modelo de vigilância digital em um questionário simplificado, de quinze questões, que se entende deter uma qualidade superior de previsão de DCNTs, já que a realização da coleta VIGITEL de 2h pode não ter foco em previsões, fornecendo apenas dados. Entretanto, esse questionário proposto no estudo como modelo preditivo tem importante diferencial, que é a possibilidade de identificar quem e onde estão aqueles indivíduos que possuem risco em desenvolver ou já apresentam as DCNTs instaladas.

Adicionado a isso, o estudo de identificação e mensuração da qualidade de vida, realizado a partir do VIGITEL 2021 com base no Método (AF)¹⁶, traz a possibilidade de melhorar e ampliar o desenvolvimento de políticas que se voltem às dimensões encontradas como de maior privação de qualidade de vida, na maioria dos respondentes entrevistados.

Além da multidimensionalidade, de serem consideradas duas linhas de corte, de possibilitar-se a decomposição em grupos e de segmentá-los segundo dimensões e indicadores, afirma-se que a medida de privação, na proposição do Método AF, é representante de avanços relacionados às outras medidas de pobreza estabelecidas anteriores, principalmente quando contribui na seara de políticas públicas.

Se faz urgente e necessário, expandir as capacidades dos indivíduos, através de uma cobertura elevada para acesso público à educação, às condições de prevenção e melhoria da saúde, e ao próprio saneamento básico, junto a outros demais problemas que atingem as pessoas menos favorecidas. Ao aumentar as capacitações nas pessoas com relação a esses âmbitos, encadeia-se a maximização de entendimento direto ou indireto dos direitos que a população tem para minimizar suas privações. A compreensão dos efeitos da pobreza nas pessoas, e de como ela afeta esses indivíduos, possibilita condições de planejamento efetivo em políticas públicas que reduzam a pobreza, oferecendo subsídios para a elaboração das mesmas e implementação adequada, focadas justamente naquelas privações que trazem sofrimento à população.

Adicionalmente, este estudo vem ao encontro das metas definidas no Plano Enfrentamento às DCNTs 2021-2030, elaborado pelo Ministério da Saúde, bem como à Estratégia de Saúde Digital para o Brasil 2020-2028 (BRASIL, 2021a, p. 66). Por fim, é notório que há disponibilidade de dados, mas existe deficiência de ferramentas para transformar esses dados em informação qualificada. Considera-se que a partir desta pesquisa foi possível comprovar que a utilização desses novos modelos de IA e das novas formas tecnológicas de modelagem de análise de dados, pode-se contribuir para predição, diagnóstico e gestão mais eficiente para a Administração da Informação e da Saúde.

¹⁶ A estimativa de um índice de qualidade de vida pelo método Alkire-Foster está inserida num conjunto de ações de vigilância das DCNTs, que possibilitam conhecer a distribuição, magnitude e tendência dessas doenças e de seus fatores de risco na população, identificando seus condicionantes sociais, econômicos e ambientais, com o objetivo de subsidiar o planejamento, execução e avaliação da prevenção e controle das mesmas. O resultado desta estimativa permite aperfeiçoar os algoritmos preditivos, objetivando promover a qualidade de vida e prevenir e controlar as DCNT.

Conclui-se que a partir do gerenciamento adequado da informação, é possível construir um sistema para extração de conhecimento e apoio a tomada de decisão, melhorando a prontidão da saúde. E, que somente por meio de abordagens abrangentes e sustentáveis poder-se-á reduzir as desigualdades sociais e promover uma sociedade mais saudável e justa.