

**PROMOÇÃO DA SAÚDE PARA IDOSOS: PROPOSTA DE SEGMENTAÇÃO
EPIDEMIOLOGICA**

**PROMOCIÓN DE LA SALUD PARA MAYORES: PROPUESTA SEGMENTACIÓN
EPIDEMIOLÓGICA**

**HEALTH PROMOTION FOR ELDERLY: A PROPOSAL EPIDEMIOLOGICAL
SEGMENTATION**

Frederico Molina Cohrs*

fcohrrs@unifesp.br

<https://orcid.org/0000-0002-6556-6852>

Alessandra Paula Ferreira Moreira Neumann*

neumann.alessandra@unifesp.br

<https://orcid.org/0000-0002-6182-9598>

Maria Aparecida Ferreira de Mello**

mariademello@technocare.net.br

<https://orcid.org/0000-0001-9542-3134>

Ivan Torres Pisa*

ivanpisa@gmail.com

<https://orcid.org/0000-0002-5106-3904>

Luiz Roberto Ramos*

lrr@uol.com.br

<https://orcid.org/0000-0003-3143-8315>

* Universidade Federal de São Paulo (Unifesp), SP, Brasil.

** Technocare, soluções e tecnologia Assistiva

Resumo

A promoção da saúde para qualquer público requer conhecimento acerca de suas características, possibilitando uma efetividade na comunicação. A população de idosos, crescente no Brasil, não é igual. Há necessidade de identificar subgrupos, para que a promoção da saúde a estes seja mais bem planejada. Para identificá-los, fez-se o uso da estratégia de segmentação. Este artigo apresenta uma proposta de segmentação epidemiológica: uma estratégia com uso de aprendizagem de máquina não supervisionado, que une dados demográficos, psicográficos e nosológicos. A proposta de segmentar e identificar os alvos para planejamento de ações é de suma importância na mudança de comportamento, e certamente contribuirá para uma população de idosos com mais autonomia e independência.

PALAVRAS CHAVE: Segmentação. Idoso. Promoção da saúde. Epidemiologia.

Resumen

La promoción de la salud para cualquier público requiere conocer sus características para que la comunicación sea más efectiva. La población de ancianos, que crece en Brasil, no es la misma. Hay subgrupos que deben identificarse para planificar mejor la promoción de la salud para esos subgrupos. Para identificar estos subgrupos se utiliza la estrategia de segmentación. Este artículo presenta la segmentación epidemiológica: una estrategia con el uso de aprendizaje automático no supervisado, que reúne datos demográficos, psicográficos y nosológicos. La propuesta de segmentar e identificar objetivos para la planificación de acciones es de suma importancia en el cambio de comportamiento, y ciertamente contribuirá a una población anciana con más autonomía e independência.

PALABRAS CLAVE: Segmentación. Anciano. Promoción de la salud. Epidemiología.

Abstract

Health promotion for any audience requires knowing its characteristics so that communication is more effective. The population of elderly people, growing in Brazil, is not the same. There are subgroups that need to be identified so that health promotion for those subgroups can be better planned. To identify those subgroups, the segmentation strategy is used. This article presents epidemiological segmentation: a strategy with the using unsupervised machine learning, bringing together demographic, psychographic, and nosological data. The proposal to segment and identify targets for action planning is of paramount importance in changing behavior and will certainly contribute to an elderly population with more autonomy and independence.

KEYWORDS: Segmentation. Elderly. Health promotion. Epidemiology.

1. Introdução

Um dos maiores desafios na promoção do envelhecimento saudável, ou com menores perdas funcionais possíveis, é a promoção da saúde visando estimular a alteração de hábitos e comportamentos da população idosa. Nas últimas décadas tem ocorrido um aumento significativo das pesquisas que aplicam a economia comportamental em saúde. O grande desafio é “construir infraestruturas institucionais e culturais que apoiem as pessoas idosas e as movam naturalmente para levar vidas saudáveis” (BEHAVIORAL, 2018). Além das diversas teorias sobre mudança de comportamento, a utilização dos princípios da economia comportamental tem se mostrado adequada para o estímulo de mudança do sistema. Esta tem permitido escalabilidade, além de ter possibilitado alcançar os grandes grupos populacionais, viabilizando melhorias da saúde ao longo do processo de envelhecimento.

Entretanto, a alteração de hábitos e comportamentos de um grupo populacional passa por uma correta e precisa comunicação interna. Cria-se, então, necessidade de estratégias específicas, para que o mesmo seja alcançado efetivamente. A fim de realizar uma comunicação eficaz, com planejamento adequado de estímulos visando mudanças de hábito e comportamento, é necessário compreender a fundo as características das pessoas idosas. Estas podem apresentar "crenças limitantes" que as impeçam de alterar seus hábitos. Para tal, é preciso estimar como as pessoas fazem escolhas subjetivas com resultados a longo prazo, além de compreender quais serão as diferenças dos resultados probabilísticos dos valores reais. As decisões diárias das pessoas são geralmente mais complicadas do que as estudadas em laboratórios (BEHAVIORAL, 2018).

Promover a saúde de forma geral, para indivíduos e comunidade, traz, sem dúvida, resultados muito satisfatórios. Todavia, ao proporcionar a saúde de grupos específicos os resultados podem ser ainda

melhores. Para tal, faz-se necessário focar na comunicação e na promoção da saúde a determinados segmentos e/ou subgrupos. A segmentação permite que os gestores transmitam ideias e mensagens com precisão. Segmentar implica buscar eficiência e otimizar comunicação e resultados de ações. Assim, conhecer bem as necessidades e comportamentos de um grupo auxilia no melhor planejamento e execução da promoção (CRESCITELLI; SHIMP, 2012).

Desta forma, a busca pela segmentação epidemiológica pode ser uma solução para a identificação dos subgrupos (segmentos) de idosos, fornecendo elementos que permitam o planejamento e execução de comunicações mais efetivas. Assim como Políticas Públicas que promovam o envelhecimento saudável, evitando os gastos que o envelhecimento com comprometimento funcional acarreta.

Brusilovsky (1999) define ainda que o segmento é um grupo de objetos similares (casos, pontos, observações, exemplos, membros, consumidores, pacientes, locais), enquanto a análise de *cluster* é um conjunto de técnicas dirigidas para o particionamento de dados, objetivando agrupar uma coleção de dados em *clusters*. Essas técnicas permitem identificar um número de *clusters* e formatos desconhecidos inicialmente; um grau de associação ou similaridade que seja forte entre os membros do mesmo segmento, e fraco entre os membros de cluster diferentes.

A segmentação não é, com isso, um simples agrupamento, mas sim um processo interpretável de objetos sob estudo. Pode-se, então, afirmar que uma classificação supervisionada não é considerada análise de *cluster*, assim como categorizações simples (segmentação de médicos por especialidade, assumindo que cada especialidade médica é conhecida; segmentação por sexo, nível educacional, taxa de respostas conhecidas) ou ainda resultado de uma consulta a uma base de dados, que por sua vez também não é um exemplo de segmentação.

Desse modo, o presente ensaio busca apresentar propostas de segmentação epidemiológica para a promoção da saúde de idosos.

2. Aspectos da promoção da saúde para idosos

A promoção da saúde visa capacitar indivíduos e comunidades a adotar comportamentos saudáveis e a fazer mudanças que reduzam o risco de desenvolver doenças crônicas. Para tal, existe o desenvolvimento de programas, com uma ampla gama de intervenções sociais e ambientais, destinados a promover e proteger a saúde e a qualidade de vida de indivíduos. Através do trabalho em relação ao aprendizado e à experiência, eles são projetados para ajudar indivíduos a melhorar sua saúde, expandindo seus conhecimentos e influenciando suas atitudes em uma abordagem que envolve diferentes atores e que foca em uma abordagem multissetorial (WHO, 1998).

A Conferência de Ottawa em 1986 levantou a ideia de permitir que indivíduos e comunidades aumentassem o controle sobre os determinantes da saúde, em busca da qualidade de vida. Com isso, a busca pela saúde da população deve adotar uma abordagem de cima para baixo, na qual os governantes desenvolvam estratégias para melhorar a saúde. Isso reforça a abordagem holística da promoção da saúde, que inclui a educação nessa área e vários serviços de suporte que permitam às pessoas melhorar sua saúde física e mental, bem como seu bem-estar social e econômico (WHO, 1986).

Para a Organização Mundial da Saúde – OMS (1986), há três elementos para a promoção da saúde: boa governança para a saúde, literacia em saúde, e cidades saudáveis. A boa governança direciona para uma linha central de políticas governamentais, que levam em consideração as implicações presentes em todas as decisões a serem tomadas, dando prioridade a políticas que diminuam as chances de adoecimento das pessoas.

A OMS afirma que é através de legislação que emerge o apoio à urbanização saudável, pois ao criar cidades que estimulam a caminhada, reduz-se a poluição do ar e das águas. Quanto à literacia, esta mostra que as pessoas precisam adquirir conhecimentos, habilidades e informações para fazer escolhas saudáveis. Já as cidades possuem um papel fundamental a desempenhar na promoção da saúde, e isso ocorre através do planejamento urbano e da construção de medidas preventivas nas comunidades (WHO, 1998).

A Conferência em Nairóbi, em 2009, recorreu a uma ação comunitária para fortalecer a saúde, melhorar os processos participativos, desenvolver e aplicar conhecimento na promoção da saúde. Estimula-se, com isso, a construção de políticas públicas sólidas, a criação de um ambiente favorável à saúde, o fortalecimento da liderança e da força de trabalho e o desenvolvimento de habilidades pessoais (WHO, 2009).

No Brasil, em 2003, um avanço para garantias ao idoso foi a publicação do Estatuto do Idoso (BRASIL, 2003). E em 2006, o Ministério da Saúde publicou a portaria 2.528, que dispõe sobre a Política Nacional de Saúde da Pessoa Idosa (BRASIL, 2006). Em suas diretrizes, a Política Nacional de Saúde da Pessoa Idosa busca, entre outros: a) promoção do envelhecimento ativo e saudável; b) atenção integral e integrada à saúde; c) estímulo às ações intersetoriais para a integralidade da atenção; d) indicação de recursos para assegurar a qualidade da atenção à saúde da pessoa idosa; e) promoção de cooperação nacional e internacional das experiências na atenção à saúde da pessoa idosa; f) apoio ao desenvolvimento de estudos e pesquisas.

Apesar das políticas publicadas pelo governo brasileiro, ainda havia necessidade de ampliação dos cuidados com os idosos. Assim, foi organizada a Rede de Atenção à Saúde no âmbito do Sistema Único de Saúde (SUS), na Portaria nº 4.279/2010 (BRASIL, 2010), e planejadas ações para promover mais estímulo de promoção da saúde (estratégia 12 no eixo II) no Plano de Ações Estratégicas para o Enfrentamento das Doenças Crônicas Não Transmissíveis (DCNT) no Brasil (2011-2022) (BRASIL, 2011).

O Planejamento Estratégico do Ministério da Saúde, para os anos de 2011 a 2015, procurou garantir atenção integral à saúde da pessoa idosa. Em 2013, o governo publicou o Decreto nº 8.114 que estabeleceu o Compromisso Nacional para o Envelhecimento Ativo. Esse decreto foi revogado e substituído pelo Decreto nº 9.921/2019 (BRASIL, 2019), solidificando uma Política Nacional da Pessoa Idosa, criada inicialmente em 2006.

As políticas públicas para promoção da saúde são construídas com visão municipal, estadual e, eventualmente, nacional. A construção dessas políticas, em geral, é o resultado de discussões com especialistas de diversas áreas, tais como: geriatras, enfermeiros, gerontólogos, educadores físicos, fisioterapeutas, psicólogos, fonoaudiólogos e outros profissionais. Para a população de idosos, o planejamento de promoção da saúde deve focar na independência e autonomia do idoso.

Outro aspecto não visível na Política Nacional de Saúde da Pessoa Idosa, normatizada pela Portaria GM/MS nº 2.528, de 19 de outubro de 2006 (BRASIL, 2006) e suas decorrentes ações para sua implementação, é a diferenciação entre envelhecimento físico e fragilidade emocional. Estudos têm demonstrado que os adultos mais velhos tendem a ter uma perspectiva mais positiva do que os adultos mais jovens. Isso foi recentemente testado numa condição extrema do mundo real. Carstensen foi coautor de um estudo em abril de 2020 — precisamente quando a pandemia COVID-19 tinha começado a elevar praticamente todos os aspectos das nossas vidas — que descobriu que os idosos eram mais resistentes do que os jovens. Num inquérito de quase 1.000 adultos com idade entre 18 e 76 anos, os idosos estavam bem cientes de que tinham um maior risco para a saúde em comparação aos adultos mais jovens, e todos tinham um estresse financeiro comparável. Entretanto, os participantes mais velhos relataram tanto emoções mais

positivas quanto menos negativas (CARSTENSEN; HERSHFIELD, 2021). Esse conhecimento poderia permitir uma melhor comunicação a respeito dos riscos para a saúde, com ações de promoção.

Promover a saúde traz resultados melhores quando se conhece o público-alvo. Quer seja adulto ou idoso, é necessário conhecer o modo como o coletivo pensa sobre sua realidade e planeja seu futuro. Carstensen e Hershfield (2021) apontam que os idosos tendem a se perceber mais jovens do que realmente são. Os idosos sabem que dispõem "menos anos em seu futuro", no entanto, os que possuem 50 anos assumem se sentir cerca de cinco anos mais jovens (um desconto de 10%) e os de 70 anos tendem a dizer que se sentem na faixa dos 56-60 anos (um desconto de 15%-20%). A partir disso, o questionamento feito é de como pode o Brasil planejar todas as suas ações de promoção da saúde da pessoa idosa sem considerar a heterogeneidade desse grupo (BRASIL, 2006), somada aos diferentes contextos do nosso território nacional. Ademais, também é indagado o risco que corre o país ao promover a alteração da percepção do processo de envelhecimento que têm os profissionais que prestarão os serviços aos idosos.

3. Segmentação epidemiológica para profissionais da saúde

Mais importante do que identificar quais são as ações necessárias para a promoção do envelhecimento saudável, que garante à população um seguimento de sua produção de riqueza para a sociedade, é promover as mudanças necessárias a fim de que a população adote efetivamente hábitos para uma vida saudável. Para conduzir essas mudanças, deve-se fazer uso dos conceitos de marketing social (CHICHIREZ, 2018), que “visa abordar os problemas sociais, em vez de produzir ganhos financeiros”. Para Smith (1956), uma das características do marketing social é ser uma estratégia de posicionamento, na qual deve-se deixar claro para o cidadão qual é a vantagem de um determinado comportamento em relação a outros. Esse posicionamento se dá por meio de uma comunicação bem-feita e bem direcionada. Uma complexidade atual são os esforços de comunicação que muitas vezes não conseguem se conectar com a mentalidade das pessoas idosas, e não reconhecem que há vários grupos de idosos. Isto é, as pessoas idosas não podem ser interpretadas como um grupo único (CARSTENSEN; HERSHFIELD, 2021). Comunicar-se eficazmente com a população idosa é um desafio. A heterogeneidade do processo de envelhecimento torna a idade cronológica um preditor cada vez mais pobre de comportamento social e de saúde, este último representado pela funcionalidade cognitiva e física (LOWSKY et al., 2014). Faz-se necessário considerar também as diferenças perceptuais relacionadas aos níveis de educação, bem-estar financeiro e conhecimentos especializados (DANNEFER et al., 2019).

O setor privado consegue identificar as diferenças entre pequenas e grandes populações através da criação de segmentos. Segmentos são, pois, grupos de indivíduos com características semelhantes entre si, e muito diferentes dos outros grupos. A segmentação é a ação de agrupar dados de uma população em segmentos ou subgrupos semelhantes, auxiliando a lidar melhor com a diversidade de características dos indivíduos (SMITH, 1956).

Um projeto com marketing social foi apresentado por GRIER e BRYANT (2005), no qual foi identificado sucesso na diminuição de acidentes automobilísticos por parte de motoristas alcoolizados. O projeto *Road Crew* (tripulação da estrada) identificou que os motoristas, mesmo se vendo embriagados, recusavam-se a deixar seus veículos estacionados próximos aos bares para pegar uma carona para casa. Assim, fazendo uso de tentativa de mudança de comportamento, o projeto oferecia levá-los aos bares que desejassem, de um bar a outro, e depois para casa. Apesar das críticas de que esse projeto poderia estimular um aumento de consumo de álcool, foi identificada, ao final do primeiro ano, uma redução de 17% nos acidentes causados por motoristas alcoolizados, trazendo uma economia de US\$ 610.000,00 ao Estado de Wisconsin; além de não ter sido identificado um aumento do consumo de álcool.

O uso de segmentação de dados tem o exemplo de SHIN e MAUPOME (2017), os quais usaram

esta estratégia, de identificar grupos diferentes, para o acesso à informação sobre saúde por parte de imigrantes mexicanos nos Estados Unidos. Havia grupos que preferiam a informação em espanhol, enquanto outros preferiam a informação em inglês. Esse conhecimento auxilia na melhor construção de estratégias de comunicação e informação sobre saúde.

Um experimento de segmentação também foi realizado por COHRS et al. (2013) que identificaram três grupos de idosos em um estudo longitudinal. Em linhas gerais, o primeiro segmento era composto, predominantemente, por mulheres com alguma dependência física, algum comprometimento cognitivo, e com doenças interferindo nas atividades diárias. O segundo segmento constituído, majoritariamente, por mulheres também, viúvas, fisicamente independentes, convivendo com poucos familiares, sobrevivendo economicamente de aposentadoria ou pensão e mentalmente capazes. O terceiro segmento, predominando o sexo masculino, integrado por indivíduos fisicamente independentes, casados, com interesse em manter prática sexual, com baixo número de comorbidades, vivendo apenas com o dinheiro da aposentadoria e apresentando uma boa capacidade cognitiva. Três segmentos com características muito diferentes entre si, mas que propiciam, aos planejadores de políticas públicas de saúde, a possibilidade de direcionar maiores cuidados aos indivíduos do primeiro segmento, cumprindo assim um dos princípios do SUS: equidade, integralidade e universalidade. Berkowitz (2011) aponta vários benefícios para realizar a segmentação: 1) auxilia no direcionamento de estratégias de mercado, aumentando a probabilidade de que uma ideia seja aceita; 2) auxilia na identificação de concentrações homogêneas de indivíduos para os quais estratégias de promoção da saúde possam ser melhor desenhadas; 3) auxilia na identificação do melhor canal de distribuição para um serviço em particular; 4) mensagens direcionadas a um segmento específico são geralmente mais efetivas; 5) focar recursos em um segmento pode trazer economia de escala.

Há várias estratégias de segmentação descritas, sendo que as mais usadas são: demográfica, geográfica, psicográfica, comportamental, e de coorte (BERKOWITZ, 2011; CRESCITELLI, 2013).

A segmentação demográfica usa variáveis como idade, sexo ou raça. Há um aumento de organizações de saúde que fazem segmentação por renda. A segmentação geográfica identifica desejos diversos em diferentes áreas do país. Moradores do Norte do Brasil possuem características diferentes das dos moradores no Sul. Para o setor de saúde, o foco está em distribuição de doenças, como descrevem Souza et al. (SOUZA et al. 2020) ao relatarem diferenças críticas de doenças infecto-parasitárias entre as várias regiões do Brasil.

As diferenças geográficas observáveis na área da saúde podem estar no próprio município, como aponta Ribeiro e Silva (2016) sobre as internações por causas respiratórias no município de São Paulo.

O estilo de vida é uma informação importante na segmentação psicográfica, assim como a classe social, a escolha de fontes de informação e a influência por pares. Estes já são dados mais complexos coletados na área da saúde. Estudos como o de Fett et al. (2009), sobre mudança de estilo de vida em mulheres sedentárias, é um exemplo do uso dessas variáveis. Já o de Mercken et al. (2012), sobre influências em uso de fumígenos entre adolescentes, no qual observaram que adolescentes selecionam amigos com base em comportamento semelhante ao de fumar, indica que o fator de comportamento de fumar é mais significativo que a influência de amigos.

A estratégia de segmentação por comportamento visa criar subgrupos de pessoas com uso semelhante de produtos ou serviços, com fidelidades parecidas, e que buscam benefícios equivalentes. Programas de fidelização que oferecem benefícios a seus usuários estão cada vez mais presentes em empresas de saúde. Liu et al. (2014) mostram que ao apresentar um indivíduo a programas de recompensas, ocorre um encorajamento à inscrição em plataformas de prevenção da saúde usando a internet.

A segmentação por coorte tem como base, segundo Motta et al. (1999, apud Feitosa e Ikeda, 2011),

o agrupamento de pessoas com características e histórias de vida semelhantes, como as pessoas que vivenciaram o ataque das torres gêmeas em Nova York em 2001. Semelhantemente, é possível identificar no texto de Feitosa e Ikeda (2011) que o conceito de coorte usado pelos profissionais de marketing diverge do conceito utilizado pela área da saúde. Neste último, uma coorte é um estudo longitudinal.

Das estratégias de segmentação descritas, nenhuma contempla um dado muito importante na saúde, quer seja em estudos populacionais ou em amostras, que é o lidar com dados nosológicos.

Assim, os autores apresentam aqui uma estratégia de segmentação, unindo os conceitos da segmentação psicográfica, geográfica, demográfica e dados nosológicos, formando assim a segmentação epidemiológica.

3.1. Inteligência artificial para segmentação epidemiológica

Para realizar as estratégias de segmentação, há técnicas de inteligência artificial em mineração de dados. Dentre estas, existe o aprendizado de máquina não supervisionado.

O aprendizado de máquina é um conjunto de métodos usados para detectar padrões em dados. Esses padrões auxiliam na predição de dados futuros e na tomada de decisões. Murphy (2012) descreve a existência de dois tipos de aprendizado de máquina: supervisionada ou preditiva e o não supervisionado. O primeiro aprendizado se faz por meio do uso de um conjunto de dados com desfecho previamente conhecido que será usado para treinamento. Após o treinamento, o algoritmo consegue prever um desfecho com base em novos dados apresentados a um modelo construído.

O segundo tipo de aprendizado não depende de um conjunto de dados com desfecho previamente conhecido. O que os algoritmos fazem é buscar um padrão ou agrupar conjuntos de dados semelhantes (SIDEY-GIBBONS; SIDEY-GIBBONS 2019). Isso é chamado de análise de *cluster*.

Análise de *Cluster* ou “clusterização” é uma técnica de mineração de dados, um método não supervisionado de classificação de observações em grupos menores que o conjunto original, denominados *clusters* ou segmentos (ROIGER; GEATZ 2003). Um *cluster* é uma coleção de dados semelhantes, mas diferentes dos pertencentes aos outros *clusters* (HAN; KAMBER, 2006). Problemas de “clusterização” têm sido abordados em diversos contextos e por pesquisadores em muitas disciplinas, o que reflete o seu grande apelo e utilidade como uma das etapas na análise exploratória de dados, apresentando-se como um problema difícil combinatoriamente, com diferentes pressupostos (JAIN; MURTY; FLYNN, 1999).

Han e Kamber citam a segmentação como sendo também conhecida como “segmentação de dados”, uma vez que a “clusterização” particiona grandes conjuntos de dados em grupos de acordo com suas similaridades (HAN; KAMBER, 2006).

Na realização de uma “clusterização” em mineração de dados, algumas condições são comuns (HAN; KAMBER, 2006):

- a. escalabilidade dos dados: alguns algoritmos de segmentação funcionam melhor com pequena quantidade de dados, outros, com grande massa de dados. Ainda há algoritmos que podem ser usados com pequena ou grande quantidade de dados;
- b. habilidade de lidar com diferentes tipos de atributos: muitas análises podem requerer uso de variáveis binárias ou categóricas, além das numéricas;
- c. possibilidade de *clusters* com formas arbitrárias: alguns algoritmos, por usar medidas euclidianas ou de Manhattan, tendem a encontrar *clusters* com formato esférico. Porém, é importante que o algoritmo a ser usado possa identificar *clusters* com formas arbitrárias;

- d. alta dimensionalidade: alguns algoritmos lidam bem com baixa dimensionalidade, como o olho humano que lida com até três dimensões. No entanto, é importante que o algoritmo possa lidar com várias dimensões;
- e. habilidade para lidar com dados ruidosos: bases de dados do mundo real possuem valores extremos, valores ausentes, dados inseridos erroneamente. Alguns algoritmos são sensíveis a tais dados e podem conduzir a uma segmentação sem qualidade;
- f. segmentação incremental e nenhuma sensibilidade com a ordem de entrada dos dados: o algoritmo deve permitir novas inserções de dados e não ser sensível à ordem de entrada;
- g. pouco domínio da área de conhecimento para determinar os parâmetros de entrada: quanto mais conhecimento da área de estudo for necessário, maior a dificuldade de controle na segmentação, especialmente em bases com alta dimensionalidade;
- h. segmentação com base em parametrizações; usabilidade e possibilidade de interpretação: os resultados devem estar intimamente ligados a uma semântica específica para facilidade de interpretação.

A “clusterização” pode ser julgada na interpretação de utilidade dos resultados. Um ponto importante a ser ressaltado é que não há uma medida de sucesso, já que são usados argumentos heurísticos para o julgamento (BRUVILOVSKY, 2010).

Há algoritmos para “clusterização” disponíveis, cada qual indicado para um grupo de investigações, problemas específicos, podendo lidar com tipos distintos de variáveis.

Os métodos de “clusterização” mais conhecidos podem ser classificados nas seguintes categorias: a) hierárquicos; b) particionamento; c) baseados em densidade; d) baseados em grade; e) baseados em modelos; f) segmentação de dados ultra dimensionais; g) segmentação baseada em parâmetros e restrições (HAN; KAMBER, 2006)

O método mais comumente utilizado é o hierárquico (GELBARD; GOLDMAN; SPIEGLER, 2007), que constrói (aglomera) ou quebra (divide) uma hierarquia de *clusters*. A representação tradicional dessa hierarquia é uma árvore, chamada de dendrograma, com elementos individuais em uma extremidade e um único *cluster* contendo todos os elementos em outra extremidade.

Um estudo comparativo realizado por Gelbard et al. (2007) aplicando 11 algoritmos de “clusterização” em quatro bases de dados públicas, aponta os métodos não hierárquicos como desempenho superior na “clusterização”, sendo os mais bem pontuados os métodos: *TwoStep* e *K-means*. Os resultados deste estudo sugerem um cuidado no tipo de conjunto de dados para a escolha do método de “clusterização”, uma vez que os resultados dificilmente são dados dependentes do domínio, mas sim do algoritmo e dado dependente (GELBARD; GOLDMAN; SPIEGLER, 2007).

Para a realização de segmentação epidemiológica, optou-se pela utilização de uma “clusterização” do tipo particionamento, cujos algoritmos disponíveis mais conhecidos são *K-means*, *K-medoides* e *TwoStep*.

O algoritmo *K-means* (VELMURUGAN, 2010) orienta cada ponto a um cluster cujo centro (também chamado de centroide) esteja mais próximo. O centro é a média de todos os pontos no *cluster*, ou seja, suas coordenadas são a média aritmética para cada dimensão separadamente sobre todos os seus pontos.

A fragilidade do *K-means* (VELMURUGAN, 2010) é o fato de lidar com a média. Se uma determinada variável tiver valores discrepantes, a média não é a melhor representação do grupo. Assim, pode-se então usar uma técnica baseada em objeto representativo, que é o *K-medoide* (VELMURUGAN,

2010). Ao invés de selecionar o valor médio do objeto no *cluster* como ponto de referência, um objeto é selecionado para representar o *cluster*. Cada objeto restante é agrupado com o objeto representativo com o qual é mais similar. O método de particionamento é então realizado com base no princípio de minimização da soma das diferenças das distâncias entre cada objeto e seus pontos de referência correspondentes.

Um problema na utilização dos métodos acima descritos recai na necessidade prévia de informar ao algoritmo a quantidade de *clusters* que se deseja. Com vistas a isso, optou-se pela utilização do método descrito a seguir, *TwoStep*.

O método *TwoStep* (THEODORIDIS; KOUTROUMBAS, 2006) é uma análise escalonável, apropriado para grandes conjuntos de dados, que possuam uma mistura de variáveis contínuas e categóricas. Seu funcionamento, cujo nome indica, é baseado em dois passos. Na primeira etapa, o algoritmo forma os *pré-clusters*, que são *clusters* dos casos originais, usados no local dos dados brutos. Estes são utilizados com objetivo de ter um menor número de casos para a próxima etapa e de reduzir o tamanho da matriz que contém distâncias entre todos os casos pareados possíveis. Isso ocorre porque é usada uma abordagem de agrupamento sequencial. O algoritmo faz uma varredura dos registros um a um e decide se o registro atual deve fundir-se com os *clusters* formados anteriormente ou iniciar um novo *cluster* com base no critério da distância (SPSS Inc, 2001).

Após a finalização da pré-segmentação, todos os casos em um mesmo *pré-cluster* são tratados como uma única entidade através de uma técnica baseada em hierarquização. As suas características são utilizadas como novos casos. O tamanho da matriz de distâncias não é mais dependente do número de casos, mas do número de *pré-clusters* (BACHER; WENZIG; VOGLER, 2004). Duas medidas de distância estão disponíveis, euclidiana e log-verossimilhança. A log-verossimilhança pode lidar com tipos mistos de atributos, sendo a vantagem deste algoritmo. A medição da dispersão das variáveis categóricas é feita por uma entropia. Igualmente ao que acontece na segmentação hierárquica, os *clusters* com as menores distâncias são unidos em cada passo (BACHER; WENZIG; VOGLER, 2004).

A vantagem do método *TwoStep* para segmentação de dados é que não se faz necessário informar a quantidade de *clusters* final. Caso este parâmetro não seja fornecido, o algoritmo encontra a quantidade ótima. Entende-se como ótima a melhor quantidade dentre diversas testadas. Os subgrupos formados podem ainda passar por uma engenharia de características para que possam ser nominados.

Em síntese, a segmentação epidemiológica deve fazer uso de dados que possam identificar o local de moradia e/ou trabalho, dados demográficos, dados psicográficos e dados nosológicos. A partir dessas informações e por meio de análise de *cluster*, identifica-se segmentos com indivíduos que apresentem características semelhantes entre si e que sejam diferentes dos sujeitos pertencentes a outros segmentos.

Aos termos segmentos bem definidos, a comunicação é facilitada por melhor compreender quem irá decodificar a mensagem encaminhada, resultando em melhores planejamentos em promoção da saúde.

4. Considerações finais

Promover a saúde do idoso pode apresentar melhores resultados se os órgãos governamentais não trabalharem de forma genérica. A razão se dá pelo fato de as pessoas possuírem condições clínicas distintas, assim como diferentes níveis de funcionalidade, compreensão do estado de saúde, qualidade de vida, educação financeira, além de diferentes relações familiares e redes de contato e atenção. Apesar de comunicação e ações de promoção da saúde sem diferenciação apresentarem preços mais atrativos, o

desejo principal é que o maior número de pessoas compreenda a necessidade do autogerenciamento da saúde.

A indústria consegue identificar as diferenças entre as populações por meio da criação de segmentos. Segmentos são grupos de indivíduos com características semelhantes entre si, e muito diferentes dos outros grupos. A segmentação é a ação de segmentar dados de uma população para achar os segmentos ou subgrupos, auxiliando a lidar melhor com a diversidade de características dos indivíduos.

O planejamento e as ações para a promoção da saúde para a pessoa idosa têm muito a ganhar ao usar a segmentação epidemiológica. São temas de importância na preservação da saúde da pessoa idosa planejar promoções para cessação de uso de fumígenos, uso abusivo de bebidas alcoólicas, uso correto de medicação, prevenção de lesões e quedas, busca por melhor qualidade de sono, cuidado com a saúde mental e preservação de processos cognitivos.

Foi constatado que em documentos oficiais brasileiros, que tratam de assuntos relacionados à promoção da saúde, não há referência nenhuma à figura do idoso no Programa Nacional de Promoção da Saúde. Existe apenas menção à citação sobre promoção da saúde no Estatuto do Idoso, sem sequer considerar as categorias de idosos contidas naquele documento.

Ressalta-se que a promoção da saúde é baseada em uma visão holística em relação à mesma. A promoção da saúde tem uma abordagem participativa, com foco não apenas na saúde individual, mas também no bem-estar de toda uma comunidade, e nas políticas públicas pensadas para ela. Portanto, indo além de estilos de vida saudáveis, não é responsabilidade exclusiva do setor da saúde. Este conceito está integrado na ação social, na qual os indivíduos são capacitados por medidas sociais para assumir o controle de sua própria saúde, incentivando cada sujeito a abraçar a ideia de bem-estar, aumentando assim o controle sobre como vive cotidianamente.

Contribuições dos autores: Frederico Molina Cohrs participou de todas as etapas do estudo, desde a concepção, análise e interpretação, redação do conteúdo intelectual e aprovação da versão final. Alessandra Paula Ferreira Moreira Neumann participou da redação do conteúdo intelectual, revisão crítica e aprovação da versão final. Maria Aparecida Ferreira de Mello participou da redação do conteúdo intelectual, revisão crítica e aprovação da versão final. Ivan Torres Pisa coorientou e participou de todas as etapas do estudo, desde a concepção até a aprovação da versão final. Luiz Roberto Ramos orientou e participou de todas as etapas do estudo, desde a concepção até a aprovação da versão final.

REFERÊNCIAS

BACHER, J.; WENZIG, K.; VOGLER, M. SPSS TwoStep Cluster - A First Evaluation Friedrich-Alexander-Universität Erlangen-Nürnberg, 2004. Disponível em: <http://www.opus.ub.uni-erlangen.de/opus/volltexte/2004/81/pdf/a_04-02.pdf>. Acesso em: 17 fev. 2020.

BEHAVIORAL, Board et al. Behavioral Economics and the Promotion of Health Among Aging Populations. National Academies of Sciences, Engineering, and Medicine. **Proceedings of a Workshop— in Brief**. Washington, DC: The National Academies Press, 2018. doi: <https://doi.org/10.17226/25190>

BERKOWITZ E. **Market segmentation**. In: Essentials of health care marketing. 3rd ed. Sudbury, MA: Jones & Bartlett Learning; 2011. p. 195–231.

BRASIL, Presidência da República. **Lei 13.466** de 12 de julho de 2017 [Internet]. 13466 jul 12, 2017. Disponível em: <http://www.planalto.gov.br/ccivil_03/_Ato2015-2018/2017/Lei/L13466.htm> Acesso em: 17 fev. 2020.

_____, Presidência da República. **Lei 10.741** de 01 de outubro de 2003 [Internet]. 2003 [citado 19 de junho de 2020]. Disponível em: <http://www.planalto.gov.br/ccivil_03/leis/2003/L10.741.htm>. Acesso em: 16 abr. 2021.

_____, Ministério da Saúde. **Portaria 2.528** de 19 de outubro de 2006 [Internet]. 2006 [citado 19 de junho de 2020]. Disponível em: <http://bvsms.saude.gov.br/bvs/saudelegis/gm/2006/prt2528_19_10_2006.html>. Acesso em: 16 abr. 2021.

_____, Ministério da Saúde. Secretaria de Políticas de Saúde. **Portaria nº 4.279**, de 30 de dezembro de 2010. Disponível em: <https://bvsms.saude.gov.br/bvs/saudelegis/gm/2010/prt4279_30_12_2010.html> Acesso em: 16 abr. 2021.

_____, **Plano de ações**. Estratégias para o enfrentamento das doenças crônicas não transmissíveis (DCNT) no Brasil: 2011-2022. 1ª edição. Brasília, DF: Ministério da Saúde; 2011. 154 p. (Série B. Textos básicos de saúde).

_____, Presidência da República. **Decreto nº 9.921**, de 18 de julho de 2019 [Internet]. 2019 [citado 4 de outubro de 2020]. Disponível em: <http://www.planalto.gov.br/ccivil_03/_ato2019-2022/2019/decreto/D9921.htm>. Acesso em: 18 abr. 2021.

BRUSILOVSKY, P. **Cluster Analysis vs. Market Segmentation** [Internet]. 2010 [citado 27 de agosto de 2021]. Disponível em: <<http://www.bisolutions.us/Cluster-Analysis-vs.-Market-Segmentation.php>>. Acesso em: 27 abr. 2021.

CARSTENSEN, L. L. ; HERSHFIELD, H. E. Além dos estereótipos: Usando a teoria da seletividade socioemocional para melhorar o envio de mensagens para adultos mais velhos. **Direções atuais na ciência psicológica**, v. 30, n. 4, p. 327-334, 2021. Disponível em: <review.ucla.edu/wp-content/uploads/2021/04/CarstensenHershfieldMessagingarticle.pdf> Acesso em: 21 fev. 2020.

CHICHIREZ, C. M.; PURCĂREA, V. L. Health marketing and behavioral change: a review of the literature. **Journal of medicine and life**, v. 11, n. 1, p. 15, 2018.

COHRS, F. M. et al. Aplicação de Análise de Cluster em dados integrados de um estudo prospectivo: projeto epidioso como cenário. **Journal of Health Informatics**, v. 5, n. 1, 29 mar. 2013.

CRESCITELLI, E.; SHIMP, T. A. Segmentação. In Comunicação de Marketing: integrando propaganda, promoção e outras formas de divulgação. São Paulo: **Cengage Learning**, p. 86-112, 2012.

DANNEFER, D. H.; C., KELLEY, J. Beyond the "haves" and "have nots". **Generations**, v. 42, n. 4, p. 42-49, 2019.

FEITOSA, W. R.; IKEDA A. A. Segmentação de Mercado com Base em Coortes: uma investigação qualitativa. São Paulo. **Revista brasileira de gestão de negócios**, v. 13, p. 359-375, 2011.

FETT, C. A.; FETT, W. C. R.; PADOVAN, G. J.; MARCHINI, J. S. Mudanças no estilo de vida e fatores de risco para doenças crônicas não transmissíveis e sistema imune de mulheres sedentárias. **Rev Nutr.** v. 22, n. 2, p. 245-255, abril de 2009.

- GELBARD, R.; GOLDMAN, O.; SPIEGLER I. Investigating diversity of clustering methods: An empirical comparison. **Data Knowledge Engineering**, v.63, n. 1, p. 155-166, 2007.
- GRIER, S.; BRYANT, C. A. SOCIAL MARKETING IN PUBLIC HEALTH. **Annual Review of Public Health**, v. 26, n. 1, p. 319–339, 21 abr. 2005.
- HAN, J.; KAMBER, M. **Cluster analysis**. In: Data mining: concepts and techniques. 2 ed San Francisco, California, USA: Morgan Kaufmann; 2006. p. 770.
- JAIN, A. K.; MURTY, M. N.; FLYNN, P. J. Data clustering: a review. **ACM Comput. Surv.**, v. 31, n. 3, p. 264–323, 1999.
- LIU S.; HODGSON C.; ZBIB A. M.; PAYNE, A. Y.; NOLAN, R. P. The Effectiveness of Loyalty Rewards to Promote the Use of an Internet-Based Heart Health Program. **Journal of Medical Internet Research**, v. 16, n. 7, p. e163, 2 de julho de 2014.
- LOWSKY, D. J., OLSHANSKY, S. J.; BHATTACHARYA, J.; GOLDMAN, D. P. Heterogeneity in healthy aging. **Journals of Gerontology**, Series A: Biological Sciences and Medical Sciences, v. 69, n. 6, p. 640-649, 2014. <https://doi.org/10.1093/gerona/glt162>
- MERCKEN L.; STEGLICH C.; SINCLAIR P.; HOLLIDAY J.; MOORE L. A longitudinal social network analysis of peer influence, peer selection, and smoking behavior among adolescents in British schools. **Health Psychology**, v. 31, n. 4, p. 450, julho de 2012.
- MURPHY, K. P. **Machine learning: a probabilistic perspective**. Cambridge, MA: MIT Press, 2012.
- RIBEIRO H.; SILVA E. N. Desigualdades intraurbanas em internações hospitalares por doenças respiratórias e circulatórias em uma área da cidade de São Paulo. **Cadernos Metrópole**, v. 18, p. 461-480, dez. 2016.
- ROIGER, R.; GEATZ, M. **Data Mining**. A tutorial-based primer. [s.l.] Addison Wesley, 2003.
- SHIN, Y.; MAUPOME, G. Segmentation of Mexican-Heritage Immigrants: Acculturation Typology and Language Preference in Health Information Seeking. **Journal of immigrant and minority health**, v. 19, n. 5, p. 1163–1173, out. 2017.
- SIDEY-GIBBONS J. A. M.; SIDEY-GIBBONS C. J. Machine learning in medicine: a practical introduction. **BMC Medical Research Methodology**, v. 19, n. 1, p. 1-18, dezembro de 2019. Disponível em: <<https://bmcmmedresmethodol.biomedcentral.com/articles/10.1186/s12874-019-0681-4>>. Acesso em: 24 mar. 2020.
- SMITH, W. R. Product differentiation and market segmentation as alternative marketing strategies. **Journal of marketing**, v. 21, n. 1, p. 3-8, 1956. Disponível em: <<https://journals.sagepub.com/doi/10.1177/002224295602100102>>. Acesso em: 24 mar. 2020.
- SOUZA H. P.; OLIVEIRA, W. T. G. H.; SANTOS, J. P. C.; TOLEDO, J. P.; FERREIRA, I. P. S.; SOUSA Esashika S. N. G.; DELÁCIO, A. Infectious and parasitic diseases in Brazil, 2010 to 2017: considerations for surveillance Enfermedades infecciosas y parasitarias en Brasil de 2010 a 2017: aspectos para la vigilancia sanitaria. **Revista Panamericana de Salud Publica**, v. 44, p. 10 ed., 2020.
- SPSS, INC. The SPSS **TwoStep Cluster Component**. A scalable component enabling more efficient customer segmentation. SPSS Inc, 2001.

THEODORIDIS, S.; KOUTROUMBAS, K. **Pattern Recognition**. 3. ed. [s.l.] Academic Press, 2006.

VELMURUGAN. Computational Complexity between K-Means and K-Medoids Clustering Algorithms for Normal and Uniform Distributions of Data Points. **Journal of Computer Science**, v. 6, n. 3, p. 363–368, 1 mar. 2010.

WHO, World Health Organization. **Carta de Ottawa**. 1986. Disponível em: <https://bvsmms.saude.gov.br/bvs/publicacoes/carta_ottawa.pdf> Acesso em: 28 mar. 2020.

_____, Department of Communications. **What is health promotion?** Disponível em: <<http://www.who.int/features/qa/health-promotion/en/>>. Acesso em: 5 dez. 2018.

_____, World Health Organization. **Health Promotion Glossary**, Geneva, 1998. Disponível em: <<https://www.who.int/healthpromotion/about/HPR%20Glossary%201998.pdf>> Acesso em: 28 mar. 2020.

Recebido em: 09/08/2021

Aceito em: 17/09/2021

Endereço para correspondência:

Frederico Molina Cohrs

fcohrrs@unifesp.br



Esta obra está licenciada sob uma [Licença Creative Commons Attribution 4.0](https://creativecommons.org/licenses/by/4.0/)