

Big data: aplicações e limites em saúde.

P. S. P. LACERDA

Mestre em Ciência e Tecnologia em Saúde, Universidade Mogi das Cruzes, Análise e desenvolvimento de Sistemas, Negócios, Centro Universitário Ítalo Brasileiro, UniÍtalo, São Paulo, SP, Brasil.

E-mail: paulo.lacerda@italo.edu.br.

COMO CITAR O ARTIGO:

P. S. P. LACERDA. **Big data: aplicações e limites em saúde..** URL: www.italo.br/portal/cepep/revistaeletronica.html. São Paulo SP, v.8, n. 3, p. 68-83, jul/2018

RESUMO

Na era contemporânea, a Internet é o principal recurso por busca de informações sobre uma vida saudável. No mundo todo, alguns estudos mostram que milhares de pessoas fazem uso de mecanismo de busca de internet para garimpar as informações sobre saúde. A saúde é uma preocupação tanto individual quanto coletiva. No campo da saúde, com avanço da tecnologia, principalmente a partir da década de 2000, novas ferramentas tecnológicas estão sendo usadas para melhorar a garimpagem, armazenamento e filtragem dessas informações, transformando-as em valores para o indivíduo ou para sociedade. O Big Data é uma dessas ferramentas. O Big Data é uma tecnologia que permite processar e analisar dados com base volume, velocidade e variedade dessas informações. Nova era de fonte de dados, novos desafios. O propósito principal desse trabalho é compreender as abordagens e limites da aplicação do Big Data na saúde pública por meio de uma revisão de literatura, além de identificar as áreas de aplicação bem como seus limites. A fim de atingir os objetivos estabelecidos, uma metodologia de revisão da literatura foi criada. Como resultado dessa metodologia, dez artigos foram encontrados e categorizados de acordo com afinidade de contexto em: a) aplicado à medicina social; b) aplicado a saúde pública; c) relacionado a conceitos gestão em saúde; d) relacionado à preocupação com a ética. Os resultados permitiram concluir que o Big Data é uma fonte de enorme potencial ainda a ser explorado.

Palavras-chave: big data, dados, saúde.

ABSTRACT

In the contemporary era, the Internet is the primary resource for seeking information about a healthy life. Worldwide, some studies show that thousands of people make use of the internet search engine to digest health information. Health is a concern both individually and collectively. In the field of health, with advanced technology, especially since the 2000s, new technological tools are being used to improve the mining, storage and filtering of this information, transforming them into values for the individual or for society. Big Data is one such tool. Big Data is a technology that allows you to process and analyze data based on volume, speed and variety of this information. New era of data source, new challenges. The main purpose of this paper is to understand the approaches and limits of the application of Big Data to public health through a literature review, as well as to identify the areas of application as well as their limits. In order to reach the established objectives, a literature review methodology was created. As a result of this methodology, ten articles were found and categorized according to context affinity in: a) applied to social medicine; b) applied to public health; c) related to health management concepts; d) related to the concern with ethics. The results allowed us to conclude that Big Data is a source of enormous potential yet to be explored.

Keywords: big data, data, health.

INTRODUÇÃO

Em busca da saúde saudável, pessoas usam a Internet como recurso para essas informações. Estudos mostram que mais de 90 milhões de americanos fazem uso dessa tecnologia por meio de mecanismo de busca. Estas consultas são registradas em arquivos de massa chamado de Big Data que são livremente acessíveis ao público (WILLIAMS; SMITH, 2015).

O termo Big Data é usado para descrever tecnologias que permitem o gerenciamento e a análise de dados com base no conceito conhecido com 3Vs: volume, velocidade e variedade (TAMBE, 2014). Um volume de dados que tradicionais sistemas de banco de dados não são capazes de suportar, analisar e transformá-los em valor (ARIMA, 2016).

A partir do início da década de 2000, há evolução crescente pela análise de dados e um grande interesse na potencialidade de utilização do Big Data em diversas áreas, inclusive na saúde pública (WAYENA *et al.*, 2018). Setores públicos e privados estão investindo nas tecnologias e nos recursos analíticos necessários para explorar por inteiro o valor Big Data.

A enorme contribuição do Big Data à saúde pública pode ser caracterizado pelo volume de informações obtido sobre ambiente ou contexto de vida (social, econômico, cultural). Potencial que pode ser aplicado em diversas áreas da saúde, como por exemplo, na medicina social. Segundo Dimeglio *et al.* (2016), a medicina social visa levar em conta o contexto social como um determinante da saúde da população. Dimeglio *et al.* (2016) afirmam que a saúde não é descrita como o único resultado de características individuais, em particular determinantes

biológicos, mas também é determinada pelo contexto, pelo ambiente, pela estrutura da sociedade em que os indivíduos e as populações vivem. Dimeglio *et al.* (2016) em seu estudo pressupõem que os dados de várias bases de dados e diversidade geram informações e prover conhecimento sobre saúde e o ambiente que a constrói.

Para o *National Institutes of Health* (NIH), esse potencial do Big Data gera oportunidades para novas colaborações providas de fontes distintas de dados que desempenharão um papel fundamental no avanço do conhecimento sobre a causa de doenças, na melhoria do atendimento ao paciente e na promoção de comunidades mais saudáveis (CRUMP; SUNDQUIST; WINKLEBY, 2015). O crescimento exponencial permite que tratamentos médicos sejam desenvolvidos e aprovados mais rapidamente, uma individualização da medicina e uma evolução na saúde pública (NATURE, 2016).

OBJETIVO GERAL

Compreender as abordagens e limites da aplicação do Big Data na saúde pública por meio de uma revisão de literatura.

OBJETIVOS ESPECÍFICOS

Identificar áreas da saúde de aplicação do Big Data.

Identificar limites de aplicação do Big Data.

METODOLOGIA

Para esse estudo foi realizada uma revisão da literatura entre os meses de janeiro e fevereiro de 2018. A busca por artigos foi realizada através de bases eletrônicas e busca manual da literatura. O portal Periódico Capes foi a base principal de consulta dos artigos.

Para a seleção dos artigos foram utilizadas as palavras-chaves em português: Big Data (no título) e saúde pública (no assunto). Em inglês foram utilizadas as palavras-chaves: Big Data (no título), *public health* (no assunto). Na combinação dos termos foi usado o parâmetro *AND*. Os termos sobre saúde foram validados pelos Descritores em Ciências da Saúde (DeCS) - BiREME. Os termos sobre tecnologia validados nos dicionários de Oxford e Wikipédia.

Como critério de exclusão foi escolhido artigos compreendidos entre os anos de 2015 e 2018, artigos somente revisados por pares, artigos em idiomas português ou inglês e artigos compreendidos na categoria tipo de material 'artigos'.

Logo em seguida, pelo próprio portal, filtro de tópicos foi aplicado utilizando a palavra chaves Big Data e *public health*. Buscou-se analisar os detalhes dos artigos e descartar os que não continham as palavras chaves *public health* e Big Data. Posteriormente, ler e compreender os resumos com a finalidade de qualificar seleção, e posteriormente, uma leitura completa dos artigos restantes descartando todos os artigos que não tinham similaridade com os objetivos geral e específico desse trabalho.

RESULTADOS E DISCUSSÃO

Foram encontrados 223 artigos com as palavras chaves em inglês e somente um artigo com as palavras chaves em português. Após aplicação dos critérios de exclusão restaram 154 artigos. Pós aplicação do filtro de tópico Big Data, sobraram 67 artigos e posteriormente, *public health*, restaram 46 artigos. Dos artigos restantes, pós-leitura dos detalhes e resumos, restaram 21 artigos. Dos 21 artigos restantes, 10 artigos foram descartados, pois ao final da leitura completa, verificou-se que o contexto não estava coerente com os objetivos desse estudo. Os 11 artigos restante foram divididos em quatro categorias: aplicado à medicina social (a), aplicado a saúde pública (b), relacionado a conceitos gestão em saúde (c), relacionado à preocupação com a ética (d).

a) Aplicado à medicina social

Dimeglio *et al.* (2016) relatam que na medicina social, o contexto social é um fator determinante para a saúde da população. Segundo os pesquisadores, o contexto social e a sociedade que cada indivíduo ou uma população está inserido são fatores determinantes num quadro de saúde.

Nesse contexto, a contribuição potencial do Big Data aparece significativa na medida em que informações sobre uma ampla gama de características do meio ambiente ou contexto de vida (social,

econômico, cultural) podem ser coletadas e serem conectadas com dados de saúde no desenvolvimento de modelos sobre estilo de vida e condições socioeconômicas, culturais e ambientais gerais.

Os pesquisadores salientam que a resposta do “*why?*” para saúde seja mais importante do que “*what?*”, pois na saúde, entender o porquê é essencial para o desenvolvimento de intervenções na área. Segundo os autores, o uso dessas tecnologias gera novas hipóteses para a pesquisa e gera novas ideias que terão que ser testadas de forma adequada. Sob tais condições, uma visão mais completa e mais global da saúde por meio da adoção de uma abordagem mais abrangente devido ao volume de informações proporcionado pelo Big Data.

b) Aplicado à saúde pública

Estudos na área de saúde pública relatam que as exposições ambientais são reposáveis pela maioria dos fatores de riscos para muitas doenças. Nesse sentido, pesquisas urgem por métodos mais abrangentes e sistemáticos que qualifiquem os impactos sobre a saúde humana (ADLAKHA, 2017). Adlaka (2017) ressalta que condições de vizinha como ambiente físico (disponibilidade de calçadas, ciclovias, parques, etc.), ambiente social (taxa de pobreza, coesão de comunidade, desemprego) e determinada aspectos ambiental (taxa de poluição, emissões de gases, resíduos biodegradáveis) também contribuem para uma diminuição da saúde pública da população residente.

Pesquisas recentes buscam quantificar esse impacto por meio de método, como por exemplo, de extensa observação de campo. Esse tipo de método gera uma limitação de visão de comportamento em momentos

específicos. Nesse sentido, Adlakha (2017) afirma que o uso de novas abordagens tecnológicas (câmeras, sensores, *Global Position System* – GPS) e aplicações de Big Data ajudam a enumerar padrões globais de atividade física e construir características ambientais com amplas aplicações nos campos de planejamento na saúde pública. O uso do Big Data e as tecnologias emergentes podem fornecer um caminho para um custo mais reduzido e uma melhor qualificação de resultados no rastreamento preciso de comportamentos em ambientes espaciais e temporais maiores.

Para Adlakha (2017), no futuro, novas pesquisas devem se concentrar no desenvolvimento do uso de tecnologias emergentes, como Big Data, para estabelecer dados de linha de base, bem como fornecer generalização ao método.

Apte *et al.* (2017) destacam que um dos maiores fatores de risco para saúde pública é a poluição do ar. Afirmam também que embora seja essencial a medição da qualidade do ar, as extensões das observações da poluição terrestre também são limitadas. Desde início do ano de 2000, avanços nos métodos de avaliação da exposição à poluição do ar ajudaram a abordar as limitações de cobertura de dados e resolução espacial.

Segundo os autores, o uso de tecnologia emergentes e inovadora como *Google Street View* (SV), Big Data e técnicas de mineração de dados ajudam a ampliar e qualificar as informações e agregam mais valor nos resultados das pesquisas. Esses dados resultantes da alta resolução e precisão sobre a qualidade do ar em

todas as principais áreas urbanas podem ter implicações transformadoras no gerenciamento ambiental, na ciência da poluição do ar, na epidemiologia, na conscientização sobre saúde pública e na política.

c) Relacionado a conceitos gestão em saúde

O conceito de saúde única é entendido por um esforço colaborativo baseado em abordagem sistêmica de várias disciplinas que trabalham localmente, nacionalmente e globalmente para atingir a saúde ideal para pessoas, animais e o meio ambiente (ASOKAN; ASOKAN, 2015). Segundo os pesquisadores, anualmente 16 milhões de pessoas morrem em todo o mundo devido a causas infecciosas.

Nesse cenário, há uma necessidade especial de aumento da eficácia de avaliação em saúde. O uso de novas fontes de dados inexploradas que podem transmitir dado e informações de forma rápida e precisa para detectar tendências de doenças, surtos, etc. se faz necessário (ASOKAN; ASOKAN, 2015).

O Big Data sinergiza diferentes sistemas, oferece um custo de geração de dados e agregada muita riqueza nos resultados (ASOKAN; ASOKAN, 2015). Já para Salathé (2016), o uso de Big Data na área de farmacovigilância pode ser usado para fins bastante distintos da saúde pública. O uso generalizado da Internet e das mídias sociais em particular tem um efeito dramático não apenas na vigilância de doenças infecciosas, mas também na vigilância do uso de drogas e eventos relacionados.

Stieb, Boot e Turner (2017) descrevem sobre o potencial do Big Data para ajudar na tomada de decisão. Exemplifica ao ilustrar, como exemplo, a saúde ocupacional e ambiental. Ambas, na área de pesquisa, fizeram

uso generalizado de grandes conjuntos de dados. Já Mooney e Pejaver (2017) afirmam que diversos autores fizeram uso do Big Data na área de vigilância em saúde, epidemiologia e outros aspectos em gestão em saúde. Os autores afirmam que usos típicos de Big Data em pesquisas etiológicas e clínica, triagem, vigilância em saúde e tomadas de decisões clínicas.

d) Relacionado à preocupação com a ética

Segundo dicionário web, ética é definida como um valor social que identifica, qualifica e guia princípios universais e crenças e ações humanas. A palavra "ética" vem do grego *ethos* e significa aquilo que pertence ao "bom costume", "costume superior", ou "portador de caráter" (Wikipédia).

Salermo *et al.* (2017) ressaltam que o avanço de tecnologia como Big Data e a profusão da tecnologia em área da saúde como a epidemiologia contribuíram para o surgimento de novas questões éticas: o uso das informações sigilosas; riscos, veracidade de fontes como redes sociais; dados coletados diferentes do objetivo da pesquisa original. Os autores questionaram que um ponto ético crítico é o impacto potencial na autonomia quando os dados são vinculados a várias fontes de dados sem a permissão do indivíduo ou do consumidor ou o consentimento informado, podendo violar termos de confidencialidade, integridade, privacidade.

No combate a quebra de diretrizes de segurança ética, legislações e diretrizes nacionais e internacionais são implementadas (WAYENA *et al.*, 2015). Alguns tópicos relacionados ao Big Data são

discutidos pelos autores como privacidade, transparência, responsabilidade, risco, integridade entre outros.

Em outro artigo, Wayena *et al.* (2018), os autores consideram que ao adotar o uso de Big Data, políticas devem ser estabelecidas seguindo os princípios de segurança. Segundo Wayena *et al.* (2018), três áreas devem ter foco mais abrangente: explorar plenamente os dados acessíveis a quem de direito com finalidade para o bem público; devido a alta complexidade e sofisticação dos métodos de gerenciamento desses dados, processos interativos que envolvam as partes interessadas devem manter um alto nível de fatores críticos para o desenvolvimento e manutenção da confiança; estratégias eficazes para garantir a integração de dados, interoperabilidade e padrões de segurança são essenciais para garantir segurança e proteção dos dados.

CONCLUSÃO

O uso de Big Data em saúde demonstra ser uma prática em evolução. Dessa forma, o uso de novas fontes de dados torna-se relevante para área da saúde. Tendo em vista os aspectos observados, quatro áreas em destaque: medicina social, saúde, cidade e a preocupação com a ética. Entende-se que vários aspectos relativos à segurança e ética foram citados como integridade, confidencialidade, responsabilidade e transparência.

É imprescindível que todos se conscientizem de que uma nova era chegou e novos desafios serão expostos à comunidade acadêmica bem como a sociedade em geral. Então, faz-se necessário a criação de políticas

e diretrizes integradas com o Big Data e computação bem como o gerenciamento e manutenção dessas diretrizes.

Conclui-se que o uso do Big Data em saúde pública é uma realidade, mas o futuro é uma grande promessa para melhorar a saúde. Pois Bi Data possui um grande potencial ainda a ser explorado e requer uma abordagem cuidadosa e diferenciada.

REFERÊNCIAS

ADLAKHA, D. Quantifying the Modern city: emerging Technologies and Big data for Active Living Research. *Public Health* v. 5, p. 105, 2015.

APTE, J. S et al. High-Resolution Air Pollution Mapping with Google Street View Cars: Exploiting Big Data. *Environmental Science and Technology*, v. 51, p. 6999–7008, 2017.

ARIMA, H. Utilizing Big Data for Public Health. *J Epidemiol*, v. 26, n.3, p. 105, 2016.

ASOKAN, G.V.; Asokan, V. Leveraging 'big data' to enhance the effectiveness of 'one health' in an era of health informatics. *J Epidemiol Global Health*, 2015.

BOURNE, P. E. Confronting the Ethical Challenges of Big Data in Public Health. *PLoS Comput Biol*, v. 11. p. 2, 2015.

CRUMP, C; SUNDQUIST, K; WINKLEBY, M.A. Transnational research partnerships: leveraging big data to enhance US health. *J Epidemiol Community Health Month*, 2015.

DIMEGLIO, C. et al. Expectations and boundaries for Big Data approaches in social medicine. *Journal of Forensic and Legal Medicine*, 2016.

MOONEY, S.J.; PEJAVER, V. Big Data in Public Health: Terminology, Machine Learning, and Privacy. *Annual Review of Public Health*, v.39, p. 1-18, 2017.

NATURE. The power of big data must be harnessed for medical progress. But how?. Nature, v. 539, p. 467, 2016.

SALATHÉ, M. Digital Pharmacovigilance and Disease Surveillance: Combining Traditional and Big-Data Systems for Better Public Health. The Journal of Infectious Diseases, v. 214 (s. 4), p. 399-403, 2016.

SALERMO, J. et al. Ethics, Big Data and Computing in Epidemiology and Public Health. Annals of Epidemiology, 2017.

STIEB, D.M.; BOOT, C.R.; TURNER, M.C. Promise and pitfalls in the application of big data to occupational and environmental health. BMC Public Health, v. 17, p. 372, 2017.

TAMBE, P. Big Data Investment, Skills, and Firm Value. Management Science, v.60, n. 6, p. 1452-1469, 2014.

WAYENA, E. et al. Policy implications of big data in the health sector. Bull World Health Organ, v. 96, p. 66–68, 2018.

_____. Ethical Challenges of Big Data in Public Health. PLoS Comput Biol, v. 11, p. 2, 2015.

WILLIAMS, R.F.; SMITH, G.P. Using 'Big Data' to Optimize Public Health Outreach. JAMA Dermatology, v. 151, n. 4, 2015.

WIKIPEDIA. Ética. Disponível em: <<https://pt.wikipedia.org/wiki/Ética>>. Acesso em: 14.mar.018.