



**UNIVERSIDADE  
FEDERAL  
DE PERNAMBUCO**

**Universidade Federal de Pernambuco  
Centro de Tecnologia e Geociências  
Departamento de Engenharia Biomédica  
Graduação em Engenharia Biomédica**

**Luana Nunes da Silva**

# **Transformando a Saúde com Big Data e Inteligência Artificial: Uma Análise Abrangente**

Recife  
2024

Luana Nunes da Silva

# **Transformando a Saúde com Big Data e Inteligência Artificial: Uma Análise Abrangente**

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia Biomédica pela Universidade Federal de Pernambuco.

Orientador (a): Profa. D.Sc. Cristine Martins Gomes de Gusmão

**Recife, 25 de Junho de 2024.**

Ficha de identificação da obra elaborada pelo autor,  
através do programa de geração automática do SIB/UFPE

Silva, Luana Nunes da.

Transformando a Saúde com Big Data e Inteligência Artificial: Uma  
Análise Abrangente / Luana Nunes da Silva. - Recife, 2024.

47 p. : il., tab.

Orientador(a): Cristine Martins Gomes de Gusmão

Trabalho de Conclusão de Curso (Graduação) - Universidade Federal de  
Pernambuco, Centro de Tecnologia e Geociências, Engenharia Biomédica -  
Bacharelado, 2024.

9.

Inclui referências.

1. Big Data. 2. Inteligência Artificial em Saúde. 3. Big Data em Saúde. I.  
Martins Gomes de Gusmão, Cristine. (Orientação). II. Título.

600 CDD (22.ed.)

Luana Nunes da Silva

# **Transformando a Saúde com Big Data e Inteligência Artificial: Uma Análise Abrangente**

Monografia apresentada como requisito parcial para obtenção do diploma de Bacharel em Engenharia Biomédica pela Universidade Federal de Pernambuco.

Aprovado em: 25/06/2024.

## **BANCA EXAMINADORA**

---

Profa. D.Sc. Cristine Martins Gomes de Gusmão  
(Orientadora) Universidade Federal de Pernambuco

---

Prof. D.Sc. Emery Cleiton Cabral Correia Lins  
(Examinador Interno) Universidade Federal de Pernambuco

# Agradecimentos

Agradeço primeiramente a minha mãe e irmão, Marcélia Nunes e Ygor Nunes, pelo amor, apoio e incentivo incondicional que sempre me deram. Agradeço também aos meus amigos, que sempre estiveram ao meu lado, nos momentos bons e nos momentos ruins. Agradeço aos meus professores, que me ensinaram muito ao longo do curso. Em especial, agradeço à minha orientadora, Cristine Gusmão, pelo apoio e orientação. Agradeço à equipe de apoio psicossocial da Universidade Federal de Pernambuco, que me ajudou a superar alguns momentos difíceis.

# Resumo

A era do Big Data tem transformado setores diversos, desde entretenimento à medicina, por exemplo. Neste contexto, o campo da saúde tem vivenciado um aumento expressivo no volume de dados, abrangendo textos, sons e imagens, desafiando bancos de dados tradicionais. Tal crescimento necessita de uma ciência de dados robusta, que apoie decisões clínicas e melhore a eficácia dos tratamentos. Contudo, para processar esses volumosos registros eletrônicos, muitas vezes não estruturados, é necessário utilizar métodos avançados. A integração desses dados e sua transformação em visões acionáveis prometem revolucionar o atendimento ao paciente, aproximando-o de uma medicina cada vez mais personalizada. A análise de dados em saúde, envolvendo mineração de registros médicos, Big Data e Inteligência Artificial, possuem ampla diversidade temática e geográfica. Os estudos abordam desafios, potenciais aplicações e transformações para decisões clínicas. Há lacunas na literatura, especialmente sobre desafios em países em desenvolvimento, destacando a necessidade de mais pesquisas e revisões abrangentes sobre o tema. A saúde, crucial para o bem-estar das sociedades, enfrenta desafios e oportunidades com a tecnologia e a expansão de dados. Estes podem otimizar diagnósticos e tratamentos, mas trazem preocupações de integração, privacidade e ética. É vital analisar tais avanços detalhadamente, promovendo parcerias e capacitação. Garantindo a segurança dos dados e considerando a ética, a tecnologia pode transformar a assistência médica de forma revolucionária.

**Palavras-chave:** Big Data; Inteligência Artificial em Saúde; Big Data em Saúde.

# Abstract

The era of Big Data has been transforming various sectors, from entertainment to medicine, for example. In this context, the field of healthcare has experienced a significant increase in data volume, encompassing texts, sounds, and images, challenging traditional databases. Such growth necessitates a robust data science that supports clinical decisions and improves treatment efficacy. However, processing these voluminous, often unstructured electronic records requires the use of advanced methods. Integrating this data and transforming it into actionable insights promise to revolutionize patient care, bringing it closer to increasingly personalized medicine. Data analysis in healthcare, involving mining of medical records, Big Data, and Artificial Intelligence, encompasses a wide thematic and geographic diversity. Studies address challenges, potential applications, and transformations for clinical decisions. There are gaps in the literature, especially regarding challenges in developing countries, highlighting the need for more research and comprehensive reviews on the topic. Health, crucial for societal well-being, faces challenges and opportunities with technology and data expansion. These can optimize diagnoses and treatments, but bring concerns of integration, privacy, and ethics. It is vital to analyze such advancements in detail, promoting partnerships and capacity building. Ensuring data security and considering ethics, technology can revolutionize healthcare provision.

**Keywords:** Big Data; Artificial Intelligence in Healthcare; Big Data in Healthcare.

# Sumário

Capítulo 1 - Introdução	1
Objetivos	3
Objetivo Geral	3
Objetivos Específicos	3
Estrutura do Documento	3
Capítulo 2 – Proposta Metodológica	4
Capítulo 3 – Análise de Dados	30
Capítulo 4 - Considerações Finais	33
Referências Bibliográficas	37

---

# Capítulo 1 - Introdução

Aspectos comuns da vida cotidiana moderna transformam-se em dados volumosos através do aprendizado de máquina. A Netflix, através de seu sistema de recomendação, identifica quais filmes as pessoas mais gostam de assistir e o Google identifica, através dos históricos de pesquisas de um usuário, o que ele quer saber (BEAM & KOHANE, 2018), por exemplo.

Na medicina não é diferente, exemplos recentes demonstram que o Big Data e o aprendizado de máquina podem gerar algoritmos com desempenho equivalente ao de médicos humanos (BEAM & KOHANE, 2018).

O termo “Big Data” tornou-se frequente no ecossistema acadêmico, e impulsionou o desenvolvimento de programas universitários e pesquisas. Porém, Big Data não se refere apenas a grandes volumes de dados, mas à sua complexidade influenciada por fatores econômicos, culturais e políticos, tornando a análise sociodemográfica uma tarefa desafiadora (SALDANHA, BARCELLOS & PEDROSO, 2021).

Nos últimos anos, o volume de dados de saúde tem crescido rapidamente devido a sistemas de coleta de informações avançados e acessíveis (ARIMA, 2016). Textos, sons e imagens, crescem rapidamente, tornando o armazenamento e processamento desafiadores, bancos de dados relacionais tradicionais não conseguem lidar com a natureza maciça e diversa desses dados (GOLI-MALEKABADI, 2016).

O sistema de saúde atual representa uma complexa rede fundamentada em informações, que requer a constante observação dos pacientes, a transmissão e o compartilhamento contínuo de dados, bem como a utilização de análises sofisticadas de grandes conjuntos de informações para garantir a oferta de serviços de saúde essenciais (SHARMA, CHEN & SHETH, 2018).

A indústria médica coleta uma enorme quantidade de dados, dos quais, a grande maioria são registros eletrônicos (KRIKSCIUNIENE & SAKALAIUSKAS, 2022). Estes dados não podem ser processados e analisados com a utilização de métodos tradicionais de estatística ou de análise de dados devido à complexidade e

---

ao volume, portanto, a descoberta de conhecimento a partir de dados clínicos brutos é um grande desafio para o sistema de saúde (KRIKSCIUNIENE & SAKALAIUSKAS, 2022).

Assim, surge a “ciência de dados”, campo interdisciplinar que auxilia na análise desses grandes conjuntos de dados e na tomada de decisão baseada neles. As técnicas estatísticas atuais, novas fontes de dados e infraestrutura computacional possibilitam análises mais profundas da estrutura social (SALDANHA, BARCELLOS & PEDROSO, 2021).

Esses bancos de dados necessitam de tecnologias específicas para sua transformação em valor, e são extremamente úteis em áreas como a saúde, podendo ajudar a avaliar seus sistemas, a eficácia de tratamentos e a entender condições médicas, que beneficiam o atendimento ao paciente (ARIMA, 2016).

O uso de Big Data na tomada de decisões clínicas aumentou a eficácia no diagnóstico e tratamento e, tecnologias como telemedicina, aprimoram a capacidade dos hospitais primários, enquanto mecanismos de transmissão adequados reforçam a confiança neles (JIANG et al., 2021).

A análise de prontuários médicos, que, em sua maioria, são não estruturados e com vocabulário técnico, é um desafio interdisciplinar, e seus objetivos variam, e incluem a assistência à decisão clínica e apoio à gestão hospitalar (PEREIRA et al., 2015).

O sistema possui camadas de coleta, gerenciamento e serviço de dados e os resultados indicam que a nuvem e o Big Data podem melhorar a atuação dos sistemas de saúde, oferecendo aplicações e serviços de saúde inteligentes (ZHANG et al., 2017).

A valorização da gestão de sistemas de saúde baseados em valor está transformando dados de saúde de meros relatórios de fatos em conhecimento acionável que impulsiona a inovação de tratamentos, previsões terapêuticas e cuidados de saúde personalizados (ABIDI & ABIDI, 2019). A análise avançada de dados de saúde permite o envolvimento dos pacientes, serviços domiciliares abrangentes, integração de diversas fontes de dados e medicina personalizada (ABIDI & ABIDI, 2019).

---

# Objetivos

## Objetivo Geral

Realizar uma revisão bibliográfica sobre os avanços, desafios e aplicações de métodos de Big Data e IA na saúde e sua influência na melhoria da assistência médica.

## Objetivos Específicos

1. Identificar e analisar os principais avanços tecnológicos no tratamento e análise de dados na área de saúde nos últimos anos.
2. Explorar os desafios enfrentados pelos profissionais de saúde e instituições na coleta, armazenamento e interpretação de grandes volumes de dados em saúde.
3. Discutir as implicações éticas, legais e sociais do uso de Big Data e tecnologias relacionadas no contexto da assistência médica.
4. Propor recomendações para a integração eficaz de novos métodos e tecnologias no ecossistema de saúde com base nas descobertas da revisão.

Por meio desses objetivos, a pesquisa busca fornecer uma visão abrangente e equilibrada dos avanços e desafios no tratamento de dados em saúde, contribuindo significativamente para a literatura existente e para as práticas no campo da assistência médica.

## Estrutura do Documento

Após este capítulo introdutório que apresenta uma visão abrangente sobre a importância do estudo tanto da área de Big Data quanto de Inteligência Artificial, esta monografia está estruturada como segue:

- **Capítulo 2 - Proposta Metodológica:** Este capítulo tem como objetivo a exploração da ascensão do Big Data na saúde, abordando sua aplicação interdisciplinar desde bioinformática a informática médica. Enquanto o tratamento de dados faltantes e a integração de Inteligência Artificial oferecem avanços notáveis, enfrenta-se desafios como recursos limitados,

---

questões éticas e infraestrutura insuficiente, especialmente em países em desenvolvimento. A análise destes dados promete melhorar diagnósticos, eficiência clínica e tomada de decisões em saúde pública.

- **Capítulo 3 - Análise de Dados:** Apresenta os resultados da análise de dados, incluindo o montante de artigos identificados, selecionados, lidos e analisados. Traz uma tabela com as análises realizadas, destacando os principais avanços tecnológicos, desafios enfrentados, e as implicações práticas dos estudos revisados. Este capítulo também aborda os critérios definidos para a análise dos artigos.
- **Capítulo 4 - Considerações Finais:** Neste capítulo, são apresentados os resultados e as contribuições da pesquisa, destacando as principais descobertas e suas implicações para a prática e a teoria. As dificuldades enfrentadas durante o estudo são discutidas, e são propostas sugestões para futuras pesquisas. Conclui com uma reflexão sobre o impacto das tecnologias de Big Data e Inteligência Artificial na saúde e as perspectivas futuras.
- **Referências Bibliográficas:** Lista todas as fontes consultadas e citadas ao longo do trabalho, fornecendo a base teórica e empírica que sustentou a pesquisa.

Esta estrutura é projetada para garantir uma abordagem abrangente e sistemática da pesquisa, facilitando a compreensão dos avanços e desafios no tratamento de dados em saúde e fornecendo uma base sólida para futuras investigações e aplicações práticas.

---

## Capítulo 2 – Proposta Metodológica

Este capítulo tem como objetivo a exploração da ascensão do Big Data na saúde, abordando sua aplicação interdisciplinar, desde bioinformática a informática médica. Enquanto o tratamento de dados faltantes e a integração de Inteligência Artificial oferecem avanços notáveis, enfrenta-se desafios como recursos limitados, questões éticas e infraestrutura insuficiente, especialmente em países em desenvolvimento. A análise destes dados promete melhorar diagnósticos, eficiência clínica e tomada de decisões em saúde pública.

A era do Big Data na saúde é assinalada pelo rápido aumento dos dados médicos gerados a partir de sistemas de informação hospitalar (SIH) e com ela a necessidade distribuída e interdisciplinar da realização de novos métodos de análises (LI, ZHANG & TIAN, 2016).

Seu uso estende-se para diversas áreas científicas na saúde: bioinformática, imagens médicas, informática de sensores, e informática médica, porém, a adoção plena da análise enfrenta obstáculos em termos de evidências práticas, recursos humanos capacitados e investimentos financeiros (NAZIR et al., 2020).

Um dos aspectos mais importantes nesse campo é o tratamento de dados faltantes em estudos de intervenção. Rioux e Little (2021) discutem a relevância desse processo ao explorar as abordagens já utilizadas, as atuais práticas em andamento e as perspectivas futuras que deveriam ser adotadas.

Essa análise é especialmente relevante no âmbito da saúde pública, onde a utilização do Big Data tem o potencial de impactar positivamente as tomadas de decisões com a utilização de *analytics* para garantir a privacidade em sistemas de saúde baseados em *Intelligence of Things* (IoT) ou Inteligência das Coisas e nuvem, mantendo a utilidade dos dados (SHARMA, CHEN & SHETH, 2018). É por isso que sua análise desempenha um papel crucial na evolução das práticas e pesquisas em saúde e levou à adoção de técnicas para melhorar a qualidade dos serviços (BELLE et al., 2015).

A transformação digital de dados envolve tecnologia médica avançada e a disponibilidade de novos modelos computacionais para lidar com dados médicos volumosos (ABOUDI & BENHLIMA, 2018).

---

A aplicação de inteligência artificial (IA) no tratamento de Big Data em saúde é enfatizada por Kriksciuniene e Sakalauskas (2022) ao destacarem os seus métodos e suas potencialidades no setor (KRIKSCIUNIENE & SAKALAIUSKAS, 2022). Da mesma forma, Beam e Kohane (2018) abordaram o potencial transformador do aprendizado de máquina, uma subárea da IA, no cuidado à saúde.

A complexidade dos dados em saúde impulsionou a utilização contínua da inteligência artificial em diagnósticos, engajamento de pacientes e atividades administrativas, embora desafios de implementação e questões éticas limitem a automação generalizada de empregos na área de saúde (DAVENPORT & KALAKOTA, 2019).

Os avanços recentes na área têm o potencial de aumentar a segurança dos pacientes e a eficiência clínica, por meio de diagnósticos aprimorados e seleção de tratamentos, previsão de riscos e maior eficiência geral (COHEN & KOVACHEVA, 2021).

Isso permite a extração e visualização de informações relevantes, apesar dos desafios de armazenamento, recuperação de informações e eficiência de custos (NAZIR et al., 2020).

Além disso, países em desenvolvimento enfrentam desafios adicionais, como a falta de infraestrutura adequada e recursos para capitalizar plenamente os benefícios do Big Data em saúde (MUHUNZI, KITAMBALA, & MASHAURI, 2023).

A metodologia deste estudo está focada em uma revisão bibliográfica abrangente, e visa identificar, avaliar e sintetizar os avanços e desafios no tratamento de dados em saúde e as aplicações práticas dos métodos mais recentes neste campo.

Os artigos foram selecionados com base em sua relevância para os avanços e desafios no tratamento de dados em saúde, bem como sua abordagem à aplicação de métodos modernos.

A diversidade das fontes dos artigos escolhidos abrange uma ampla gama de editores e fóruns, cada um com seu foco e público específicos, o que destaca a natureza interdisciplinar dos Big data e da inteligência artificial (IA) na saúde. Essa diversidade é crucial por várias razões:

- 
1. **Alcance Amplo de Público:** A inclusão de jornais e editores como o Healthcare Management Forum, IEEE Xplore, JAMA Network Open, Hindawi, entre outros, garante que a pesquisa e as discussões alcancem um amplo público, incluindo profissionais de saúde, pesquisadores, engenheiros e formuladores de políticas. Este amplo alcance é essencial para fomentar a colaboração interdisciplinar.
  2. **Variedade de Perspectivas:** Cada publicação tem seu foco editorial e prioridades. Por exemplo, o *Healthcare Management Forum* enfatiza liderança e gestão dentro da saúde, enquanto o IEEE Xplore foca em inovações tecnológicas e de engenharia. Essa variedade assegura que as discussões sobre Big data e IA na saúde sejam abrangentes, cobrindo aspectos técnicos, éticos, de gestão e clínicos.
  3. **Insights Acadêmicos e Práticos:** A mistura de jornais revisados por pares, como os publicados pela Hindawi ou encontrados na Rede JAMA, com fóruns mais orientados para a prática, como o *Healthcare Management Forum*, oferece um equilíbrio entre pesquisa de ponta e sua aplicação em ambientes de saúde. Esse equilíbrio é crítico para traduzir a pesquisa em prática.
  4. **Perspectivas Globais e Locais:** A seleção de artigos inclui pesquisas de vários países e regiões, conforme mostra a Figura 1, refletindo o interesse global em Big data e IA na saúde. Essa perspectiva global é essencial para compreender e abordar os desafios universais na saúde, como melhorar o cuidado ao paciente, gerenciar sistemas de saúde e combater doenças. Ao mesmo tempo, estudos específicos, como aqueles focados em sistemas de saúde de países particulares, fornecem insights sobre desafios e inovações locais.
  5. **Inovação e Tradição:** Incluir fontes de editores estabelecidos como John Wiley & Sons, que adquiriu a Hindawi, ao lado de plataformas mais novas e jornais de acesso aberto, reflete a natureza evolutiva da publicação científica. Essa evolução é particularmente relevante em campos de rápida movimentação como IA e Big data, onde o acesso aberto e a publicação rápida podem acelerar a inovação.

---

**Figura 1. Distribuição Geográfica das Fontes de Publicação utilizadas nesta pesquisa.**



Fonte: Autor

As fontes utilizadas são predominantemente artigos de revistas científicas, publicações em conferências e trabalhos de pesquisa reconhecidos na área da saúde e tecnologia.

Seu processo de revisão é composto por:

**Palavras-chave e Tópicos:** A busca focou em tópicos relacionados a "Big Data", "inteligência artificial", "aprendizado de máquina" e "sistemas de saúde", conforme evidenciado pelos títulos e temas dos artigos.

**Fontes de Dados:** Foram consultadas bases de dados acadêmicas reconhecidas, visto que as publicações pertencem a jornais e conferências respeitados.

**Filtros Temporais:** A maioria das publicações é recente, destacando-se a relevância e atualidade do tema.

**Diversidade Geográfica:** A busca não se restringiu a uma região específica, indicando uma abordagem global ao tópico.

**Consideração de vários aspectos:** A seleção inclui artigos que abordam aspectos técnicos, metodológicos, aplicativos e desafios do uso de Big Data e IA na saúde.

Proposta de Análise:

---

**Categorização:** Os artigos foram categorizados com base em temas principais, como técnicas de IA, aplicativos de saúde, desafios éticos, gestão de Big Data, etc.

**Síntese:** Para cada categoria, uma síntese dos principais insights, técnicas e descobertas seria realizada.

**Identificação de Tendências:** Analisar os artigos para identificar tendências emergentes, como novas técnicas ou aplicações.

**Desafios e Lacunas:** Avaliar os artigos para identificar desafios comuns enfrentados na aplicação de Big Data e IA na saúde e quaisquer lacunas na pesquisa atual.

**Aplicações Práticas:** Destacar estudos que fornecem insights práticos ou demonstrações de aplicativos no mundo real.

**Contribuições Metodológicas:** Avaliar artigos que introduzem ou propõem novas metodologias para a análise de dados de saúde.

A análise proposta visa extrair insights valiosos dessa rica coleção de literatura, fornecendo um entendimento robusto e abrangente do campo. Este entendimento permitirá identificar tendências emergentes, desafios persistentes e oportunidades para futuras pesquisas e aplicações práticas na área da saúde. Além disso, a síntese das informações encontradas orientará a definição de estratégias eficazes para a implementação de tecnologias de Big Data e Inteligência Artificial em ambientes clínicos e de pesquisa.

Ao concluir este capítulo metodológico, o próximo capítulo, **Capítulo 3 - Análise de Dados**, apresentará detalhadamente os resultados obtidos a partir dessa análise. Será feita uma discussão sobre os avanços tecnológicos, as dificuldades enfrentadas e as aplicações práticas mais relevantes identificadas na literatura revisada, oferecendo uma visão clara das contribuições mais significativas para a área de saúde.

---

## Capítulo 3 – Análise de Dados

A partir das bases de dados pesquisadas, os resultados indicam uma diversidade de estudos sobre a análise de dados em saúde, com ênfase na mineração de registros médicos eletrônicos, Big Data e a integração com Inteligência Artificial, conforme mostra a Tabela 1.

Os estudos revisados abordam diferentes aspectos deste campo, incluindo desafios de privacidade, tratamento de dados ausentes, potenciais aplicações em saúde pública e análise de dados para decisões clínicas.

As publicações analisadas vêm em formatos variados, como revisões, estudos de corte e investigações específicas de técnicas e métodos. Eles exploram temas como o potencial da Inteligência Artificial na análise de dados em saúde, a mineração de registros médicos eletrônicos e a transformação de grandes volumes de dados para decisões clínicas.

A distribuição geográfica dos estudos é notavelmente diversificada, com contribuições oriundas de diferentes regiões, incluindo Canadá, Estados Unidos, Japão e até mesmo estudos que se concentram especificamente nos desafios enfrentados pelos países em desenvolvimento.

No entanto, apesar da riqueza de informações encontradas, é evidente que existem lacunas na literatura, particularmente quando se trata de desafios específicos relacionados à infraestrutura e ao conhecimento em países em desenvolvimento.

A complexidade inerente à mineração de registros médicos eletrônicos e à interpretação correta dos dados colhidos também sugere a necessidade de mais pesquisas na área para abordar essas lacunas e maximizar o potencial da análise de dados na saúde.

Esses resultados sugerem a importância de realizar uma revisão bibliográfica da literatura para analisar as informações disponíveis nos últimos 10 anos e cobrir as lacunas no conhecimento sobre o tema.

Os critérios para seleção foram:

---

**Identificação:** Artigos encontrados através da busca inicial em diferentes bases de dados.

**Seleção:** Artigos selecionados após uma primeira triagem dos títulos e resumos.

**Leitura:** Artigos lidos integralmente para avaliar sua relevância.

**Análise:** Artigos incluídos na análise final após avaliação completa.

**Tabela 1. Tabela de Análise de Artigos**

<b>Tema</b>	<b>Artigos (Primeiros Autores e Ano)</b>	<b>Quantidade de Autores</b>
Big Data em Saúde	ABIDI (2019), ABOUDI (2018), BELLE (2015), NAZIR (2020), MUHUNZI (2023)	5
Arquitetura e Implementação	ABOUDI (2018), GOLI-MALEKABADI (2016)	2
Inteligência Artificial em Saúde	ABIDI (2019), DAVENPORT (2019), KRIKSCIUNIENE (2022)	3
Machine Learning em Medicina	BENNIS (2020)	1
Análise de Dados de Saúde	ARIMA (2016), BEAM (2018), CHOI (2015), MEYSTRE (2017)	4
Cloud Computing e Big Data	GOLI-MALEKABADI (2016)	1
IoT e Qualidade de Dados	BANERJEE (2017), SHARMA (2018)	2
Utilização de Dados Médicos	CHOI (2015), NIU (2019)	2
Desafios e Oportunidades	MUHUNZI (2023)	1

---

Análise de Dados em Cloud	GOLI-MALEKABADI (2016)	1
Aplicação de IA em Saúde	COHEN (2021), KRIKSCIUNIENE (2022)	2
Text Mining em Registros Médicos	PEREIRA (2015)	1

A crescente intersecção entre Big Data e saúde tem levado os pesquisadores a explorar e propor soluções inovadoras para desafios enfrentados no setor de saúde. Dessa forma, para guiar a análise, os seguintes critérios foram definidos:

- **Relevância Temática:** O artigo se alinha diretamente com o tema central da revisão?
- **Qualidade Metodológica:** O artigo apresenta uma metodologia rigorosa e bem definida?
- **Originalidade:** O artigo oferece insights ou descobertas únicas que agregam valor ao campo de estudo?
- **Aplicabilidade Prática:** O artigo discute implicações práticas ou apresenta soluções aplicáveis?
- **Fonte de Publicação:** O artigo foi publicado em uma revista ou conferência reconhecida?

ABIDI & ABIDI (2019), discutem como os modelos de sistemas de saúde modernos estão mudando de sistemas baseados em volume para sistemas baseados em valor. Sendo essa mudança reflexo de uma nova percepção sobre o valor dos dados em saúde, que está evoluindo da simples coleta de fatos para a geração de conhecimentos aplicáveis na inovação de tratamentos, medicina personalizada, entre outros avanços.

Além disso atribuem a crescente importância da IA na área da saúde e medicina a dois principais fatores: (a) a maior disponibilidade de frameworks de IA de código livre com aplicações gerais, que democratizam o acesso e utilização dessas essas ferramentas por indivíduos sem formação na área; e (b) A habilidade de combinar diferentes tipos de dados de saúde para formar extensos bancos de

---

dados capazes de explorar questões intrincadas associadas ao gerenciamento da saúde.

Concluem que a convergência de Big Data e IA permite métodos de análise de dados de saúde inteligentes, capazes de aprender autonomamente associações não lineares e relações causais inerentes aos dados. Esses métodos são empregados para gerar serviços de saúde inovadores orientados por dados, como medicina de precisão, previsão de tendências de doenças e otimização da utilização do sistema de saúde.

Embora apresente uma abordagem sistemática na exploração das capacidades de AI e Big Data em dados em saúde, o artigo poderia apresentar uma seção explorando os critérios e metodologia usados em sua elaboração. Assim como aprofundar sobre as metodologias específicas utilizadas nos estudos citados, assim enriquecendo a discussão.

ABIDI & ABIDI (2019) conseguem oferecer uma visão abrangente e clara dos assuntos propostos, contextualizando conceitos-chaves sobre Big Data e análise de dados em saúde. A aplicabilidade do artigo é vasta, abrangendo a gestão de sistemas de saúde, desenvolvimento de políticas, prática clínica e pesquisa, fornecendo insights valiosos para a otimização de recursos, melhoria da qualidade dos cuidados e inovação em saúde.

ABOUDI & BENHLIMA (2018) defendem a importância de arquiteturas extensíveis capazes de tratar e processar grandes volumes de dados e propõe uma nova arquitetura de Big Data que integra computação em lote e em tempo real para melhorar a confiabilidade dos sistemas de saúde, gerando alertas em tempo real e fazendo previsões precisas sobre condições de saúde dos pacientes.

A arquitetura proposta e a implementação de um protótipo baseado em *Spark* e *MongoDB* demonstram potencial para detecção de casos de emergência e geração de alertas em tempo real, destacando a importância de abordagens de Big Data na gestão de saúde.

Na arquitetura apresentada são considerados dados de prontuários eletrônicos, imagens médicas, sensores eletrônicos e redes sociais. O que destaca a busca por uma arquitetura flexível capaz de lidar com dados em diferentes formatos.

---

O artigo apresenta uma revisão bibliográfica completa ao detalhar outras soluções empregadas em Big Data em saúde. E utiliza em sua solução ferramentas amplamente utilizadas no mercado de dados com alta flexibilidade e escalabilidade. proposta, suas camadas, e a implementação do protótipo, mostrando aplicabilidade prática.

Além disso, oferece contribuições originais ao combinar computação em lote e em tempo real considerando dados de diferentes formatos e origens, abordando tanto a detecção de emergências quanto a previsão de condições de saúde.

Assim como ABOUDI (2018), BELLE (2015) também destaca o uso de frameworks como Hadoop e MapReduce para análise e transformação de grandes volumes de dados, que proporcionam escalabilidade e aplicabilidade em uma variedade de situações do mundo real.

O artigo de BELLE (2015) explora a aplicação de análises de Big Data no setor de saúde, focando na gestão, análise e assimilação de grandes volumes de dados estruturados e não estruturados gerados pelos sistemas de saúde atuais.

Destaca-se a aplicação dessas análises na melhoria da entrega de cuidados de saúde e na exploração de doenças. Os autores enfatizam os desafios fundamentais inerentes ao paradigma de Big Data na saúde e discutem áreas promissoras como análises baseadas em imagem, sinal e genômica.

Ao explorar a utilização de grandes volumes de dados médicos e combinar dados multimodais de fontes diversas, o artigo identifica áreas de pesquisa com potencial para impactar significativamente a entrega de cuidados de saúde.

O artigo apresenta uma abordagem metodológica sólida ao detalhar as tecnologias e técnicas utilizadas para análise de dados em grande escala, apesar de não se aprofundar em metodologias específicas.

ARIMA (2016) discute o potencial e os desafios de utilizar grandes volumes de dados para a saúde pública. Traz uma advertência ao cenário, enfatizando que a mera existência de grandes volumes de dados não garante *insights* valiosos. Ele exemplifica com sua pesquisa no Japão, indicando que muitos desafios ainda precisam ser superados, sendo a privacidade um dos principais.

---

O editorial conclui que Big data oferece oportunidades sem precedentes para avançar na saúde pública por meio da análise de grandes volumes de dados de várias fontes. E que sua utilização eficaz pode facilitar a identificação de padrões de saúde, previsão de epidemias, melhoria dos serviços de saúde e personalização de tratamentos.

Dado o formato editorial, a profundidade metodológica é limitada. No entanto, a abordagem baseada em visão geral e opinião é apropriada para o objetivo de discutir ideias gerais e tendências no campo. O editorial contribui com uma discussão oportuna sobre o uso de Big Data na saúde pública, um tema de crescente interesse e importância. As ideias apresentadas têm ampla aplicabilidade para profissionais e pesquisadores em saúde pública, bem como para formuladores de políticas interessados em integrar análises de Big Data em suas estratégias de saúde.

Nesse ínterim é justamente onde Cohen e Kovacheva (2021) entram, apresentando uma abordagem técnica para o problema. Eles propõem uma plataforma focada na padronização de dados, visando não só melhorar a análise, mas também abordar preocupações éticas inerentes ao tratamento de dados sensíveis na saúde.

A plataforma MERLIN (Medical Record Longitudinal Information AI System) foi desenvolvida com um design modular e adaptável, permitindo a inclusão de diferentes módulos de ETL, QA/QC (Quality Assurance/Quality Control), criação de conjuntos de dados e análise de dados.

Ao abordar os principais desafios enfrentados no fluxo de trabalho de análise de dados biomédicos, tais como a qualidade dos dados — destacando a inconsistência e a incompletude dos registros eletrônicos de saúde (EHRs) —, a ausência de padronização nos dados e as restrições na infraestrutura tecnológica necessárias para armazenar e processar vastos volumes de dados biomédicos, os autores propõem uma metodologia guiada por princípios fundamentais destinados a otimizar o manuseio e armazenamento de dados, visando superar esses obstáculos de maneira eficaz. Estes princípios incluem:

- 
- **Atomicidade:** Garante que cada elemento de dado seja mantido de maneira indivisível, simplificando significativamente a administração e análise dos dados.
  - **Consistência:** Assegura a uniformidade dos dados mediante a implementação de formatos padronizados e procedimentos rigorosos de verificação, melhorando a qualidade e a confiabilidade dos dados.
  - **Traduzibilidade:** Facilita a conversão ou tradução dos dados para diversos formatos, adaptando-se às necessidades específicas das análises, e promovendo a interoperabilidade entre diferentes sistemas e plataformas.
  - **Rastreabilidade:** Permite o rastreamento detalhado da proveniência dos dados, abrangendo a metodologia de sua coleta e processamento, reforçando a integridade e a confiabilidade dos dados.

Assim como a solução proposta por ABIDI & ABIDI (2019), esta solução também considera dados com diferentes origens e formatos. Considerando dados de prontuários eletrônicos, sensores de leito e dados coletados manualmente.

Os dados coletados são tratados, padronizados e disponibilizados para análise, através de processos ETL. A estrutura modular da plataforma facilita a adição de novas fontes de dados e métodos de análise, promovendo a colaboração e a eficiência na pesquisa em saúde.

O artigo apresenta uma metodologia sólida ao sistematizar e analisar criticamente a literatura existente, identificando lacunas e sugerindo direções futuras para pesquisa. Além de contribuir originalmente ao mapear o estado da arte da mineração de texto em prontuários médicos eletrônicos e ao identificar áreas que necessitam de mais investigação.

Por fim, a plataforma proposta tem grande potencial de facilitar significativamente a coleta de dados, criação, avaliação e análise de conjuntos de dados em formatos prontos para aplicação em aprendizado de máquina. Provendo um impacto positivo no desenvolvimento de modelos de IA e na melhoria da entrega de cuidados de saúde.

Devenport e Kalakota (2019) concordam com a relevância da IA na saúde, mas também apresentam uma visão mais pragmática. Eles reconhecem o potencial

---

da IA, desde diagnóstico até a administração, porém advertem sobre os desafios de sua integração e as implicações éticas, reiterando que, apesar de promissora, a trajetória da IA a saúde é longa e repleta de obstáculos a serem superados.

Os autores também levantam questões éticas relacionadas ao uso da IA na saúde, incluindo responsabilidade, transparência e privacidade, e a necessidade de governança e monitoramento contínuos.

Em especial, discutem a existência de estudos e pesquisas sugerindo que a IA pode desempenhar tão bem quanto ou melhor que humanos em tarefas específicas da saúde, como na identificação de tumores malignos em exames de raio-x. E diante da perspectiva da substituição de profissionais de saúde por inteligências artificiais, destacam que a adoção de IA pode tornar os processos administrativos mais eficientes, especialmente em tarefas repetitivas que consomem tempo dos profissionais de saúde. Otimizando o tempo e esforço do profissional de saúde.

Concluem com uma visão otimista do papel da IA na melhoria da medicina de precisão e no apoio à comunicação com pacientes, destacando desafios significativos na adoção e integração de sistemas de IA na prática clínica diária.

O artigo não descreve explicitamente a metodologia utilizada. A ausência de uma seção dedicada sugere que os autores compilaram e sintetizam informações de uma variedade de fontes para apresentar uma visão geral do potencial da IA na saúde.

No capítulo 9 do livro *“Intelligent Systems for Sustainable Person-Centered Healthcare”*, os autores Kriksciuniene e Sakalauskas (2022) abordam os desafios do volume e complexidade dos dados em saúde, ressaltando a necessidade de métodos analíticos avançados. Os autores apresentam técnicas de inteligência artificial para análise de dados em pacientes com AVC, destacando a relevância da qualidade dos dados e o potencial da IA para otimizar decisões em saúde.

Também oferecem uma análise aprofundada de como a inteligência artificial e os métodos de aprendizado de máquina são aplicados na saúde, focando em mineração de dados e no potencial para melhorar os sistemas de saúde.

---

Ao contrário de ABIDI & ABIDI (2019), que defendem que desde as primeiras inteligências artificiais já existiam aplicações na área da saúde, Kriksciuniene e Sakalauskas afirmam o estágio inicial da utilização da IA na saúde, apontando a falta de procedimentos definidos para a seleção dos melhores métodos analíticos.

Entre os métodos de Inteligência Artificial abordados, incluem-se técnicas estatísticas clássicas e avançadas de aprendizado de máquina, como o aprendizado supervisionado, não supervisionado, por reforço, *deep learning* e métodos de ensemble. Além disso, discute-se o uso de algoritmos específicos, tais como redes neurais, árvores de decisão e algoritmos genéticos, cada um com seu papel distinto na análise e interpretação de grandes volumes de dados, proporcionando insights valiosos para aprimorar sistemas de saúde.

Dada a explosão de dados na saúde, o foco do capítulo em métodos de IA é altamente relevante, oferecendo insights sobre como aproveitar o Big Data para avançar os resultados de saúde e a eficiência do sistema.

MUHUNZI (2023) discute as oportunidades e desafios na implementação de análises de Big Data em países em desenvolvimento. Destacando a necessidade de treinamento em análises de Big Data para profissionais de saúde, a implementação de governança de dados e estruturas regulatórias, a criação de uma cultura de compartilhamento de dados, a implementação de medidas de segurança de dados e a incorporação de computação em nuvem para o armazenamento de dados de saúde.

Entre os desafios apresentados estão a infraestrutura de TI inadequada, altos custos de implementação, preocupações com privacidade e segurança de dados, posse fragmentada de dados e desafios técnicos como qualidade de dados e multidimensionalidade.

Sobre oportunidades, o artigo conclui que a análise de Big Data tem aplicações significativas em análise de imagens médicas, gestão da saúde da população, informática clínica e bioinformática, mesmo em países em desenvolvimento. Os principais desafios incluem a falta de evidências sobre os benefícios práticos, prioridades concorrentes, falta de pessoal treinado, e a necessidade de investimento financeiro substancial.

---

Esse artigo também é composto por revisão da literatura ampla. E utilizou busca por palavras-chaves como "*Big Data analytics*" e "*Big Data analytics healthcare developing countries*", no *PubMed*, *Science Direct* e *Google Scholar*, nas plataformas online para pesquisa acadêmica e científica.

GOLI-MALEKABADI (2016) propõe um modelo focado em dados NoSQL, especificamente em bases de dados baseadas em documentos, para armazenar dados de saúde na nuvem. Capaz de lidar com grandes volumes de dados não estruturados e sua escalabilidade em ambientes distribuídos. Além disso, o modelo utiliza o recurso de *sharding* para distribuir dados de forma eficiente pela nuvem, melhorando o desempenho tanto em operações de leitura quanto de escrita em comparação com sistemas baseados em SQL Server.

O estudo demonstrou que o modelo NoSQL oferece vantagens significativas em termos de flexibilidade, extensibilidade e tempo de preparação de dados, além de ser mais adequado para o armazenamento e recuperação eficientes de grandes volumes de dados de saúde.

O *MongoDB* foi escolhido como o sistema de banco de dados não relacional selecionado por sua capacidade de lidar com grandes volumes de dados não estruturados, sua escalabilidade e por ser adequado para armazenar dados em formato BSON (uma variante de texto do formato JSON), o que facilita o armazenamento de dados complexos de saúde. A *vCloud Ai* foi plataforma de nuvem na qual o modelo foi implementado, fornecendo infraestrutura como serviço (IaaS).

No contexto de "Big Data em Saúde", a solução apresentada é extremamente relevante, pois oferece uma solução potencial para os desafios de gerenciar, armazenar e acessar eficientemente grandes volumes de dados de saúde, facilitando assim a análise de dados, a pesquisa e a tomada de decisões baseada em evidências.

O artigo inicialmente demonstra uma abordagem metodológica focada na identificação de desafios específicos associados ao armazenamento de Big Data de saúde e propõe um modelo para superá-los.

---

Em NAZIR (2020) é realizado o estudo e relato da literatura detalhado sobre o gerenciamento, análise e aplicação de Big Data no setor de saúde. Que busca identificar características-chave, técnicas de última geração, soluções ótimas, aplicações e análises dentro do Big Data em saúde, enfatizando o papel da programação científica em aprimorar a entrega de saúde e a tomada de decisões.

O estudo enfatiza a importância do Big Data na transformação da saúde por meio de melhor gerenciamento, análise e programação científica, destacando o potencial para uma melhor identificação de doenças, sugestões de tratamento e cuidados com os pacientes.

Dentre todos os artigos analisados, o trabalho de NAZIR (2020) apresenta a metodologia mais detalhada e robusta. Especificando os critérios de seleção de cada artigo presente na revisão, palavras-chaves usadas, fontes das publicações e países de origem.

A metodologia de pesquisa envolve uma revisão ampla da literatura, garantindo uma análise sólida dos dados coletados e fornecendo uma base confiável para as conclusões tiradas.

Sumariza os resultados da literatura disponível publicada relacionada a Big Data em cardiologia, usando um protocolo de literatura sistemática com dados coletados de trabalhos publicados de 2015 a 2019 (incluindo parte de 2020) de conferências, jornais, livros, revistas e outras fontes online, selecionando 127 artigos mais relevantes para avaliação de qualidade.

Ressalta a crescente importância da análise de Big Data em saúde, particularmente em países em desenvolvimento, mostrando seu potencial para revolucionar os sistemas de saúde por meio da tomada de decisões baseada em dados e cuidados personalizados com os pacientes.

BENNIS (2020), aborda em sua tese de doutorado a aplicação e desenvolvimento de modelos de aprendizado de máquina na medicina, destacando a importância da avaliação de dados brutos, extração de parâmetros e criação de modelos no contexto de problemas clínicos não resolvidos. Através de estudos específicos, demonstra como o aprendizado de máquina pode aprimorar a detecção e previsão de condições clínicas, como a hipovolemia progressiva e o

---

monitoramento da reserva cardio-cerebrovascular durante simulações de hemorragia.

Ressalta a crescente importância da análise de Big Data em saúde, particularmente em países em desenvolvimento, mostrando seu potencial para revolucionar os sistemas de saúde por meio da tomada de decisões baseada em dados e cuidados personalizados com os pacientes.

O estudo sublinha a potencialidade do aprendizado de máquina para melhorar o diagnóstico, tratamento e previsão de desfechos em medicina, enfatizando a necessidade de abordagens metodológicas rigorosas para garantir a aplicabilidade e confiabilidade dos modelos desenvolvidos.

A metodologia utilizada na tese abrange várias etapas cruciais para desenvolver e avaliar modelos de aprendizado de máquina, focando na avaliação da qualidade dos dados brutos, extração de parâmetros e criação de modelos. Utiliza uma abordagem detalhada que inclui a coleta e análise de dados clínicos, aplicação de algoritmos de aprendizado de máquina para identificar padrões e prever desfechos clínicos.

Ferramentas específicas, como MATLAB, são empregadas para análise de dados e simulações, enquanto os dados são coletados de equipamentos de monitoramento padrão em unidades de cuidado intensivo neonatal.

A tese também enfatiza a importância da qualidade dos dados e a necessidade de métodos consolidados para a extração de parâmetros significativos, que são fundamentais para a criação de modelos preditivos precisos e confiáveis em contextos clínicos.

Concluindo que a qualidade e a precisão dos dados são fundamentais para o desempenho dos modelos. A extração de parâmetros baseados em conhecimento de domínio é crucial para a aplicabilidade clínica dos modelos. E que modelos simples podem ser tão eficazes quanto os complexos, dependendo do contexto clínico, com a vantagem de maior transparência e interpretabilidade.

BEAM (2018) artigo de opinião que explora o impacto do Big Data e da aprendizagem de máquina no setor de saúde, afirma que Big data e aprendizagem de máquina têm o potencial de revolucionar a saúde, melhorando a precisão dos

---

diagnósticos, personalizando tratamentos e aumentando a eficiência operacional. E que a adoção dessas tecnologias enfrenta desafios, incluindo questões de privacidade de dados, a necessidade de infraestrutura adequada e a escassez de habilidades especializadas.

O texto aborda a utilização de tecnologias avançadas de deep learning, que são formas sofisticadas de redes neurais. Que têm sido aplicadas com sucesso na análise de imagens médicas, como na identificação de retinopatia diabética em fotos de fundo de olho, alcançando uma precisão comparável ou até superior à de especialistas humanos.

Contudo, para atingir esse nível de precisão, os modelos de deep learning requerem uma quantidade substancial de dados, muitas vezes necessitando de centenas de milhares de imagens para aprender de forma eficaz as características relevantes para prever os resultados desejados.

CHOI (2015) discute a utilização e o valor dos grandes volumes de dados médicos gerados pelos sistemas de informação médica e a aplicação desses dados para a melhoria da prevenção de doenças, tratamento e operações de instituições médicas.

Afirma que é possível integrar e analisar dados de várias fontes dentro dos sistemas de informação médica para gerar novas informações úteis. E que a análise de Big Data médico pode contribuir significativamente para a prevenção de doenças infecciosas, tratamento de doenças e geração de informações estatísticas de receita para gestão de instituições médicas.

Toda essa discussão revela a importância do Big Data e da inteligência artificial na transformação do setor de saúde e a necessidade de abordagens holísticas que considerem os desafios técnicos, éticos e práticos enfrentados pelo setor.

NIU (2019) se dedica a entender como a grande quantidade de dados médicos afeta o sistema de saúde hierárquico e as decisões de tratamento dos pacientes. Através da análise, ele busca entender a influência dos dados médicos no comportamento dos pacientes e dos prestadores de serviços de saúde, propondo

---

maneiras de otimizar a utilização desses dados para melhorar os resultados de saúde e a eficiência do sistema.

O principal objetivo é analisar os efeitos da Big Data médica nas decisões de tratamento dos pacientes dentro de um sistema de saúde hierárquico, identificando como esses dados podem melhorar a eficiência e a equidade no acesso aos cuidados de saúde.

Complementando essa abordagem, realiza-se uma revisão da literatura que se debruça sobre temas como a assimetria de informações médicas, a heterogeneidade do conhecimento médico e as características dos serviços médicos como “bens de credência”. Essa revisão destaca a disparidade entre o nível de conhecimento especializado dos médicos em comparação com o dos pacientes, um fator chave na análise subsequente.

Dentro deste contexto, o artigo introduz dois modelos matemáticos: um modelo de jogo de sinalização médico-paciente para estudar as dinâmicas da confiança do paciente nos hospitais e um modelo de jogo de diagnóstico e tratamento de dois níveis para avaliar diagnósticos hierárquicos e encaminhamentos.

**Análise do Modelo de Jogo de Sinalização:** Explora as interações entre pacientes e hospitais, focando na troca de sinais para inferir a qualidade do hospital e os resultados de equilíbrio dessas interações.

**Modelo de Jogo de Diagnóstico e Tratamento de Dois Níveis:** Compara o sistema médico tradicional com o sistema médico hierárquico (HMS), analisando os efeitos dos transbordamentos do conhecimento médico na alocação de recursos médicos e no bem-estar geral do paciente.

Com o suporte do Big Data, a precisão na análise de doenças pode ser significativamente aprimorada, o que, por sua vez, melhora a utilidade total do sistema médico e a utilidade total dos pacientes. O artigo sugere que um sistema médico completo, apoiado por Big Data para tratamento médico inteligente e análise precisa, pode aumentar a utilidade geral dos pacientes e do sistema como um todo.

O Big Data tem o potencial de resolver problemas decorrentes da assimetria de informações entre pacientes e hospitais, melhorando o nível de diagnóstico e

---

tratamento dos hospitais e aprimorando a capacidade dos hospitais de transmitir sinais confiáveis sobre sua qualidade e serviços. Isso pode guiar efetivamente as decisões de tratamento médico dos pacientes, encorajando-os a optar por tratamentos em instituições médicas primárias, mitigando assim o dilema do tratamento primário de baixo nível nas comunidades sob o HMS.

A aplicação do Big Data em decisões clínicas e na eficácia dos programas de tratamento médico tem melhorado significativamente a eficácia na identificação e tratamento de doenças por médicos e hospitais. Além disso, tecnologias como tratamento médico inteligente e diagnóstico e tratamento médico remoto podem ampliar a capacidade de serviço das instituições médicas primárias, aumentando a confiança dos pacientes comuns nessas instituições.

O artigo discute como os *spillovers* de conhecimento médico, facilitados pelo Big Data e pela tecnologia da Internet, podem influenciar as decisões de tratamento e admissão feitas por médicos e pacientes dentro do design do sistema médico hierárquico. Especificamente, questiona-se se a fortificação do Big Data na tomada de decisão médica pode efetivamente alterar o atual dilema do tratamento médico de baixo nível na comunidade.

Para aproveitar plenamente os benefícios do Big Data, o artigo propõe estratégias para as instituições médicas, como prestar total atenção à transmissão de sinais em serviços médicos, desenhar um sistema de comunicação médico-paciente para as necessidades de informação dos pacientes, e aplicar tecnologia de Big Data e inteligência artificial para melhorar a eficiência do mercado de serviços médicos.

Ao concluir a investigação, o estudo resume os achados significativos, ressaltando como a aplicação de grandes volumes de dados pode refinar a eficiência do sistema médico e a qualidade do cuidado ao paciente. Além disso, sugere estratégias inovadoras para aprimorar a transmissão de informações e a utilização de Big Data na tomada de decisões médicas, marcando um passo adiante na busca por um sistema de saúde mais eficiente e equitativo.

MEYSTRE (2017) revisa a reutilização ou o uso secundário de dados clínicos, destacando seu crescimento rápido e a importância para a saúde de alta qualidade,

---

gestão de saúde aprimorada, redução de custos, gerenciamento de saúde da população e pesquisa clínica eficaz.

Aborda as motivações, desafios, questões éticas e de privacidade, integração e interoperabilidade de dados, modelos de dados e terminologias, e exemplos de reutilização de dados clínicos. E tem como objetivo fornecer uma visão ampla da pesquisa publicada sobre reutilização de dados clínicos desde 2005, sem ser uma revisão abrangente de todas as publicações no campo.

O artigo discute dois exemplos de reutilização de dados clínicos:

- **Extração de Medidas de Qualidade:** Utiliza Medidas de Qualidade Clínica (CQMs) para avaliar a eficácia, eficiência e experiência do cuidado ao paciente. A automação da extração dessas medidas a partir de anotações clínicas foi testada com sucesso em relatórios de colonoscopia e categorias de doenças como insuficiência cardíaca, alcançando alta precisão e recall na detecção de adenomas e na contagem de pólipos.
- **Sistemas de Saúde Aprendizes:** Baseiam-se no conceito de um ciclo contínuo de coleta de dados de saúde, extração de conhecimento e sua aplicação na prática clínica, visando a melhoria contínua dos serviços de saúde. Apesar do interesse global, a implementação prática desses sistemas ainda é limitada, enfrentando desafios como acesso aos dados de EHR e interoperabilidade semântica dos dados.

Os autores concluem que a reutilização de dados clínicos é essencial para alcançar uma gestão de saúde melhorada, redução de custos, gestão da saúde da população e pesquisa clínica eficaz. E que existem desafios como questões de privacidade, integração de dados e interoperabilidade.

O artigo é fundamentado em uma pesquisa da literatura em várias bases de dados, incluindo MEDLINE (acessada via PubMed), anais de conferências e a Biblioteca Digital ACM. Utilizou combinações de palavras-chave como “secondary use” com “clinical data,” “clinical information,” ou “clinical record.”

No artigo "Text Mining Applied to Electronic Medical Records: A Literature Review", PEREIRA (2015) discute o papel da mineração de texto na análise de registros médicos eletrônicos (EMRs) para avançar na saúde e apoio à decisão

---

clínica. O foco é na importância de processar dados não estruturados para identificar padrões e insights cruciais que beneficiam diagnósticos e tratamentos. Pereira examina diferentes metodologias, como Knowledge Discovery in Databases (KDD), CRISP-DM e SEMMA, detalhando etapas desde a preparação até a interpretação dos dados. O trabalho também aborda desafios, incluindo preocupações com privacidade e precisão dos dados, e sugere aprimoramentos na integração de dados e no uso de tecnologias de IA avançadas.

O estudo realiza revisão da literatura e identificação de áreas para pesquisa futura, com o objetivo de melhorar os cuidados de saúde por meio da tecnologia. Entre as principais conclusões, destaca-se que a mineração de texto em EMRs é capaz de revelar informações importantes para a saúde, enfrentando obstáculos como a variedade de formatos de dados e as questões de privacidade. As técnicas de mineração de texto abrangem desde o pré-processamento de dados, incluindo limpeza de texto e tokenização, até métodos mais complexos como desambiguação e reconhecimento de entidades, fundamentais para extrair e analisar informações de forma eficaz

Em BANERJEE (2017) utiliza-se análises de casos de uso para discutir desafios e soluções para a qualidade dos dados em diferentes níveis de um sistema de Internet das Coisas (Internet of Things - IoT). Os autores discutem a importância do controle de qualidade dos dados no contexto da IoT, especialmente para aplicações de saúde focando na saúde dos idosos e sistemas de monitoramento residencial.

Ele aborda desafios na manutenção da qualidade dos dados gerados por dispositivos IoT e propõe uma abordagem para avaliar e garantir essa qualidade em diferentes camadas de aplicação.

Propõe melhorar o controle de qualidade dos dados no contexto da Internet das Coisas ao mapear o framework OSI (Open Systems Interconnection) para a qualidade do IoT, destacando a importância de considerar diferentes níveis de qualidade (dispositivo, rede e aplicação) e introduzindo a qualidade semântica dos dispositivos. A conclusão enfatiza o impacto significativo da IoT e da Internet de Tudo (Internet of Everything - IoE) na vida cotidiana, através da melhoria da

---

qualidade dos dados da IoT, que pode transformar áreas como entretenimento, vigilância, transporte, atividades diárias, aplicações industriais e saúde.

Ainda sobre IoT SHARMA (2018) aborda o desenvolvimento de análises preservando a privacidade em sistemas de saúde baseados em IoT e nuvem, destacando a necessidade de equilibrar privacidade, eficiência e qualidade do modelo no contexto da saúde digital.

Usando o sistema do kHealth, um sistema de monitoramento e gestão de saúde móvel/digital que no momento do artigo estava em avaliação com pacientes e médicos, apresentou os desafios e discutiu soluções práticas para análises preservando a privacidade em sistemas de informação de saúde.

Discute que componentes de análise em saúde digital envolvem métodos estatísticos e de aprendizado de máquina para extrair conhecimentos valiosos dos dados de saúde. E que a análise que preserve a privacidade busca adaptar esses métodos para manter a segurança dos dados, abrangendo desde a sumarização estatística, que ajuda a entender condições médicas, até aprendizados supervisionados e não supervisionados para diagnóstico de doenças e identificação de padrões em grandes volumes de dados. Que têm como objetivo desenvolver modelos analíticos que possam ser globais, como prever surtos de doenças, ou personalizados, adaptando-se às condições de saúde individuais.

A integração dos conceitos e abordagens discutidos por Banerjee (2017) e Sharma (2018) com as ideias apresentadas por Abidi & Abidi (2019), Aboudi & Benhlina (2018), e outros autores ressalta uma tendência crescente na área da saúde digital: a utilização de tecnologias avançadas, como a Internet das Coisas (IoT) e Big Data, para aprimorar a qualidade, eficiência e personalização dos cuidados de saúde.

Banerjee (2017) destaca a importância do controle de qualidade dos dados em sistemas IoT, especialmente em aplicações voltadas para a saúde, como o monitoramento da saúde dos idosos e sistemas de monitoramento residencial. A abordagem proposta de mapear o framework OSI (Open Systems Interconnection) para a qualidade do IoT e a ênfase na qualidade semântica dos dados ressoam com a discussão de Abidi & Abidi (2019) sobre a transição de sistemas de saúde baseados em volume para sistemas baseados em valor, onde o valor dos dados de

---

saúde é amplificado pela sua qualidade e aplicabilidade em inovações de tratamento e medicina personalizada.

Sharma (2018), por outro lado, aborda o desafio de desenvolver análises de saúde digital que preservem a privacidade dentro de sistemas baseados em IoT e nuvem, um ponto crucial também para Aboudi & Benhlina (2018), que defendem a importância de arquiteturas de dados flexíveis e extensíveis em saúde. A convergência entre a necessidade de análises que preservem a privacidade e a proposta de novas arquiteturas de dados sublinha a crescente complexidade dos sistemas de saúde digitais e a necessidade de soluções inovadoras que equilibrem eficiência, privacidade e qualidade dos dados.

Ambas as discussões, sobre controle de qualidade dos dados e análises que preservam a privacidade, complementam a visão de Abidi & Abidi (2019) sobre a crescente importância da IA e do Big Data na saúde. A convergência de Big Data e IA, conforme discutido, permite métodos de análise de dados de saúde inteligentes, que aprendem autonomamente associações não lineares e relações causais, essenciais para a geração de serviços de saúde inovadores orientados por dados.

A abordagem sistemática na exploração das capacidades de AI e Big Data em saúde, sugerida pela necessidade de uma seção explorando critérios e metodologias, ecoa no trabalho de Banerjee (2017) e Sharma (2018) pela necessidade de abordagens rigorosas no controle de qualidade dos dados e na garantia da privacidade. Essa abordagem metodológica rigorosa é fundamental para superar os desafios enfrentados na integração eficaz da IoT, Big Data e IA nos sistemas de saúde, visando não apenas a otimização dos recursos e a melhoria da qualidade dos cuidados, mas também a inovação contínua em saúde.

Portanto, a conexão entre esses trabalhos ilustra uma paisagem complexa e interconectada na saúde digital, onde a qualidade dos dados, a privacidade, a aplicação de tecnologias avançadas como IoT e Big Data, e o desenvolvimento de sistemas de saúde baseados em valor são temas centrais. Juntos, eles oferecem um caminho promissor para a transformação dos cuidados de saúde, destacando a importância de abordagens holísticas que integram tecnologia, metodologia e ética.

---

## **Principais Avanços Tecnológicos no Tratamento e Análise de Dados na Área de Saúde**

Um avanço significativo na análise de dados e Big Data em saúde é a utilização de grandes fluxos de informações de mídias sociais para o monitoramento de doenças. MUHUNZI (2023), explica que tweets relacionados ao HIV mostraram uma correlação positiva com os casos de HIV, enquanto padrões em mensagens nas redes sociais foram usados para prever e diagnosticar sinais de depressão. Durante a pandemia de COVID-19, o PNUD analisou tweets para monitorar a desinformação na Tanzânia. Esses casos ilustram como a análise de big data é crucial para a vigilância em tempo real de doenças e para entender a saúde pública em diferentes regiões.

Além disso, as ferramentas de Big Data podem organizar dados clínicos, testes laboratoriais e relatórios de pacientes em formas estruturadas, melhorando a recuperação e extração de dados e informando a tomada de decisões clínicas de maneira mais econômica (MUHUNZI, KITAMBALA, & MASHAURI, 2023).

Outro avanço, segundo BELLE (2015), é o fato do custo do sequenciamento do genoma humano está diminuindo rapidamente com novas tecnologias de alto rendimento, o que impacta políticas de saúde e cuidados clínicos. A análise desses dados genômicos para recomendações rápidas é um grande desafio. Iniciativas como a medicina P4 (preditiva, preventiva, participativa e personalizada) e o perfil integrativo de ômicas pessoais (iPOP) buscam resolver esses problemas. A medicina P4 analisa dados genômicos, desenvolve diagnósticos baseados em sangue e explora novos alvos de drogas. O iPOP combina monitoramento fisiológico com sequenciamento genômico para criar perfis detalhados de saúde.

## **Desafios na Coleta, Armazenamento e Interpretação de Grandes Volumes de Dados em Saúde**

BELLE (2015) aponta que no passado, os dados de dispositivos que capturam sinais fisiológicos continuamente raramente eram armazenados, e quando eram, utilizavam software proprietário dos fabricantes, dificultando o acesso e a interoperabilidade. E que, atualmente, embora os grandes fabricantes comecem a fornecer interfaces para dados ao vivo, surgem desafios típicos de Big Data, como

---

falta de protocolos, padrões e questões de privacidade. Além disso, limitações de largura de banda, escalabilidade e custo nos sistemas de saúde dificultam a adoção generalizada desses dados de streaming.

A armazenagem e recuperação de grandes volumes de dados exigem mecanismos sofisticados para facilitar a rápida obtenção e uso conforme as necessidades analíticas. No entanto, cada sistema de saúde tende a ter seus próprios esquemas de banco de dados, o que impede a interoperabilidade. A agregação de dados enfrenta desafios como a integração de fontes díspares, a padronização e a segurança, sendo essencial simplificar a complexidade dos dados médicos e garantir sua privacidade e proveniência.

Ao longo dos últimos anos muitas instituições passaram a usar registros médicos eletrônicos, que oferecem uma fonte substancial de dados médicos. Embora esses sistemas de registros eletrônicos de saúde (EHR) tenham sido inicialmente projetados para tarefas administrativas, eles frequentemente armazenam informações dos pacientes em vários formatos, às vezes junto com sistemas antigos. Os dados dos EHR frequentemente sofrem com informações faltantes, classificação incorreta e erros. Além disso, dados de alta dimensão, como telemetria, EEG, dados genéticos e dados fisiológicos contínuos fornecem insights clínicos valiosos, mas exigem grande capacidade de armazenamento e processamento. Dessa forma, Cohen e Kovacheva (2021) concluem que os obstáculos para incorporar a pesquisa em ciência de dados no cuidado ao paciente incluem baixa qualidade dos dados, recursos limitados e requisitos rigorosos de confidencialidade do paciente.

Reforçando os pontos discutidos anteriormente, MUHUNZI (2023) expõe que a aplicação de Big Data na saúde enfrenta dificuldades devido à fragmentação e diversidade dos dados, que vêm de várias fontes e formatos, necessitando de interpretação para obter conclusões significativas. A segurança dos dados, a privacidade e a confidencialidade dos pacientes são fundamentais, especialmente quando os dados precisam ser compartilhados entre diferentes partes interessadas e instituições. A falta de protocolos e estruturas de dados adequados torna o consentimento informado e a proteção da privacidade ainda mais importantes. Além

---

disso, os desafios de integração e interoperabilidade são obstáculos significativos para a análise eficaz de Big Data na área da saúde.

### **Implicações Éticas, Legais e Sociais do Uso de Big Data e Tecnologias Relacionadas**

O uso de Big Data e IA na assistência médica levanta várias questões éticas, legais e sociais. A privacidade dos dados dos pacientes é uma das principais preocupações, uma vez que a coleta e o armazenamento de grandes volumes de dados de saúde podem expor informações sensíveis a riscos de acesso não autorizado e violações de privacidade. Implementar medidas robustas de segurança de dados, como criptografia, controle de acesso e desidentificação, é crucial para mitigar esses riscos.

Além das preocupações com a privacidade, há implicações legais significativas associadas ao uso de Big Data e IA na saúde. A regulamentação dos dados de saúde varia amplamente entre diferentes países, complicando a integração e o compartilhamento de dados em contextos internacionais. A falta de normas claras e consistentes pode resultar em problemas de conformidade e litígios legais. Portanto, as instituições de saúde devem assegurar que suas práticas de coleta, armazenamento e análise de dados estejam em conformidade com as leis e regulamentos locais e internacionais para evitar penalidades legais e proteger os direitos dos pacientes.

O papel humano é essencial para manter a empatia e a confiança na relação médico-paciente, algo que sistemas de IA não conseguem substituir. Pacientes muitas vezes preferem receber informações médicas de um clínico empático em vez de um sistema de IA. O uso de máquinas inteligentes para decisões médicas levanta questões de responsabilidade, transparência, permissão e privacidade. Muitos algoritmos de IA são difíceis de interpretar e podem cometer erros no diagnóstico e tratamento, complicando a atribuição de responsabilidade. Além disso, esses sistemas podem apresentar viés algorítmico, prevendo maior probabilidade de doença com base em gênero ou raça sem que esses sejam fatores significativos (DAVENPORT & KALAKOTA, 2019).

---

O reuso de dados clínicos para pesquisa enfrenta desafios legais e éticos, buscando equilibrar a proteção da privacidade dos pacientes e a realização de pesquisas científicas. Embora na maioria dos países não seja necessário obter consentimento legal para coletar dados clínicos, os pacientes consentem indiretamente devido aos benefícios diretos para seus cuidados. É essencial proteger o direito à privacidade e o controle dos indivíduos sobre seus dados, garantindo que qualquer violação desses direitos seja mínima e menos intrusiva possível (MEYSTRE, 2017).

---

## Capítulo 4 - Considerações Finais

O setor da saúde é reconhecido como um dos mais cruciais para o bem-estar e desenvolvimento de qualquer sociedade. Com a emergência de novas tecnologias e a expansão exponencial de dados, torna-se imperativo explorar como esses avanços podem ser aproveitados para melhorar a assistência médica.

A capacidade de coletar, processar e analisar grandes volumes de dados pode fornecer insights valiosos para profissionais de saúde, permitindo diagnósticos mais precisos, tratamentos personalizados e uma abordagem preventiva mais eficaz.

No entanto, apesar do potencial revolucionário desses avanços, existem desafios significativos associados à sua implementação. A integração de sistemas, a proteção da privacidade do paciente, a capacitação de profissionais e as implicações éticas são apenas algumas das preocupações que surgem à medida que o mundo é movido em direção a uma era de medicina orientada por dados.

Para integrar eficazmente Big Data e IA no ecossistema de saúde, é essencial abordar desafios de infraestrutura, segurança e regulamentação. Recomenda-se desenvolver uma infraestrutura robusta de TI, incluindo soluções de computação em nuvem para armazenamento e gestão de grandes volumes de dados, permitindo que até pequenas organizações de saúde superem desafios de armazenamento e reduzam custos.

Além disso, é crucial criar uma cultura de compartilhamento de dados entre as partes interessadas e instituições de saúde para enfrentar os desafios de interoperabilidade e maximizar os benefícios das análises de Big Data. A troca de informações e a colaboração entre diferentes entidades podem melhorar significativamente os resultados clínicos. Assim como, é imprescindível implementar medidas de segurança de dados para proteger informações sensíveis e garantir conformidade com regulamentos de privacidade.

A má gestão dos dados de saúde e a falta de regulamentação adequada resultam em altos custos de infraestrutura. Implementar governança e políticas eficazes pode melhorar a utilização dos dados. Padronizar formatos de dados e

---

garantir a interoperabilidade entre sistemas de saúde é fundamental para integrar novas tecnologias de maneira eficiente, promovendo a troca segura e eficiente de informações de saúde e melhorando o atendimento e os resultados clínicos.

Dessa forma, a análise detalhada desses avanços e desafios é não apenas relevante, mas indispensável. Compreender as nuances dessas tecnologias e suas aplicações no contexto da saúde pode guiar políticas públicas, informar decisões clínicas e potencialmente transformar a assistência médica, tornando-a mais eficiente, eficaz e centrada no paciente.

Além disso, se considerada a crescente demanda por serviços de saúde de qualidade e a necessidade contínua de inovação no setor, este estudo se justifica pela contribuição significativa que pode oferecer a acadêmicos, profissionais de saúde, tomadores de decisão e, em última análise, aos pacientes que são o núcleo da prática médica.

Ao destacar os avanços, benefícios e desafios no tratamento de dados em saúde, este trabalho tem o potencial de lançar luz sobre o caminho a seguir, aproveitando ao máximo as oportunidades oferecidas pela tecnologia, ao mesmo tempo, em que aborda suas complexidades inerentes.

Porém, como em toda pesquisa, nem tudo está atrelado aos benefícios, existe uma lacuna muito grande no que se diz respeito à utilização de tecnologias dentro área. As instituições de saúde e pesquisa precisam se unir e formar parcerias com instituições acadêmicas e empresas de tecnologia a fim de promover a aplicação prática das descobertas, bem como investir mais em infraestrutura de dados.

Também é de suma importância que os profissionais de saúde sejam treinados não apenas nas suas especialidades médicas, mas também em como utilizar, analisar e interpretar dados.

Além disso, com o aumento da coleta de dados, é imprescindível abordar questões de ética e de privacidade. Os sistemas de saúde devem garantir que os dados dos pacientes sejam coletados e armazenados de forma segura.

---

É importante reiterar que a aplicação prática de Big Data e IA no setor de saúde tem o potencial de revolucionar os cuidados ao paciente, mas deve ser desenvolvida com cautela e uma sólida compreensão da ciência de dados.

Esta revisão apresenta algumas limitações importantes. Ela se baseia exclusivamente em trabalhos publicados em revistas científicas e se concentra especificamente em uma seleção de estudos relacionados ao uso de Big Data e Inteligência Artificial (IA) na área da saúde. E seu objetivo não foi realizar uma análise exaustiva de toda a literatura disponível nesse campo. Há uma vasta gama de tópicos adicionais que merecem ser explorados em pesquisas futuras.

Como sugestões para estudos futuros, é essencial investigar mais a fundo a integração dessas tecnologias nos sistemas de saúde existentes, abordando desafios práticos como interoperabilidade, segurança da informação e proteção da privacidade do paciente. Isso pode envolver a colaboração entre pesquisadores, profissionais de saúde e especialistas em tecnologia para desenvolver padrões e protocolos que facilitem a troca segura e eficiente de dados entre sistemas e dispositivos.

Outra área importante de pesquisa é a avaliação do impacto dessas tecnologias na qualidade dos cuidados de saúde e na eficiência operacional dos sistemas de saúde. Estudos longitudinais podem ser conduzidos para analisar os resultados clínicos, a satisfação do paciente e os custos associados à implementação de novas tecnologias, fornecendo evidências empíricas para informar políticas e práticas futuras.

No contexto brasileiro, a implementação de sistemas baseados em Big Data e Inteligência Artificial (IA) no Sistema Único de Saúde (SUS) representa uma área promissora para estudos futuros. Embora este trabalho tenha se concentrado em aspectos gerais, é fundamental reconhecer a relevância do tema para o SUS, que desempenha um papel crucial no atendimento a uma grande parcela da população do país.

A aplicação dessas tecnologias no SUS pode contribuir significativamente para aprimorar a eficiência, qualidade e acessibilidade dos serviços de saúde oferecidos à população brasileira. Assim, estudos futuros podem se concentrar em explorar as melhores práticas para a implementação desses sistemas no contexto

---

do SUS, considerando os desafios específicos enfrentados pelo sistema de saúde pública do Brasil, como a diversidade cultural, as diferenças regionais e a limitação de recursos.

Além disso, é importante investigar como garantir que essas tecnologias sejam acessíveis e equitativas, de modo a beneficiar todas as camadas da sociedade brasileira, incluindo aquelas em situação de vulnerabilidade socioeconômica

Por fim, é importante continuar a explorar questões éticas e de privacidade relacionadas à coleta, armazenamento e uso de dados de saúde. Isso inclui garantir que os pacientes tenham controle sobre seus próprios dados e que os sistemas de saúde implementem medidas robustas de segurança

---

# Referências Bibliográficas

ABIDI, S. & ABIDI, S. (2019). Intelligent health data analytics: a convergence of artificial intelligence and Big Data. *Healthcare Management Forum*, 32(4), 178-182. <https://doi.org/10.1177/0840470419846134>

ABOUDI, N. & BENHLIMA, L. (2018). Big data management for healthcare systems: architecture, requirements, and implementation. *Advances in Bioinformatics*, 2018, 1-10. <https://doi.org/10.1155/2018/4059018>

ARIMA, H. Utilizing Big Data for Public Health. *\*J Epidemiol\**, v. 26, n. 3, p. 105, 2016. <https://doi.org/10.2188/jea.JE20160036>.

BANERJEE, T.; SHETH, A. IoT Quality Control for Data and Application Needs. *\*IEEE Intelligent Systems\**, v. 32, n. 2, p. 68-73, mar.-abr. 2017. <https://doi.org/10.1109/MIS.2017.35>.

BEAM, A. & KOHANE, I. (2018). Big data and machine learning in health care. *Jama*, 319(13), 1317. <https://doi.org/10.1001/jama.2017.18391>

BELLE, A., THIAGARAJAN, R., SOROUSHMEHR, S., NAVIDI, F., BEARD, D., & NAJARIAN, K. (2015). Big data analytics in healthcare. *Biomed Research International*, 2015, 1-16. <https://doi.org/10.1155/2015/370194>

BENNIS, FRANCISCUS, et al. (2020). Machine learning in medicine. <https://doi.org/10.26481/dis.20201113fb>

BOSWORTH, A. et al. An inventor's perspective on consumer health informatics. *\*Am J Prev Med\**, v. 40, n. 5 Suppl 2, p. S241-4, mai. 2011. <https://doi.org/10.1016/j.amepre.2011.02.004>.

---

CHOI, J. Utilization value of medical Big Data created in operation of medical information system. \*The Journal of the Korea Institute of Electronic Communication Sciences\*, v. 10, n. 12, p. 1403-1410, 2015. <https://doi.org/10.13067/JKIECS.2015.10.12.1403>

Cohen, R. and Kovacheva, V. (2021). A methodology for a scalable, collaborative, and resource-efficient platform to facilitate healthcare ai research.. <https://doi.org/10.48550/arxiv.2112.06883>

DAVENPORT, T. & KALAKOTA, R. (2019). The potential for artificial intelligence in healthcare. *Future Healthcare Journal*, 6(2), 94-98. <https://doi.org/10.7861/futurehosp.6-2-94>

GOLI-MALEKABADI, Z.; SARGOLZAEI-JAVAN, M.; AKBARI, M. K. An effective model for store and retrieve big health data in cloud computing. \*Comput Methods Programs Biomed\*, v. 132, p. 75-82, ago. 2016. <https://doi.org/10.1016/j.cmpb.2016.04.016>.

IZAD SHENAS, S. A. et al. Identifying high-cost patients using data mining techniques and a small set of non-trivial attributes. \*Comput Biol Med\*, v. 53, p. 9-18, out. 2014. <https://doi.org/10.1016/j.compbimed.2014.07.005>

JEFFS, L. et al. Building the foundation to generate a fundamental care standardised data set. \*Journal\*, [s.d.]. <https://doi.org/10.1111/jocn.14308>

JIANG, G. et al. Feasibility of capturing real-world data from health information technology systems at multiple centers to assess cardiac ablation device outcomes: A fit-for-purpose informatics analysis report. \*Journal\*, [s.d.]. DOI:10.1093/jamia/ocab117

Kriksciuniene, D. and Sakalauskas, V. (2022). Overview of the artificial intelligence methods and analysis of their application potential., 167-183. [https://doi.org/10.1007/978-3-030-79353-1\\_9](https://doi.org/10.1007/978-3-030-79353-1_9)

---

LEYENS, L. et al. Use of Big Data for drug development and for public and personal health and care. \*Journal\*, [s.d.]. <https://doi.org/10.1002/gepi.22012>.

Li, X., Zhang, Y., & Tian, Y. (2016). Medical Big Data analysis in hospital information system.. <https://doi.org/10.5772/63754>

MCPADDEN, J. et al. A Scalable Data Science Platform for Healthcare and Precision Medicine Research. \*Journal\*, [s.d.].

MEYSTRE, S. M.; et al. Clinical Data Reuse or Secondary Use: Current Status and Potential Future Progress. \*Yearb Med Inform\*, v. 26, n. 01, p. 38-52, 2017. DOI: 10.15265/IY-2017-007.

MUHUNZI, D., KITAMBALA, L., & MASHAURI, H. (2023). Big data analytics in the healthcare sector: opportunities and challenges in developing countries. a literature review.. <https://doi.org/10.21203/rs.3.rs-2869049/v1>

NAZIR, S., KHAN, S., KHAN, H., ALI, S., GARCÍA-MAGARIÑO, I., ATAN, R., ... & NAWAZ, M. (2020). A comprehensive analysis of healthcare Big Data management, analytics and scientific programming. *IEEE Access*, 8, 95714-95733. <https://doi.org/10.1109/access.2020.2995572>

NIU, W. et al. Knowledge Spillovers of Medical Big Data Under Hierarchical Medical System and Patients' Medical Treatment Decisions. \*IEEE Access\*, [s.d.], 2019. <https://doi.org/10.1109/ACCESS.2019.2908440>.

PEREIRA, L. et al. Text Mining Applied to Electronic Medical Records: A Literature Review. \*International Journal of E-Health and Medical Communications (IJEHMC)\*, v. 6, n. 3, p. [início-final], 2015. <https://doi.org/10.4018/IJEHMC.2015070101>.

---

RIOUX, C.; LITTLE, T. D. Missing data treatments in intervention studies: What was, what is, and what should be. *\*Journal\**, v. 45, n. 1, p. [início-final], 15 out. 2019. <https://doi.org/10.1177/016502541988060>.

SHARMA, S., CHEN, K., & SHETH, A. (2018). Toward practical privacy-preserving analytics for iot and cloud-based healthcare systems. *Ieee Internet Computing*, 22(2), 42-51. <https://doi.org/10.1109/mic.2018.112102519>

TAYLOR, J. A. et al. The road to hell is paved with good intentions: the experience of applying for national data for linkage and suggestions for improvement. *\*BMJ Open\**, v. 11, [s.d.]. <https://doi.org/10.1136/bmjopen-2020-047575>.

YAHAV, D.; TAU, N.; SHEPSHELOVICH, D. Assessment of Data Supporting the Efficacy of New Antibiotics for Treating Infections Caused by Multidrug-resistant Bacteria. *\*Clin Infect Dis\**, v. 72, n. 11, p. 1968-1974, 1 jun. 2021. <https://doi.org/10.1093/cid/ciaa457>.

ZHANG, Y. et al. Health-CPS: Healthcare Cyber-Physical System Assisted by Cloud and Big Data. *\*IEEE Systems Journal\**, v. 11, n. 1, p. 88-95, mar. 2017. <https://doi.org/10.1109/JSYST.2015.2460747>.