

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ESTRATÉGIA
MESTRADO PROFISSIONAL EM GESTÃO E ESTRATÉGIA-MPGE**

DISSERTAÇÃO

**PLANO ESTRATÉGICO PARA A ADOÇÃO DO BIG DATA NA SAÚDE PÚBLICA:
UM ESTUDO APLICADO NO INSTITUTO NACIONAL DE CÂNCER**

**STRATEGIC PLAN FOR THE ADOPTION OF BIG DATA IN PUBLIC HEALTH: AN
APPLIED STUDY AT THE NATIONAL CANCER INSTITUTE**

LUANA DA SILVA ALVES

2024



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ESTRATÉGIA
MESTRADO PROFISSIONAL EM GESTÃO E ESTRATÉGIA-MPGE**

**PLANO ESTRATÉGICO PARA A ADOÇÃO DO BIG DATA NA SAÚDE PÚBLICA:
UM ESTUDO APLICADO NO INSTITUTO NACIONAL DE CÂNCER**

**STRATEGIC PLAN FOR THE ADOPTION OF BIG DATA IN PUBLIC HEALTH: AN
APPLIED STUDY AT THE NATIONAL CANCER INSTITUTE**

LUANA DA SILVA ALVES

*Sob a Orientação do Professor
Saulo Barbará de Oliveira*

*Sob a Coorientação do Professor
Sandro Luís Freire de Castro Silva*

Dissertação submetida como requisito parcial para a obtenção do grau de **Mestre**, no Curso de Mestrado Profissional em Gestão e Estratégia do Programa de Pós-Graduação em Gestão e Estratégia da Universidade Federal Rural do Rio de Janeiro - UFRRJ.

Seropédica, RJ
Setembro de 2024

Universidade Federal Rural do Rio de Janeiro
Biblioteca Central / Seção de Processamento Técnico

Ficha catalográfica elaborada
com os dados fornecidos pelo(a) autor(a)

A474p

Alves, Luana da Silva, 1991-
PLANO ESTRATÉGICO PARA A ADOÇÃO DO BIG DATA NA SAÚDE
PÚBLICA: UM ESTUDO APLICADO NO INSTITUTO NACIONAL DE
CÂNCER / Luana da Silva Alves. - Rio de Janeiro, 2024.
180 f.

Orientador: Saulo Barbará de Oliveira.
Coorientador: Sandro Luís Freire de Castro Silva.
Dissertação (Mestrado). -- Universidade Federal
Rural do Rio de Janeiro, Programa de Pós-Graduação em
Gestão e Estratégia da Universidade Federal Rural do
Rio de Janeiro - UFRRJ, 2024.

1. Big Data na Saúde . 2. Saúde Digital. 3. Plano
Estratégico. 4. INCA. 5. Saúde Pública. I. Barbará de
Oliveira, Saulo, 1948-, orient. II. Luís Freire de
Castro Silva, Sandro, 1986-, coorient. III
Universidade Federal Rural do Rio de Janeiro.
Programa de Pós-Graduação em Gestão e Estratégia da
Universidade Federal Rural do Rio de Janeiro - UFRRJ.
IV. Título.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível
Superior – Brasil (CAPES) – Código de Financiamento 001.

**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ESTRATÉGIA**

LUANA DA SILVA ALVES

Dissertação submetida como requisito parcial para a obtenção do grau de Mestre(a), no Programa de Pós-Graduação em Gestão e Estratégia, Área de Concentração em Gestão e Estratégia.

DISSERTAÇÃO APROVADA EM 09/09/2024.

Prof(a). Dr(a). Sandro Luis Freire de Castro Silva
Presidente da Banca/Coorientador(a)
Membro Interno
UFRRJ

Documento assinado digitalmente
 BEATRIZ CORDEIRO JARDIM
Data: 11/09/2024 17:08:39-0300
Verifique em <https://validar.itd.gov.br>

Dr(a). Beatriz Cordeiro Jardim
Membro Externo
INCA

Documento assinado digitalmente
 MARCELO FORNAZIN
Data: 20/09/2024 14:05:43-0300
Verifique em <https://validar.itd.gov.br>

Prof(a). Dr(a). Marcelo Fornazin
Membro Externo
UFF

TERMO N° 787/2024 - PPGE (12.28.01.00.00.00.05)

(Nº do Protocolo: **NÃO PROTOCOLADO**)

(Assinado digitalmente em 20/09/2024 14:16)

SANDRO LUIS FREIRE DE CASTRO SILVA

ASSINANTE EXTERNO

CPF: ####.###.807-##

Visualize o documento original em <https://sipac.ufrrj.br/documentos/> informando seu número: 787, ano: 2024, tipo:
TERMO, data de emissão: 20/09/2024 e o código de verificação: **1a2760df42**

DEDICATÓRIA

Dedico este estudo à constante correnteza do conhecimento, que nos conduz por terras desconhecidas e nos desafia a explorar novas margens. Que ele seja uma bússola para aqueles que navegam as águas turbulentas da descoberta, guiando-os rumo a horizontes ainda não mapeados.

AGRADECIMENTOS

Não há palavras que possam expressar o quanto sou grata à minha família, em especial Vânia Cruz, Letícia Torres, Larissa Torres, Luma Torres, Julia Cruz e Caroline Alves, pelo apoio constante ao longo desta jornada desafiadora. Seu amor e confiança foram fundamentais para a minha persistência e jamais esquecerei o suporte que me ofereceram.

Ao meu pai, José Francisco Vaz Torres, e à minha avó, Norma Vaz Torres, que do céu observam cada passo que dou, dedico o mais profundo agradecimento. Mesmo que estejam fisicamente ausentes, suas presenças permanecem vivas em meu coração, e seus ensinamentos continuam a guiar-me. Agradeço por terem sido exemplos de resiliência, força e amor. Que o legado de vocês perdure como uma luz em meu caminho, inspirando-me em todas as etapas da vida.

À Caroline Alves, minha parceira de vida, agradeço por sua presença constante, paciência e apoio ao longo deste processo. Suas palavras gentis, atenção sincera e risadas espontâneas foram um alívio bem-vindo nos momentos difíceis. Compartilhar a vida contigo é um privilégio que valorizo profundamente. Estendo este agradecimento às minhas cachorrinhas, Amora e Molly, por sempre me lembrarem da importância de valorizar as coisas simples da vida.

Aos meus orientadores, Prof. Dr. Saulo Barbará e Prof. Dr. Sandro Luís Freire, expresso minha sincera gratidão pela confiança desde o início, pelo incentivo, apoio e orientação ao longo de todo este percurso. Agradeço pela paciência e pelo cuidado demonstrado durante esta caminhada. A vocês, todo o meu respeito e admiração pela dedicação e compromisso com o meu crescimento acadêmico e profissional. Levarei para sempre comigo os valiosos ensinamentos recebidos e os aplicarei em todas as áreas da minha vida. Obrigada pelo impacto positivo que tiveram em minha jornada.

À minha turma do mestrado, representada por Elisy Felício, Daiane Luz, Ana Lívia, Kelly Gonçalves, Leonardo Menezes, Edson Lima, Wallace Silva e Anderson Almeida, agradeço pela parceria, apoio mútuo e momentos compartilhados ao longo deste caminho desconhecido e repleto de desafios. Foi uma honra percorrer esta trajetória ao lado de vocês.

Aos professores do programa de mestrado em Gestão e Estratégia da Universidade Federal Rural do Rio de Janeiro, meu sincero agradecimento pelo conhecimento transmitido e pela dedicação demonstrada em nossa formação. Levo comigo as lições valiosas que cada um de vocês compartilhou.

Ao Instituto Nacional de Câncer (INCA), expresso minha sincera gratidão pela confiança depositada neste estudo. Espero verdadeiramente que os conhecimentos compartilhados possam agregar valor à instituição. Aos profissionais do INCA, meu respeito, admiração e profundo agradecimento pelo acolhimento, disponibilidade e valiosa colaboração

demonstrados ao longo desta dissertação.

Por fim, agradeço a todos aqueles que, de alguma forma, contribuíram para a realização deste trabalho, seja através de palavras de incentivo, apoio emocional, ou simples gestos de amizade. A todos vocês, o meu muito obrigado.

Epígrafe

“Ninguém pode entrar duas vezes no mesmo rio, pois quando nele se entra novamente, não se encontra as mesmas águas e o próprio ser já se modificou.”

Heráclito de Éfeso

RESUMO EXPANDIDO

Objetivo – O presente estudo teve como objetivo final propor a elaboração de um plano estratégico com recomendações práticas para a implementação do *big data* no Instituto Nacional de Câncer (INCA). Para alcançar este propósito, foram definidos e perseguidos os seguintes objetivos intermediários: a) descrever, com base na literatura, a relação entre tecnologia e o processo de evolução tecnológico na área da saúde; b) identificar os impactos no setor de saúde ocasionados pelo aumento do volume exponencial de dados e a contribuição do *big data* para melhoria da gestão e; c) desenvolver um plano estratégico com recomendações práticas que viabilizem a implementação do *big data* no INCA, por meio de um estudo aplicado.

Metodologia – Como procedimentos metodológicos foi realizada uma pesquisa aplicada de abordagem qualitativa no Instituto Nacional de Câncer (INCA). Para a coleta de dados, recorreu-se a uma pesquisa documental que precedeu a realização de entrevistas feitas usando um roteiro semiestruturado, tendo como sujeitos profissionais da área da saúde e gestores da área de Tecnologia da Informação do INCA. O levantamento da literatura foi feito usando a técnica de mapeamento sistemático da literatura. Para análise dos dados foi utilizada a técnica de análise de conteúdo, visando inicialmente transcrever os documentos e as entrevistas realizadas a fim de obter as primeiras impressões dos entrevistados sobre o tema. Em seguida, o material coletado foi explorado, com o intuito de descrever, a partir da fala dos sujeitos entrevistados, as suas impressões sobre uso de tecnologias como o *big data* no Instituto. Por fim, com base na frequência de repetição dos elementos, os dados foram categorizados para que, a partir disso, fosse possível tratar os dados brutos obtidos. A partir dos dados encontrados, elaborou-se um plano estratégico com recomendações práticas para a implementação do *big data* no INCA.

Resultados – Os resultados do mapeamento sistemático da literatura permitiram perceber que, no Brasil, os desafios para implementação do *Big Data* na saúde incluem o grau de complexidade da análise e armazenamento dos dados em tempo real, a necessidade de aquisição de tecnologias robustas por parte das instituições de saúde brasileira, a importância do fortalecimento da integração de dados entre as instituições de saúde nacional e do fortalecimento do processo de digitalização dos prontuário médicos, majoritariamente nas instituições públicas de saúde. Além disso, a pandemia da Covid-19, potencializou o uso de tecnologias no setor de saúde e aumentou, significativamente, o uso dos dados para nortear as decisões estratégicas das instituições de saúde. Também foi possível observar a importância da padronização dos registros de informações dos pacientes nos prontuários eletrônicos e a necessidade de uma mudança de perspectiva por parte das instituições de saúde, sobre o investimento para adoção de tecnologias no setor, que ainda é visto como um custo e não como uma estratégia para prover a melhoria na qualidade dos serviços de saúde prestados à população. Com base na análise dos dados coletados em campo, foi possível identificar que, especificamente no

contexto do INCA, o *big data* se apresenta como uma ferramenta de valor capaz de aprimorar a gestão de dados oncológicos, promover avanços significativos na pesquisa, diagnóstico e tratamento do câncer, além de apresentar potencial para aumentar a eficiência administrativa e operacional da instituição. O uso integrado das tecnologias de *big data*, não apenas pode aprimorar a tomada de decisões clínicas e estratégicas do Instituto, mas também melhorar a qualidade dos serviços de saúde prestado pelo INCA ao paciente com câncer.

Limitações do Método de Pesquisa – Como é próprio do processo de seleção, ao se escolher determinados métodos ou técnicas de pesquisa, reconhece-se que os métodos de pesquisa selecionados, pela sua própria natureza, possuem limitações intrínsecas na sua aplicação, tanto na fase de coleta, quanto na fase de análise de dados. Entretanto, cuidados foram tomados visando superar as limitações do método, como descrito a seguir. A pesquisa documental, que envolveu relatórios de gestão, planejamento estratégico de saúde digital, plano diretor de TI, dentre outros documentos, teve como limitação a origem dos dados, uma vez que são de fontes secundárias, ou seja, não geradas durante a pesquisa. Tal limitação foram superadas a partir da averiguação da confiabilidade das fontes responsáveis pelo seu fornecimento.

A entrevista individual com roteiro semiestruturado apresentou limitações, como a influência das relações trabalhistas em suas respostas e a inteligibilidade das perguntas realizadas e das respostas dadas. Para atenuar este risco, foi realizado o piloto de pesquisa com um grupo de servidores do INCA, para corrigir pontos dúbios e validar o roteiro de entrevista, com a pactuação do Termo de Consentimento Livre e Esclarecido.

As análises indutivas e interpretativas enfrentaram a limitação de uma possível interpretação inadequada das percepções dos participantes pela pesquisadora. No entanto, essa questão foi mitigada por meio de uma categorização cuidadosa dos dados empíricos. Por outro lado, não foi possível prolongar a etapa de coleta, uma vez que os profissionais estavam sobrecarregados com as demandas de suas funções. Porém, isso não resultou a prejuízo para a pesquisa.

Aplicabilidade do trabalho e contribuições – A iniciativa de elaboração de um plano de ação com propostas de melhorias estratégicas para a implementação do *big data* no INCA, levou em consideração o Planejamento Estratégico elaborado pela gestão do INCA, para o período de 2020-2023, em conformidade com o Plano Nacional de Saúde (PNS), tomado como ponto de partida para a elaboração estratégica do Instituto. Além disso, o fomento à produção do conhecimento científico conta com o incentivo e apoio dos gestores do setor de Tecnologia da Informação do INCA, setor sob o qual foi realizada a pesquisa. Desta forma, espera-se que, por meio deste estudo, o INCA tenha uma visão sistematizada das ações adotadas pelas instituições de saúde, que implementaram o *Big Data* na prática em saúde, para identificar as possibilidades do uso do *Big Data* e seus impactos nos serviços no campo da oncologia.

Originalidade – Tomando como base os estudos analisados ao longo desta pesquisa, percebeu-se que as instituições de saúde orientadas por dados aplicaram terapias mais assertivas, controlaram seus ativos com maior eficiência e beneficiaram os pacientes com tratamentos personalizados, fatores que melhoraram a qualidade do tratamento e aumentaram as chances de sobrevivência dos pacientes. Desta forma, pressupõem-se que, por se tratar de uma instituição pública de saúde, o INCA não está alheio a esta realidade. Neste sentido, tecnologias como o *big data* se apresentam como uma ferramenta estratégica de valor para o Instituto alcançar seus objetivos estratégicos a médio e longo prazo.

Palavras-chaves - Big Data; Saúde Pública; Saúde Digital; Plano Estratégico; INCA.

EXPANDED ABSTRACT

Objective – The ultimate objective of this study was to propose the development of a strategic plan with practical recommendations for implementing big data at the National Cancer Institute (INCA). To achieve this goal, the following intermediate objectives were defined and pursued: a) describe, based on the literature, the relationship between technology and the process of technological evolution in the healthcare field; b) identify the impacts on the healthcare sector caused by the exponential increase in data volume and the contribution of big data to improving management; and c) develop a strategic plan with practical recommendations that enable the implementation of big data at INCA through an applied study.

Methodology – The methodological procedures involved an applied qualitative research approach conducted at the National Cancer Institute (INCA). Data collection included a documentary research phase that preceded interviews conducted using a semi-structured script, involving healthcare professionals and IT managers at INCA. The literature review was conducted using the systematic literature mapping technique. Data analysis was carried out using content analysis, initially aimed at transcribing documents and interviews to obtain the first impressions of the interviewees on the topic. The collected material was then explored to describe the interviewees' impressions regarding the use of technologies such as big data at the Institute. Finally, based on the frequency of recurring elements, the data was categorized to enable the processing of the raw data obtained. From the data collected, a strategic plan with practical recommendations for the implementation of big data at INCA was developed.

Results – The results of the systematic literature mapping revealed that, in Brazil, the challenges for implementing big data in healthcare include the complexity of real-time data analysis and storage, the need for Brazilian healthcare institutions to acquire robust technologies, the importance of strengthening data integration among national healthcare institutions, and the enhancement of the digitization process of medical records, particularly in public healthcare institutions. Furthermore, the COVID-19 pandemic has significantly increased the use of technologies in the healthcare sector, driving the use of data to guide strategic decisions within healthcare institutions. It was also observed that the standardization of patient information in electronic medical records is crucial, as is a shift in perspective within healthcare institutions regarding technology adoption investments, which are still seen as a cost rather than a strategy to improve the quality of healthcare services provided to the population. Based on the analysis of field-collected data, it was identified that, specifically in the context of INCA, big data is a valuable tool capable of enhancing oncology data management, promoting significant advances in cancer research, diagnosis, and treatment, and potentially increasing the administrative and operational efficiency of the institution. The integrated use of big

data technologies can not only improve the Institute's clinical and strategic decision-making but also enhance the quality of healthcare services provided by INCA to cancer patients.

Research Method Limitations – As is typical in the selection process, choosing specific research methods or techniques acknowledges that the selected methods inherently possess limitations in their application, both in the data collection and analysis phases. However, steps were taken to overcome these limitations, as described below. Documentary research, which involved management reports, digital health strategic planning, IT master plans, and other documents, had the limitation of data origin, as they were secondary sources, not generated during the research. This limitation was addressed by verifying the reliability of the sources providing this data. The individual interview using a semi-structured script presented limitations, such as the influence of work relationships on responses and the intelligibility of the questions asked and the answers given. To mitigate this risk, a research pilot was conducted with a group of INCA staff to correct ambiguous points and validate the interview script, with the agreement of the Free and Informed Consent Term. Inductive and interpretive analyses faced the limitation of possible misinterpretation of participants' perceptions by the researcher. However, this issue was mitigated by careful categorization of the empirical data. On the other hand, it was not possible to extend the data collection phase, as professionals were overwhelmed with their job demands. Nevertheless, this did not result in any harm to the research.

Applicability and Contributions – The initiative to develop an action plan with strategic improvement proposals for the implementation of big data at INCA took into account the Strategic Planning developed by INCA's management for the 2020-2023 period, in accordance with the National Health Plan (PNS), which served as a starting point for the Institute's strategic planning. Additionally, the promotion of scientific knowledge production is supported and encouraged by INCA's IT management, under which the research was conducted. Thus, it is expected that, through this study, INCA will gain a systematized view of the actions adopted by healthcare institutions that have implemented big data in health practice, to identify the possibilities of using big data and its impacts on oncology services.

Originality – Based on the studies analyzed throughout this research, it was observed that data-driven healthcare institutions applied more assertive therapies, managed their assets more efficiently, and benefited patients with personalized treatments, factors that improved treatment quality and increased patients' chances of survival. Thus, it is assumed that, as a public healthcare institution, INCA is not isolated from this reality. In this sense, technologies like big data present themselves as a valuable strategic tool for the Institute to achieve its medium and long-term strategic objectives.

Keywords – Big Data; Public Health; Digital Health; Strategic Plan; INCA.

RESUMO

ALVES, S. L. Plano estratégico para a adoção do *big data* na saúde pública: um estudo aplicado no Instituto Nacional de Câncer. 2024. 168f. Dissertação (Mestrado Profissional em Gestão e Estratégia) – Universidade Federal Rural do Rio de Janeiro, Programa de Pós-Graduação em Gestão e Estratégia, Seropédica-RJ, 2024.

O crescente uso dos recursos do *big data* nos diversos setores da sociedade impõe a necessidade de se compreender as vantagens e benefícios desta tecnologia, especialmente no setor de saúde, onde a vida humana é o bem mais valioso. Apesar dos avanços, parte das instituições de saúde ainda não conhecem o potencial de tecnologias mais recentes como o *big data*, para gerar *insights* e auxiliar na tomada de decisões estratégicas por parte da equipe gerencial, médica e assistencial. As informações extraídas do *big data* podem gerar impactos positivos para a redução de custos operacionais, otimização da análise e diagnóstico de doenças em estágios iniciais e viabilidade de tratamentos personalizados para os pacientes portadores de doenças complexas, de forma dinâmica e eficiente, impactando diretamente a qualidade dos serviços de saúde. Tendo em vista que as tecnologias são predominantemente desenvolvidas por países de alto poder aquisitivo e com uma perspectiva de saúde diferente da do Brasil, este trabalho tem por objetivo avaliar que caminhos podem ser adotados por uma instituição pública de saúde brasileira para implementação e uso do *big data* em suas decisões estratégicas. Como procedimentos metodológicos foi realizada uma pesquisa aplicada de abordagem qualitativa no Instituto Nacional de Câncer (INCA). Para a coleta de dados, recorreu-se a uma pesquisa documental que precedeu a realização de entrevistas feitas usando um roteiro semiestruturado com os profissionais da área da saúde e gestores da área de Tecnologia da Informação do INCA. Para análise dos dados foi utilizada a técnica de análise de conteúdo. Quanto ao produto tecnológico elaborado nesta dissertação, foi elaborado um relatório técnico conclusivo com recomendações estratégicas para implementação do *big data* no INCA.

Palavras-chave: Big Data; Saúde Pública; Saúde Digital; Plano Estratégico; INCA.

ABSTRACT

ALVES, S. L. **Strategic plan for the adoption of *big data* in public health:** an applied study at the National Cancer Institute. 2024. 168f. Dissertation (Professional Master's in Management and Strategy) - Federal Rural University of Rio de Janeiro, Graduate Program in Management and Strategy, Seropédica-RJ, 2024.

The growing use of big data resources in various sectors of society necessitates an understanding of the advantages and benefits of this technology, particularly in the healthcare sector, where human life is the most valuable asset. Despite the advancements, some healthcare institutions are still unaware of the potential of more recent technologies, such as big data, to generate insights and assist in strategic decision-making by management, medical, and care teams. The information extracted from big data can have a positive impact on reducing operational costs, optimizing the analysis and diagnosis of diseases at early stages, and enabling personalized treatments for patients with complex conditions in a dynamic and efficient manner—key factors for providing quality healthcare services to the population. Considering that these technologies are predominantly developed by countries with high purchasing power and a different healthcare perspective than Brazil, this study aims to evaluate the paths that a Brazilian public healthcare institution can adopt for the implementation and use of big data in its strategic decisions. Methodologically, an applied research with a qualitative approach was conducted at the National Cancer Institute (INCA). Data collection involved documentary research followed by interviews conducted using a semi-structured script with healthcare professionals and IT managers at INCA. Content analysis was used to analyze the data. As for the technological product developed in this dissertation, a conclusive technical report with strategic recommendations for implementing big data at INCA was prepared.

Keywords: Big Data; Public Health; Digital Health; Strategic Plan; INCA.

LISTA DE QUADROS

Quadro 1: Características do big data com base na coleção de V's.....	25
Quadro 2: Quadro-síntese dos conceitos de big data	27
Quadro 3: Apresentação da versão inicial do roteiro de entrevista semiestruturado	59
Quadro 4: Adequações sugeridas pelos avaliadores do plano piloto	60
Quadro 5: Apresentação do roteiro de entrevista semiestruturado após a aplicação do plano piloto	62
Quadro 6: Categorização dos resultados	66
Quadro 7: Categorias principais utilizadas na etapa de análise dos resultados	70
Quadro 8: Categorias emergentes que complementam a análise dos resultados	73
Quadro 9: Percepção dos entrevistados sobre os principais desafios que podem ser enfrentados pelo INCA para garantir a qualidade dos dados de saúde gerenciados pela instituição	82
Quadro 10: Caminhos que podem ser adotados pelo INCA para viabilizar a integração e interoperabilidade dos sistemas de saúde ativos na instituição	85
Quadro 11: Síntese sobre a Aplicabilidade dos Insights Extraídos do Big Data no INCA	90
Quadro 12: Boas práticas para o uso ético do big data no INCA	95
Quadro 13: Detalhamento do Produto Tecnológico	100
Quadro 14: Riscos Identificados e suas Descrições	105
Quadro 15: Frequência dos Riscos Identificados	110
Quadro 16: Impacto Potencial dos Riscos Identificados	115
Quadro 17: Riscos de Baixo Impacto Identificados	120
Quadro 18: Estratégias de Resposta aos Riscos Identificados para a Implementação do Big Data no INCA	125
Quadro 19: Critérios do PICOC adotados	130
Quadro 20: String de busca final	135
Quadro 21: Critérios de inclusão e exclusão	140
Quadro 22: Lista de artigos selecionados para a composição da Fundamentação Teórica	145

LISTA DE FIGURAS

Figura 1: Visão geral do percurso metodológico da pesquisa	17
Figura 2: Visualização da nuvem de tags estática de termos-chave que aparecem em resumos de artigos relacionados a saúde digital	22
Figura 3: Estrutura para aplicação do big data analytics na saúde	34
Figura 4: Objetivos estratégicos do INCA (2020-2023)	38
Figura 5: Pilares estratégicos do INCA (2020-2023)	45
Figura 6: Diagrama metodológico do Estudo de Caso	45
Figura 7: Visão geral do processo de seleção dos artigos	61
Figura 8: Análise de Conteúdo segundo Bardin (2016)	48
Figura 9: Distribuição das unidades do INCA	52
Figura 10: Objetivos estratégicos do INCA (2020-2023)	55
Figura 11: Proposta para implementação do framework de melhores práticas de governança de dados no INCA	60
Figura 12: Ciclo PDCA para manutenção da qualidade dos dados de saúde no INCA	62
Figura 13: Fluxo de big data no INCA	96
Figura 14: Etapas para implementação do big data no INCA	112
Figura 15: Visão geral do processo de seleção dos artigos	137

LISTA DE SIGLAS E ABREVIASÕES

Sigla	Significado
API	<i>Application Programming Interface</i>
BD2K	<i>Big Data to Knowledge</i>
BI	<i>Business Intelligence</i>
HIAE	Hospital Israelita Albert Einstein
IA	Inteligência Artificial
INCA	Instituto Nacional de Câncer
INFORMATIZA APS	Programa de Apoio à Informatização e Qualificação dos Dados da Atenção Primária à Saúde
ML	<i>Machine Learning</i>
MS	Ministério da Saúde
MSL	Mapeamento Sistemático da Literatura
NIH	<i>National Institutes of Health of the United States</i>
NPCR	<i>National Program of Cancer Registries</i>
PDCA	<i>Plan-Do-Check-Act</i>
PEP	Prontuário Eletrônico do Paciente
PIUBS	Programa de Informatização das Unidades Básicas de Saúde
PNS	Plano Nacional de Saúde
PROADI-SUS	Programa de Apoio ao Desenvolvimento Institucional do Sistema Único de Saúde
RHC	Registro Hospitalar de Câncer
SAS	Secretaria de Atenção à Saúde
SETI	Serviço de Tecnologia da Informação do INCA
SIS	Sistema de Informação em Saúde
SUS	Sistema Único de Saúde
TICs	Tecnologia da Informação e da Comunicação

SUMÁRIO

1. INTRODUÇÃO	22
1.1. Apresentação do tema	22
1.2. Problema de Pesquisa	24
1.3. Objetivos.....	27
1.3.1. Objetivo Final	27
1.3.2. Objetivos Intermediários.....	27
1.4. Relevância da Pesquisa.....	27
1.5. Oportunidades	28
1.6. Viabilidade	28
1.7. Contribuições da Pesquisa.....	29
1.8. Percurso da Pesquisa	30
2. FUNDAMENTAÇÃO TEÓRICA.....	34
2.1. Tecnologias da informação e da comunicação (TICS) na saúde	34
2.2. Conceituação de <i>big data</i>	38
2.2. Informação: a força motriz do <i>big data</i>	43
2.4. O uso estratégico do <i>big data</i> no setor de saúde	45
2.5. Desafios para implementação do <i>big data</i> na saúde	49
2.6. Projeções futuras para o uso do <i>big data</i>	53
2.7. Possíveis caminhos para implementação do <i>big data</i> em uma instituição de saúde.....	55
3. METODOLOGIA	60
3.1 Caracterização da Pesquisa	62
3.2 Seleção dos Sujeitos da Pesquisa	62
3.3 Coleta de Dados e Plano Piloto	63
3.4 Análise dos Dados	69
3.5 Limitações do Método de Pesquisa.....	72
4. ESTUDO APLICADO	73
4.1 Caracterização da unidade de estudo	73
4.2 Produto Tecnológico	78
5. RESULTADOS	80
5.1 Descrição dos Participantes	80
5.2 Categorias de Análise	81
5.3 CP 1: Desafios e Oportunidades para Implementação do <i>Big Data</i> no INCA	85
5.3.1.1 CE 1: Governança dos dados	86
5.3.1.2 CE 2: Manutenção da qualidade dos dados gerenciados pelo INCA.....	90
5.3.1.3 CE 3: Integração e Interoperabilidade dos Sistemas Ativos no INCA.....	99
5.4 CP 2: Benefícios potenciais da adoção do <i>big data</i> no INCA.....	105
5.5 CP 3: Preocupações éticas relacionadas à implementação do <i>big data</i> no INCA.....	116
5.6 CP 4: O uso do <i>big data</i> para tomada de decisões estratégicas no INCA.....	118
6. PLANO ESTRATÉGICO PARA A IMPLEMENTAÇÃO DO BIG DATA NO INCA	124
6.1 Apresentação do Produto Tecnológico: Plano Estratégico para Implementação do Big Data no INCA	
124	
6.2 Relatório Técnico Conclusivo: Propostas estratégicas para implementação do <i>big data</i> no INCA.....	127

6.2.1	Contextualização	127
6.2.2	Plano Estratégico para Implementação do Big Data no INCA - Ferramenta 5W2H.....	129
6.2.3	Mapeamento do Processo para Implementação do Big Data no INCA	135
7.	PROPOSTA DE UM PLANO DE AÇÃO PARA MITIGAR OS RISCOS ASSOCIADOS A IMPLEMENTAÇÃO DO BIG DATA NO INCA.....	139
7.1	Plano de Ação: Análise dos Riscos associados a Implementação do Big Data no INCA	139
7.2	Identificação dos riscos.....	141
7.3	Análise de frequência e impacto potencial dos riscos identificados.....	143
7.4	Frequência dos Riscos	143
7.5	Impacto Potencial dos Riscos	144
7.6	Considerações sobre os Riscos de Baixo Impacto.....	145
7.7	Estratégia de resposta aos riscos identificados	146
8.	CONCLUSÃO.....	149
8.1	Síntese dos Resultados e Contribuições	149
8.2	Reflexão sobre os Desafios e Oportunidades	150
8.3	Trabalhos Futuros	150
8.4	Considerações Finais	151
REFERÊNCIAS		153
APÊNDICE A: MAPEAMENTO SISTEMÁTICO DA LITERATURA (MSL)		160
APÊNDICE B: ROTEIRO DE ENTREVISTA SEMIESTRUTURADO - PROFISSIONAIS DE SAÚDE DO INCA.....		168
ANEXO A: TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO		171
ANEXO B: TERMO DE ANUÊNCIA PARA AUTORIZAÇÃO DA PESQUISA NO INCA... ..		174
ANEXO C: PARECER CONSUBSTANIADO DO COMITÊ DE ÉTICA EM PESQUISA		176

1. INTRODUÇÃO

1.1. Apresentação do tema

As Tecnologias de Informação e Comunicação (TICs) se tornaram componentes fundamentais para o desenvolvimento de diferentes áreas da sociedade, incluindo o setor de saúde (ADJEKUM et al, 2018). A digitalização de processos e os avanços nas tecnologias digitais, impactaram fortemente este setor, levando a uma mudança de paradigma que hoje é comumente chamada de saúde digital, e-saúde ou *mHealth* (SILVENNOINEN, 2020).

A saúde digital compreende o uso de recursos das TICs para dar suporte à saúde e áreas correlatas, abrangendo não só a prestação de serviços e assistência médica ao paciente, mas também, as áreas de vigilância e pesquisa (BALTAXE et al, 2019; SAYANTAN, 2020). Sua força motriz está na capacidade dos sistemas integrados de coletar, armazenar e analisar uma grande quantidade de dados, permitindo uma ampla compreensão das informações geradas nas etapas pré e pós consulta de uma parcela significativa da população (HOSSAIN et al, 2019).

A utilização de tecnologias como o *big data*, *Machine Learning* e Inteligência Artificial (IA) fazem parte da saúde digital, onde um exemplo disso é o projeto do *National Institutes of Health of the United States* (NIH), que apoia o desenvolvimento e a integração de ferramentas tecnológicas para expandir o uso do *big data* e da ciência de dados na área biomédica, intitulado *big data to Knowledge* (BD2K). O projeto visa desenvolver e fornecer novas abordagens e ferramentas metodológicas, para potencializar a análise do *big data* e, a partir disso, gerar conhecimentos fundamentais para os profissionais e pesquisadores da área (NIH, 2020).

No Brasil, buscando contribuir para o desenvolvimento de soluções relevantes para a saúde nacional, o Hospital Israelita Albert Einstein (HIAE) apresentou em 2021, um projeto de *big data* e Inovação em parceria com Universidades Federais e Hospitais Universitários para o Ministério da Saúde, dentro do Programa de Apoio ao Desenvolvimento Institucional do Sistema Único de Saúde (PROADI-SUS). A principal finalidade da proposta é promover a parceria entre o setor público e privado, para que, por meio de um trabalho colaborativo, seja possível transferir tecnologias e implementar ações estratégicas em conjunto para o desenvolvimento igualitário dos profissionais de saúde. Ressalta-se também, que o projeto favorece o compartilhamento de conhecimentos entre

profissionais, pesquisadores e professores da área da saúde e tecnologia como forma de criar estratégias, metodologias e ferramentas inteligentes para solucionar problemas complexos (BRASIL, 2021).

Ao analisar o cenário da saúde pública nos últimos anos, observa-se indícios de que os desafios atuais do setor vão além da necessidade de integração entre as frentes do poder público, privado e acadêmico, e inclui a complexidade de análise de grandes volumes de dados estruturados e não estruturados e a necessidade de investimentos significativos para aquisição dos recursos (tecnológicos e humanos), fundamentais para viabilizar e potencializar os benefícios do *big data* no setor. Além disso, é preciso que as instituições se preocupem com as boas práticas para manipulação dos dados dos pacientes, a fim de garantir a segurança e privacidade na coleta e compartilhamento de dados sensíveis (SCHMARZO, 2018; SAYANTAN, 2020; CHIARELLO, 2021).

Do ponto de vista da Gestão de Saúde Pública, os dados de saúde são utilizados como subsídio para a gestão do planejamento estratégico e a tomada de decisões (WONG et al, 2022). Um exemplo disso, está na análise de dados demográficos, que permite ao gestor público identificar as regiões com menores taxas de encaminhamento populacional aos serviços de assistência médica (D'AGOSTINO et al, 2021).

Esses conhecimentos são utilizados para a implementação de políticas públicas destinadas a promoção de ações de intervenção de saúde populacional, congruentes com as particularidades de cada região, garantindo assim, o direito individual de acesso a saúde de forma igualitária (SAUNDERS et al, 2020). Essas tendências foram percebidas pelo Ministério da Saúde, que adotou iniciativas como a produção do plano de Estratégia de Saúde Digital do Brasil (FERREIRA; CALVOSA, 2021), contendo recomendações para a elaboração de estratégias nacionais, buscando acelerar o processo de digitalização no Sistema Único de Saúde (SUS) (BRASIL, 2012).

Para compreender como o processo de inovação tecnológica se aplica às unidades do SUS, que são plurais, atuam em diferentes segmentos sociais e com diferentes finalidades, é necessário observar o processo de evolução das TICs na área da saúde, tendo em vista que este é um dos pilares para a aplicabilidade do *big data* em uma organização de saúde (FILHO; PEREIRA, 2021). Além disso, dada a abrangência da terminologia do *big data*, é importante observá-lo sob diferentes óticas, para que a partir desse entendimento, seja possível identificar as ferramentas de gestão mais eficazes e aplicáveis ao setor de saúde pública.

1.2. Problema de Pesquisa

Nos últimos anos, observou-se um aumento da busca feita por diferentes setores por ferramentas analíticas, eficazes para a análise de grandes volumes de dados (*big data*). Isso se deve ao fato de as organizações estarem procurando maneiras de melhorar a tomada de decisão, aumentar a vantagem competitiva e potencializar o desempenho de suas atividades em seu nicho de atuação (MARCONI, 2020; HUSSAIN et al, 2021).

Esta tendência também é perceptível nas instituições de saúde, onde houve uma mudança no modelo de gestão, que deixou de ser centrado na doença e passou a ser centrado no paciente¹ (SENTHILKUMAR et al, 2018). A complexidade do sistema de saúde fez com que as instituições percebessem a importância de gerenciar e analisar as suas bases de dados, para que dessa forma, pudessem ter uma visão sistêmica do cenário em que estão inseridas, melhorando assim seus processos e serviços (LEE et al, 2020).

Em paralelo, a introdução de prontuários eletrônicos, o compartilhamento de informações por meio de sensores, equipamentos médicos e plataformas digitais de saúde, fizeram com que os fluxos de dados dos pacientes aumentassem exponencialmente (LUPTON et al, 2016; CARDONA, 2022). Outro ponto relevante está no fato de que a prática em saúde gera quantidades significativas de dados por meio da inclusão de imagens, registros clínicos, dados genômicos e comportamentos de saúde nos sistemas integrados (GUPTA et al, 2019) todos estes fatores impulsionaram o aumento dos dados de saúde e mudaram a perspectiva de uso dos dados dos pacientes por parte das instituições (CARDONA, 2022).

Uma questão frequentemente levantada quando se trata do uso de dados na saúde é o uso estratégico do *big data* (CHIARELLO, 2021), fator fundamental para que as informações extraída dos dados sejam capazes de apoiar a tomada de decisões das instituições e fomentar a adoção de ações inteligentes no setor (SAYANTAN et al, 2020). Por meio do uso estratégico do *big data*, as instituições de saúde identificam tendências, aumentam a sua capacidade de prever demandas futuras e alocam recursos de maneira mais eficientes (SAYANTAN et al, 2020).

Além disso, a compreensão aprofundada das necessidades da população, obtida

¹ No paradigma tradicional, centrado na doença, o enfoque predominante residia no diagnóstico e tratamento de condições médicas específicas, por vezes negligenciando o contexto do paciente em sua totalidade. Com a mudança para um modelo centrado no paciente, adotou-se uma perspectiva holística, que integra não apenas o estado clínico, mas também os valores, contextos social, emocional e psicológico do indivíduo.

por meio da análise de dados demográficos e epidemiológicos, permite que estratégias personalizadas e direcionadas sejam formuladas, tomando como base o fluxo dos dados em tempo real (AL MAYAHI et al, 2018). Esse fluxo, capacita os profissionais de saúde a tomar decisões clínicas embasadas em informações verossímeis e com maior assertividade, viabilizando o diagnóstico precoce de doenças, gerando melhorias nos cuidados prestados ao paciente e aumentando a qualidade dos serviços de saúde prestados à população (SAYANTAN et al, 2020; YUNITTA, 2022).

No campo da oncologia, por exemplo, cerca de setecentos e quatro mil pessoas são diagnosticadas com câncer por ano no Brasil (INCA, 2022) e, ter acesso ao diagnóstico precoce da doença, poderá reduzir o tempo do paciente na fila de espera pelo tratamento (INCA, 2020). Neste cenário, o *big data* se apresenta como uma importante ferramenta no diagnóstico e tratamento de doenças multifatoriais, uma vez que, a partir da análise dados dos pacientes, é possível definir os tratamentos mais adequados para cada quadro da doença, acompanhar a evolução dos pacientes com câncer e melhorar a prática em saúde (BERTUCCI et al, 2019).

Quando o uso integrado de tecnologias é direcionado para melhorar a qualidade do tratamento do paciente e aumentar as chances de sobrevivência do indivíduo, é gerado valor para a área oncológica, que neste caso, está diretamente relacionado com a melhoria da qualidade de vida e com a redução do risco de morte do paciente com câncer (CHEUNG, 2020).

No Brasil, a coleta, armazenamento, processamento e análise dos dados de saúde dos pacientes diagnosticados com câncer atendidos em uma unidade hospitalar, pública, privada ou filantrópica, ocorre por meio do Registro Hospitalar de Câncer (RHC). Segundo o Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA, 2010), o RHC é um centro de coleta de dados de saúde, que permite acompanhar e avaliar a qualidade da assistência prestada ao paciente oncológico por meio da análise dos dados armazenados nessas bases (INCA, 2010).

O conjunto de dados de saúde coletados pelos hospitais são enviados para compor a Base Nacional dos Registros Hospitalares (BNRH). Esses dados são disponibilizados por meio do IntegradorRHC, Sistema de Informação em Saúde (SIS) desenvolvido pelo INCA para consolidação dos dados médicos (INCA, 2010). É importante ressaltar que o envio regular dos dados coletados para este sistema é obrigatório para os hospitais habilitados na Atenção Especializada em Oncologia do SUS e facultativo para os hospitais

não habilitados (INCA, 2020).

De acordo com o Ministério da Saúde (2022), até o ano de 2020 existiam 4.134 hospitais no Brasil. Desses, 1.598 estavam habilitados na BNRH, enquanto 2.536 não estavam. Dos 1.598 hospitais habilitados na BNRH, 1.468 eram hospitais gerais e 130 hospitais especializados em câncer (MINISTÉRIO DA SAÚDE, 2022).

Dos 130 hospitais especializados em atendimento oncológico no Brasil, 95 eram habilitados para a Atenção Especializada em Oncologia do SUS (MINISTÉRIO DA SAÚDE, 2022). Esses hospitais estão concentrados majoritariamente nas regiões Sudeste e Sul do país, com destaque para os estados de São Paulo, Rio de Janeiro, Minas Gerais, Paraná e Rio Grande do Sul, que juntos abrigam a maior parte das unidades. As capitais dessas regiões, como São Paulo, Rio de Janeiro, Belo Horizonte, Curitiba e Porto Alegre, concentram os principais centros de referência, enquanto em outras regiões, como Norte e Nordeste, a distribuição é mais dispersa e menos densa, refletindo disparidades no acesso ao tratamento especializado no cenário nacional (TEMPORÃO, 2021).

Os hospitais habilitados para a Atenção Especializada em Oncologia pelo SUS oferecem um atendimento integral aos pacientes com câncer, abrangendo desde o diagnóstico até o tratamento e o acompanhamento contínuo. Além disso, essas instituições são obrigadas a coletar dados sobre novos casos de câncer entre os pacientes atendidos, o que fortalece o monitoramento da incidência e da mortalidade por câncer no Brasil. Essa coleta de dados desempenha um papel crucial no planejamento e na avaliação dos programas de prevenção e controle do câncer, contribuindo para a formulação de políticas públicas mais eficazes (INCA, 2022).

Gupta et al (2019) explicam que a coleta e armazenando dos dados médicos dos pacientes fazem com que as instituições de saúde extraiam novos conhecimentos por meio da análise de diferentes tipos e fontes de dados. Neste sentido, os riscos de que registros de saúde relevantes deixem de ser considerados, como geralmente acontece em processos manuais, são mitigados a partir da interoperabilidade tecnológica e gerenciamento de grandes volumes de dados de saúde (GUPTA et al, 2019).

Complementando esta ideia, Yunitta et al (2022) afirmam que o uso de recursos de informática, a interligação das bases de dados de saúde, a troca automática de informações entre as instituições de saúde e a implantação de rotinas que permitem acompanhar a rastreabilidade e evolução do câncer, tende a aumentar e ganhar maior relevância para as instituições que compõem o sistema público de saúde (YUNITTA et

al, 2022; BERTUCCI et al, 2019).

Diante do potencial do uso de informações extraídas por meio do *big data* para apoiar o tratamento de doenças complexas como o câncer, o INCA foi escolhido como objeto de estudo para a coleta e análise de dados que subsidiarão as reflexões apresentadas nesta dissertação. Assim, este estudo busca responder à seguinte questão: *Quais estratégias podem ser adotadas pelo INCA para a implementação eficaz do big data na instituição?*

1.3. Objetivos

Para o desenvolvimento deste estudo são propostos os seguintes objetivos:

1.3.1. Objetivo Final

O presente estudo teve como objetivo final propor a elaboração de um plano estratégico com recomendações práticas para a implementação do *big data* no Instituto Nacional de Câncer (INCA). Além disso, foi desenvolvido um plano de ação complementar, na forma de um Produto Técnico-Tecnológico (PTT), que aborda os principais desafios que poderão surgir durante o processo de implementação do *big data* e as estratégias que poderão ser adotadas pela instituição para mitigá-los.

1.3.2. Objetivos Intermediários

Para alcançar o objetivo final deste estudo, foram definidos e perseguidos os seguintes objetivos intermediários:

- a) Descrever, com base na literatura, a relação entre tecnologia e o processo de evolução tecnológico na área da saúde;
- b) Identificar os impactos no setor de saúde ocasionados pelo aumento do volume exponencial de dados e a contribuição do *big data* para melhoria da gestão; e
- c) Desenvolver um plano estratégico com recomendações práticas que viabilizem a implementação do *big data* no INCA, por meio de um estudo aplicado.

1.4. Relevância da Pesquisa

Os aspectos relevantes capazes de justificar a elaboração de um plano estratégico

para implementação do *big data* no INCA são:

- a) Oferecer ao INCA uma visão sistematizada do uso do *big data* como uma ferramenta estratégica no setor de saúde;
- b) Identificar a potencialidade da adoção do *big data* no setor de saúde e descrever os benefícios dessa adoção;
- c) Descrever possíveis caminhos que podem ser adotados por uma instituição de saúde com contexto semelhante ao do INCA, para a implementação eficaz do *big data*; e
- d) Identificar as possibilidades de uso do *big data* e seus impactos nos serviços da área de oncologia.

1.5. Oportunidades

A pesquisa mostra-se oportuna devido ao potencial do uso do *big data* no setor de saúde e ao fato de que este projeto foi fundamentado com base no Planejamento Estratégico do INCA para o período de 2020-2023, alinhado ao Plano Nacional de Saúde (PNS), que serviu como base para definição das estratégias do Instituto a médio e longo prazo. Além disso, o apoio e incentivo dos gestores do setor de Tecnologia da Informação do INCA fortalecem a produção de conhecimento científico, sendo este o setor sob o qual a pesquisa será conduzida.

1.6. Viabilidade

Os principais aspectos de viabilidade levantados foram:

- a) De acesso: o fato de a pesquisadora ter tido vínculo empregatício na organização estudada, entre março de 2019 e setembro de 2022, e contar com o apoio institucional para o desenvolvimento do Plano Estratégico como uma proposta de melhoria organizacional, permite o acesso aos profissionais de saúde e as informações necessárias para a elaboração deste estudo. Além disso, o fato de ter participado ativamente do fornecimento da estrutura tecnológica para a prática da assistência hospitalar, dá a pesquisadora conhecimento sobre às tecnologias do Instituto.
- b) De tempo de execução: o tempo estimado para execução das etapas de levantamento de requisitos, identificação de oportunidades de evolução, mapeamento dos fluxos da proposta e elaboração do Plano Estratégico para implementação do *big data* no INCA é de 24 meses, a contar de fevereiro de 2022.

- c) De custo: o desenvolvimento do Plano Estratégico é uma iniciativa da área de Tecnologia da Informação do INCA e o levantamento das informações necessárias para a concretização do projeto serão colhidas na instituição. Desta forma, os custos para elaboração do Plano de Ação serão absorvidos pelas atividades institucionais, sem que exista a necessidade de investimentos iniciais.

1.7. Contribuições da Pesquisa

Espera-se que este estudo seja capaz de contribuir para o Instituto Nacional de Câncer (INCA), para instituições com contextos de saúde semelhantes ao do INCA, para a academia e para a sociedade em geral. As principais contribuições podem ser organizadas nas seguintes categorias:

- a) Contribuições para o INCA e instituições de saúde com contextos semelhantes: Este trabalho visa oferecer ao INCA e a outras instituições com contextos de saúde semelhantes ao do Instituto uma análise prática das estratégias que podem ser implementadas para adoção do *big data* em suas rotinas e processos organizacionais. O plano estratégico proposto é flexível e pode ser adaptado a diferentes realidades, beneficiando tanto o Instituto quanto outras organizações de saúde interessadas em aprimorar seus processos por meio do uso do *big data*.
- b) Contribuições para a academia: Este estudo explora as percepções de profissionais de saúde, gestores e de TI do INCA, uma instituição pública de saúde, sobre o uso do *big data*, com foco na oncologia. Os resultados e metodologias desenvolvidas podem servir de base para futuras pesquisas sobre o tema. Trabalhos derivados desta dissertação foram apresentados em eventos acadêmicos relevantes, como o EMPRAD (2023) e o EnANPAD (2024), sendo que um deles foi premiado como o melhor PTT no EnANPAD. Esses trabalhos estão disponíveis nos anais eletrônicos dessas associações acadêmicas.
- c) Contribuições para a sociedade: As estratégias propostas para a implementação do *big data* no INCA têm como objetivo contribuir para a melhoria da qualidade do atendimento ao paciente e para a otimização da utilização de recursos no tratamento oncológico. A adoção eficaz do *big data* pode resultar em diagnósticos precoces da doença, na personalização dos tratamentos de câncer, e em uma gestão mais eficiente dos dados dos pacientes, contribuindo assim para a elevação dos padrões de cuidado na

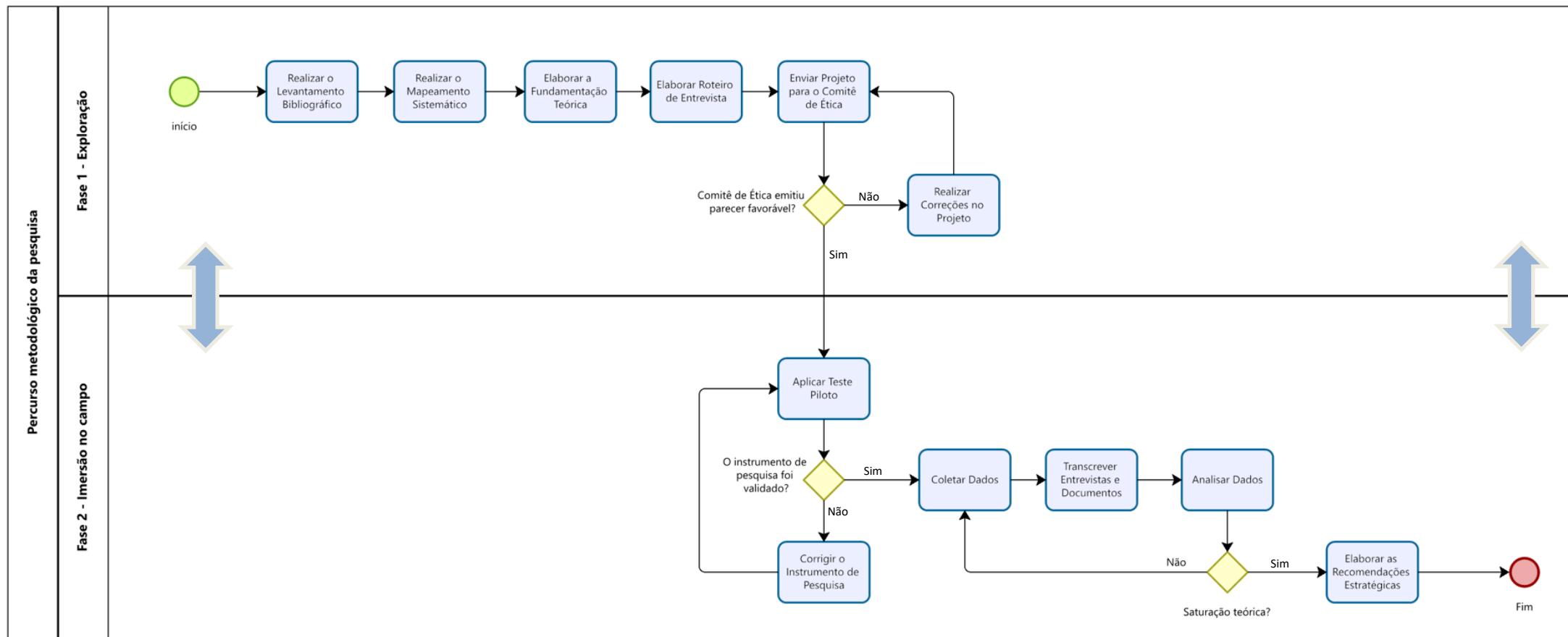
saúde e gerando potenciais benefícios para a sociedade.

1.8. Percurso da Pesquisa

A realização da presente pesquisa conta com duas fases: uma de exploração, que consiste na construção da parte teórica deste estudo e outra de imersão no campo. Durante a fase de exploração, foi realizado um levantamento bibliográfico para compreender o campo de pesquisa sobre o uso do *big data* no setor de saúde, além de um mapeamento sistemático da literatura (MSL), onde diversas abordagens teóricas foram estudadas, tendo como pilares fundamentais os seguintes temas: (1) Evolução tecnológica na área da saúde; (2) Compreensão do conceito de *big data* e; (3) Aplicabilidade do *big data* na saúde.

A fase de imersão no campo teve como objetivo aproximar o conhecimento teórico do campo de pesquisa, realizado por meio de um estudo aplicado no Instituto Nacional de Câncer José Alencar Gomes da Silva (INCA). A Figura 1, apresenta uma visão macro sobre o desenvolvimento deste estudo.

Figura 1- Visão geral do percurso metodológico da pesquisa.



Fonte: Elaboração própria (2024), adaptado de Silva (2021).

Na Fase 1, o levantamento bibliográfico e o mapeamento sistemático da literatura (MSL) subsidiaram a composição da fundamentação teórica desta pesquisa. No processo de seleção dos artigos foi adotado como critério de inclusão a necessidade do estudo abordar o uso do *big data* no setor de saúde. Em contrapartida, foram excluídos os artigos de revisão, os estudos que tratavam o problema de forma lateral, ou seja, apenas sobre a perspectiva da saúde privada e os artigos que não tratavam o uso do *big data* no setor de saúde.

Os resultados do MSL permitiram perceber que, no Brasil, os desafios para implementação do *big data* na saúde incluem o grau de complexidade da análise e armazenamento dos dados em tempo real, a necessidade de aquisição de tecnologias robustas por parte das instituições de saúde brasileira, a importância do fortalecimento da integração de dados entre as instituições de saúde nacional e do fortalecimento do processo de digitalização dos prontuário médicos, majoritariamente nas instituições públicas de saúde. Além disso, a pandemia da Covid-19, potencializou o uso de tecnologias no setor de saúde e aumentou, significativamente, o uso dos dados para nortear as decisões estratégicas das instituições de saúde.

Além disso, foi possível observar a importância da padronização dos registros de informações dos pacientes nos prontuários eletrônicos e a necessidade de uma mudança de perspectiva por parte das instituições de saúde sobre o investimento para adoção de tecnologias no setor, que ainda é visto como um custo e não como uma estratégia para prover a melhoria na qualidade dos serviços de saúde prestados à população. Ao final do MSL, 41 estudos foram selecionados e utilizados como base para a construção da Metodologia (Capítulo 3). Os procedimentos metodológicos do MSL estão apresentados de forma detalhada no Apêndice A – Mapeamento Sistemático da Literatura (MSL).

A Fase 2, detalhada nos Capítulos 4 e 5, corresponde à metodologia adotada na fase de imersão no campo, onde foi realizada uma pesquisa aplicada, de abordagem qualitativa e descritiva no INCA. Para a coleta de dados, foi realizada uma pesquisa documental, que precederá a realização de entrevistas por meio de roteiros semiestruturados, com os profissionais da área da saúde e gestores da área de Tecnologia da Informação do INCA.

Para análise dos dados, foi utilizada a técnica de análise de conteúdo, visando inicialmente transcrever os documentos e as entrevistas realizadas, a fim de obter as primeiras impressões dos entrevistados sobre o tema. Em seguida, procurou-se explorar o material coletado, com o intuito de descrever, a partir da fala dos sujeitos entrevistados, as suas

impressões sobre as possibilidades do uso do *big data* e seus impactos nos serviços na área da oncologia. Por fim, com base na frequência de repetição dos elementos, os dados serão categorizados para que, a partir disso, seja possível tratar os dados brutos obtidos.

A partir dos resultados encontrados, foi estruturado um plano estratégico com recomendações práticas para a implementação do *big data* no INCA. Além disso, foi desenvolvido um plano de ação complementar, na forma de um Produto Técnico-Tecnológico (PTT), que aborda os principais desafios que poderão surgir durante o processo de implementação do *big data* e as estratégias que poderão ser adotadas pela instituição para mitigá-los.

2. FUNDAMENTAÇÃO TEÓRICA

Este capítulo apresenta as reflexões teóricas, com base nos estudos selecionados no levantamento bibliográfico, a respeito dos seguintes temas inerentes à pesquisa realizada: (1) Evolução tecnológica na área da saúde; (2) Compreensão do conceito de *big data* e; (3) Aplicabilidade do *big data* na saúde.

2.1. Tecnologias da informação e da comunicação (TICS) na saúde

As TICs são mais do que uma ferramenta de apoio ao desenvolvimento e funcionamento de determinada estrutura ou organização, sendo, muitas vezes, compreendida como um dos seus pilares estruturais (JIMENEZ et al, 2021). Por meio do uso das TICs é possível ter uma visão sistêmica e integrada dos diferentes setores que compõem uma estrutura organizacional, que impactam na capacidade produtiva e no posicionamento estratégico adotado pela organização, tendo em vista a busca por se manter competitiva em um mercado em constante transformação (HAN et al, 2018).

Apesar das frequentes abordagens na literatura sobre o uso das TICs no setor produtivo, sua presença pode ser percebida em diferentes setores (XIA et al, 2020). Isso se deve, principalmente, ao fato de as ferramentas tecnológicas serem utilizadas como mediadoras do processo de comunicação na atualidade assumindo um papel fundamental no processo de desenvolvimento econômico, político e social (JANETT; YERACARIS, 2020).

De forma geral, as TICs referem-se aos dispositivos de computação e comunicação que gerenciam informações e são aplicáveis a inúmeros processos, gerando impactos em diferentes áreas. Devido a essa característica, também são chamadas de tecnologias transversais (FERREIRA; CALVOSA, 2021). A adoção das TICs possui relação direta com os resultados e objetivos estratégicos que a organização almeja alcançar, seu uso impacta diretamente na eficiência das operações e no aumento da produtividade, uma vez que, as TICs complementam e potencializam outros recursos competitivos da organização (ERAZO; ACHICANOY, 2016).

Na área da Administração Pública, sua adoção possui múltiplos objetivos, entre eles: potencializar o alcance e a melhoria contínua da qualidade dos serviços prestados; viabilizar a eficácia, a eficiência e a transparência dos atos administrativos; auxiliar no processo de fiscalização das ações governamentais; e potencializar o exercício da cidadania por meio da facilidade de acesso do cidadão, aos serviços públicos ofertados na Internet (REIS; DE LESSA CARVALHO, 2020).

No setor de saúde, a intensificação da adoção das TICs transformou as práticas dos

profissionais, tanto nas esferas pública quanto na iniciativa privada, gerando impactos nos três níveis de atenção à saúde: o primário, o secundário e o terciário (CARLOTO; DINIS, 2018). Desta forma, neste setor as TICs estão presentes desde as ações mais simples, como a visualização de exames e agendamentos de consultas através das plataformas digitais, até as mais complexas, como o diagnóstico e monitoramento de doenças multifatoriais (JIMENEZ et al, 2021).

Recentemente, o trabalho elaborado por Penteado et al (2021) analisou o campo de pesquisa da informática médica no período de 1960 até 2020 e apresentou a evolução temporal da área, mostrando variações em relação aos termos usados para sua nomenclatura e as principais tecnologias desenvolvidas nos últimos 30 anos. Este estudo sugere que o desenvolvimento tecnológico e a maturidade da área da saúde podem ser classificados em quatro temas principais: Informática médica; Informática em saúde; Saúde eletrônica e; Saúde digital (PENTEADO et al, 2021); logo, entende-se que, dependendo do cenário em que as TICs estão inseridas, a natureza da sua compreensão pode ser modificada.

No setor de saúde pública, termos como *eHealth*, *mHealth* e Saúde Digital são utilizados para caracterizar o uso de um conjunto de tecnologias, onde o seu significado varia de acordo com o foco da sua aplicação (SAVOY et al, 2021). De modo geral, esses campos de conhecimento compreendem desde os sistemas de informação hospitalares de suporte à decisão clínica, o que incluí a integração de informações de saúde, armazenamento de imagens, textos e sinais vitais, teles saúde e os registros eletrônicos de pacientes, até, mais recentemente, os tópicos de coleta de grandes volumes de dados por meio de equipamentos eletrônicos versáteis, redes sociais, computação em nuvem e a aplicação de inteligência artificial para a descoberta de padrões e para a análise de dados médicos (PENTEADO et al, 2021).

Shaw et al (2018) e Soriano et al (2019), definem *eHealth* como a prática de cuidados de saúde suportados por soluções de Tecnologia da Informação e Comunicação, conectados através da Internet (SORIANO et al, 2019). Seu foco está no usuário. Nessa perspectiva, as TICs são utilizadas como um meio capaz de ampliar e facilitar o acesso do cidadão aos serviços e informações de saúde (SHAW et al, 2018).

A *mHealth* ou *Mobile Health*, pode ser definida como as práticas de saúde realizadas por meio de aparelhos móveis, como smartphones, assistentes digitais e dispositivos de monitoramento (FATEHI et al, 2020). Seu foco está nas tecnologias sem fio e móveis. Com o *mHealth*, médicos e pacientes podem ter acesso a informações para prevenção, controle, tratamento e monitoramento de doenças crônicas por meio de aplicativos e plataformas mobile

(ZEEB et al, 2020).

Vale ressaltar que o *mHealth* e *eHealth* compõem a Saúde Digital, que utiliza as TICs para produzir e disponibilizar informações confiáveis em saúde para a população, profissionais da área e gestores públicos. (KLOC et al, 2019). Seu foco está em aumentar a capilaridade dos serviços e no aprimoramento da coordenação dos sistemas de saúde, por meio das tecnologias (WIENERT, 2019).

Apesar da similaridade dos conceitos apresentados, o termo saúde digital é mais abrangente e incorpora os recentes avanços tecnológicos (WONG et al, 2022; GHASSEMI; MOHAMED, 2022). Penteado et al (2021) complementam ao explicar que, o uso de tecnologias como big data, computação em nuvem, dispositivos móveis, internet das coisas e inteligência artificial, impactaram diretamente na maneira como os serviços de saúde são prestados ao cidadão, levando a uma mudança paradigmática na área da saúde (PENTEADO et al, 2021).

A Figura 2 apresenta as palavras mais frequentes encontradas nos resumos dos artigos utilizados para compor esta seção e que abordam o tema da saúde digital.

Figura 2 - Visualização da nuvem de *tags* estática de termos-chave que aparecem em resumos de artigos relacionados a saúde digital.



Fonte: Elaboração própria (2023), com base nos artigos selecionados no MSL.

Na saúde pública brasileira, a saúde digital tem como principal objetivo ampliar o acesso do cidadão aos serviços de saúde. Para alcançar este objetivo, foram criados programas que visam à informatização e a integração dos dados de saúde dos cidadãos entre os estabelecimentos de saúde e os órgãos de gestão em saúde das unidades federativas, por meio

da adoção de inovações tecnológicas e o fomento ao amplo acesso as informações médicas (OPAS, 2021), (BRASIL, 2021). Três das principais iniciativas resultantes destes esforços são, o Programa de Apoio à Informatização e Qualificação dos Dados da Atenção Primária à Saúde (Informatiza APS), o Programa de Informatização das Unidades Básicas de Saúde (PIUBS) e o programa Conecte-SUS, criados pelo Ministério da Saúde (MINISTÉRIO DA SAÚDE, 2021).

O PIUBS surgiu em 2017 como uma proposta do Ministério da Saúde (MS) para apoiar o processo de digitalização da saúde pública brasileira. Inicialmente, o projeto consistia na implementação do Prontuário Eletrônico do Paciente (PEP) nas unidades básicas de saúde do SUS. Após os resultados obtidos nestas unidades, como a melhoria da qualidade dos registros de saúde dos pacientes e a facilidade de acesso a essas informações, o uso do projeto se expandiu para as unidades de atenção primária, secundária e terciária do sistema de saúde pública nacional (SOUZA, 2018).

No ano de 2019, foi criado o Informatiza APS, programa que faz parte da estratégia de saúde digital do Ministério da Saúde, o Conecte SUS, cujo objetivo é apoiar a informatização das unidades de saúde públicas e a qualificação e unificação dos dados em saúde, disponíveis nas instituições públicas e privadas de todo o Brasil (GADENZ, 2021). No entanto, apesar do aumento do desenvolvimento de políticas nacionais voltadas para a implementação de estratégias digitais no Brasil, Fornazin et al (2022) alertam que existe um cenário de divisão digital que faz com que parte da população não tenha direito a utilizar estes benefícios, seja por desinformação ou pela exclusão social, sendo este um paradigma no que tange a ampliação do acesso do cidadão a saúde.

Sendo assim, a perspectiva teórica adotada neste estudo, é a de Borovskaya, Masych e Fedosov (2020), que abordam o uso das TICs sob o prisma da saúde social, que se refere à aplicação integrada dessas tecnologias para melhorar a qualidade de vida, o bem-estar e os serviços de saúde da população. Especificamente no setor de saúde, as TICs devem atuar como transformadoras de processos, se apresentando como catalisadoras para melhorar a qualidade de vida dos pacientes, otimizando os fluxos de trabalho dos profissionais e integrando recursos na área da saúde (MAYER-SCHONBERGER; CUKIER, 2022), sem deixar de observar aspectos sociais, econômicos e ambientais.

2.2. Conceituação de *big data*

Na literatura pode-se encontrar diversas definições de *big data*. Este conceito evoluiu nos últimos anos, no entanto, ainda não é claramente compreendido e não há uma unanimidade quanto a sua definição. Dada a abrangência do termo, o *big data* pode ser tratado como: uma grande quantidade de dados digitais, um grande repositório de dados, uma ferramenta, tecnologia ou um fenômeno (cultural ou tecnológico) (GUPTA et al, 2019; HRIPCSAK, 2021; CARDONA et al, 2022; YUNITTA et al, 2022).

Na visão de Lupton e Mike (2016), o *big data* pode ser considerado como um conjunto massivo de dados digitais, gerados continuamente e que são produzidos por meio das interações de tecnologias online. Seguindo o mesmo raciocínio, Gupta et al (2019) definem o *big data* como uma volumosa base de diferentes tipos de dados, com potencial para gerar *insights* e que representa um desafio para as técnicas tradicionais de armazenamento e análise de dados.

Tanto Lupton e Mike (2016) quanto Gupta et al (2019), ressaltam que, extrair informações úteis do *big data* não é uma tarefa simples e ressaltam que informações de valor são resultantes de um tratamento adequado dos dados. Gupta et al (2019) complementam esta ideia ao explicar que, para extrair informações relevantes do *big data* é fundamental que a equipe responsável pela manipulação e gerenciamento dos dados seja composta por profissionais multidisciplinares e qualificados.

Sob uma lente teórica diferente, Carter (2011), Knapp (2018) e Hripcsak (2021), percebem o *big data* como ferramentas, processos e procedimentos que permitem a uma organização criar, manipular e gerenciar volumosos conjuntos de dados e bases de armazenamento. Sob este ponto de vista, o *big data* é identificado como uma ferramenta capaz de reunir diferentes tipos e fontes de dados, que permitem aos usuários gerenciar grandes quantidades de informações (KNAPP, 2018). O autor seminal Carter (2011), se referiu as tecnologias de *big data* como parte de uma geração de tecnologias e arquiteturas, projetadas para extrair valor de um conjunto de dados, em alta velocidade.

Combinando essas duas abordagens, Hripcsak (2021) define o *big data* como um sistema complexo, onde o gerenciamento dos dados é dependente de *hardwares* e *softwares* robustos, sendo fundamental que a organização conte com grupo de profissionais capacitados e com expertise para extrair informações relevantes e de fácil compreensão (HRIPCSAK, 2021).

Embora existam diversas definições, o conceito estabelecido por Beyer e Laney (2012) que se concentra em listar as características do *big data* com base no modelo "3 V's": Volume,

Velocidade e Variedade, é provavelmente a mais popular dentre elas, podendo ser observada em parte significativa dos estudos destinados a esta área de conhecimento (ZHIHAN, 2020; LIANG, 2021). Posteriormente, outros autores estenderam o modelo dos “3 V’s” e, como resultado, uma gama de recursos foram adicionados à lista, sendo: variabilidade, veracidade, disponibilidade e valor dimensões do *big data*, sendo estas amplamente aceitas na comunidade acadêmica (YUNITTA et al, 2022).

É importante compreender que o *big data* é composto por dados estruturados e não-estruturados como textos, vídeos, imagens, registros das redes sociais e informações geradas em tempo real, por meio de sensores e dispositivos interconectados (LEE, 2020). Portanto, ao descrever o *big data*, não se pode esquecer que o termo se refere mais a um fenômeno social do que a uma tecnologia específica (CARDONA, 2022).

Agrawal (2020), Al Mayahi (2019), Zhang (2018), Gandomi (2018) e Oiszak (2018) descrevem o fenômeno do *big data* a partir de características relacionadas a sua natureza, incluindo uma coleção de V’s. O Quadro 1, descreve os recursos apontados por estes últimos autores como fatores capazes de transformar um ativo de informação de alto volume em valor.

Quadro 1 - Características do *big data* com base na coleção de V’s.

Identificação	Definição	Desafio
Volume	Refere-se à quantidade de dados armazenados em um repositório de dados.	Analizar e extrair informações relevantes de um conjunto de dados complexos.
Velocidade	Rapidez com que novos dados são gerados.	Gerenciar dados não-estruturados de forma eficaz e em tempo real.
Variedade	Heterogeneidade de dados. Diferentes tipos de dados de saúde.	Obter <i>insights</i> a partir da análise de grandes volumes de dados heterogêneos, disponíveis de maneira holística.
Variabilidade	Inconsistência de dados.	Corrigir a interpretação de dados que podem variar de forma

Identificação	Definição	Desafio
		significativa, dependendo do contexto.
Veracidade	Refere-se a confiabilidade e qualidade dos dados.	Adotar estratégias e práticas seguras para coletar, armazenar, recuperar e analisar os dados.
Visualização	Capacidade de interpretar os dados e criar conhecimentos.	O desafio está em gerar insights a partir das informações extraídas do <i>big data</i> .
Valor	Descoberta de <i>insights</i> e conhecimentos que aumentam as chances de uma organização alcançar um determinado objetivo.	Descobrir os conhecimentos ocultos em grandes quantidades de dados complexos.

Fonte: Elaboração própria (2023), com base em Cardona (2022), Agrawal e Prabakaran (2020), Zhang (2018) e Gandomi (2018).

O valor do *big data* está diretamente relacionado com a análise dos dados e extração de informações úteis do *big data* para as partes interessadas. Estes *insights* são capazes de auxiliar no processo de tomada de decisões estratégicas das instituições e facilitar a resolução de problemas estruturais (CHRISTENSEN et al, 2019).

Jindal et al (2018) complementam esta ideia ao abordar o conceito de valor sobre a perspectiva do uso do *big data* no setor de saúde por meio das tecnologias. O autor cita como exemplo os modelos assistenciais de saúde, onde os serviços prestados ao paciente foram ampliados por meio da computação em nuvem, que forneceu escalabilidade e flexibilidade para as plataformas digitais de serviços médicos, aumentando a segurança e a velocidade do compartilhamento das informações sensíveis (SAYANTAN et al, 2020).

Buscando elucidar o entendimento do que é o *big data* na visão dos autores que se debruçaram sobre o tema, o Quadro 2 foi elaborado para sintetizar os diferentes conceitos de *big data* apresentados ao longo deste capítulo.

Quadro 2 - Quadro-síntese dos conceitos de *big data*.

Perspectiva	Conceito	Autor(es)
Repositório de dados	<i>O big data</i> é um conjunto volumoso de dados, cujo processamento exige o uso de ferramentas tecnológicas robustas e desenvolvidas especificamente para este fim. Cabe ressaltar que, quanto maior o conjunto de dados mais complexo é obter algum valor a partir dele.	Ohlhorst (2012)
	<i>O big data</i> pode ser compreendido como conjuntos massivos de dados digitais, gerados continuamente e que são produzidos por meio das interações de tecnologias <i>online</i> .	Lupton e Mike (2016)
	<i>O big data</i> é um conjunto de diferentes tipos de dados estruturados e não-estruturados em crescimento exponencial, com potencial para gerar <i>insights</i> e que representa um desafio para as técnicas tradicionais de armazenamento e análise de dados.	Gupta et al (2019)
Ferramenta Tecnológica	<i>O big data</i> faz parte de um robusto grupo de tecnologias desenvolvidas para processar uma vasta variedade de dados em alta performance.	Carter (2011)
	<i>O big data</i> é uma ferramenta capaz de reunir diferentes tipos e fontes de dados em um repositório robusto e eficiente, viabilizando o gerenciamento e armazenamento de informações para as organizações.	Knapp (2018)
	<i>O big data</i> é um sistema complexo, onde o processo de gerenciamento destes dados é dependente de <i>hardwares</i> e <i>softwares</i> de alta performance, sendo fundamental que a organização conte com grupo de profissionais capacitados e com <i>expertise</i> para extrair novos conhecimentos <i>do big data</i> .	Hripcak (2021)

<i>Big data</i> com base na coleção de V's	<p><i>O big data</i> é um conjunto de dados que pode ser definido por meio de três características: volume, uma vez que <i>o big data</i> é composto por enormes quantidades de dados; velocidade, pois os dados são criados em tempo real; e variedade, considerando que <i>o big data</i> é constituído por dados estruturados, semiestruturados e não estruturados.</p>	Beyer e Laney (2012)
Fenômeno (cultural ou tecnológico)	<p>O fenômeno <i>do big data</i> impactou profundamente a sociedade contemporânea em diferentes segmentos. Seu potencial para armazenar uma grande quantidade de dados complexos em tempo real, modificou a maneira tradicional de análise e processamento dos dados. Desta forma, as organizações orientadas por dados, precisam adotar estratégias e práticas seguras para coletar, armazenar, recuperar e analisar os dados com veracidade, para que desta forma, seja possível descobrir os conhecimentos ocultos do <i>big data</i>.</p>	Zhang (2018) Gandomi (2018) Oiszak (2018) Al Mayahi (2019) Agrawal e Prabakaran (2020)

Fonte: Elaboração própria (2023).

Neste estudo, o *big data* é compreendido como fluxos e não estoques de informações, o que reforça a importância da mudança de visão das organizações sobre a amplitude e aplicabilidade do *big data* em diferentes setores (LIANG, 2021). Especificamente no setor de saúde, entende-se que o *big data* não pode ser definido apenas como uma ferramenta robusta para o armazenamento de grandes quantidades de dados, mas principalmente, como um ativo de valor para gerar *insights* para a tomada de decisão por meio do gerenciamento, processamento e análise dos dados.

2.2. Informação: a força motriz do *big data*

Uma das razões fundamentais para a existência do *big data* é a extensão com que uma informação pode ser gerada e disponibilizada na sociedade contemporânea. O aumento do fluxo dos dados foi potencializado pelo processo de digitalização, onde as informações analógicas foram convertidas em formato digitais e legível por máquinas (RESNYANSKY, 2019).

A digitalização alcançou ampla popularidade por meio dos projetos de digitalização em massa, onde o *Google Print Library* foi pioneiro. Criado em 2004, pela empresa global de tecnologia *Google* em parceria com as universidades norte-americanas *Harvard*, *Stanford* e *Oxford*, o projeto tinha como objetivo converter acervos de bibliotecas de livros impressos em coleções digitais, que posteriormente foram disponibilizados para consultas públicas *online e offline*, por meio das plataformas do *Google* (BORISSOVA, 2018). No entanto, Lemos (2021) alerta que, o processo de digitalização não é uniforme, pois as desigualdades de renda e acesso as TICs dentro dos territórios, acentuam a desproporcionalidade deste processo.

Recentemente, Ricarte (2020) realizou um estudo quantitativo para analisar a expansão do processo de digitalização na vida cotidiana em uma perspectiva global e adotou como recorte temporal os anos de 2019 e 2020. Os resultados apontaram que, as medidas restritivas impostas pelos governos para tentar conter o aumento da taxa de contágio do novo Coronavírus, levou a uma maior utilização dos meios digitais quando comparado ao ano anterior. Isso se deve ao aumento da necessidade de manutenção das relações sociais e da prestação dos serviços essenciais a população, além das atividades educacionais e de trabalho, que aumentaram não só o tráfego de compartilhamento de dados, mas a importância do armazenamento e análise desses fragmentos de informações por parte das organizações.

Na percepção de Mayer-Schonberger e Cukier (2022), a tendência tecnológica onde as representações da vida cotidiana são captadas através dos dados originou o fenômeno da Dataficação, que para os autores é a evolução do processo de digitalização (MAYER-SCHONBERGER; CUKIER, 2022). Segata e Rifiotis (2021) explicam que a diferença fundamental é que, a digitalização permite que as informações analógicas sejam transferidas e armazenadas em um formato digital estruturado, enquanto a dataficação visa organizar a versão digitalizada de sinais analógicos para gerar *insights* e criar uma memória encadeada dos dados coletado (SEGATA; RIFIOTIS, 2021).

Há um consenso entre os autores de que, ambos os fenômenos têm como consequência o aumento exponencial do volume dos dados, que são potencializados pela integração dos artefatos tecnológicos conectados em rede, tornando o ambiente propício para a expansão de

tecnologias como *o big data* (RICARTE, 2020; LEMOS, 2021; SEGATA; RIFIOTIS, 2021). Além disso, a rapidez com que as tecnologias estão avançando, aumenta o reconhecimento do valor dos dados pelas organizações, que cada vez mais incorporam em suas práticas gerenciais e de produção a mentalidade das organizações orientadas por dados (MAYER-SCHONBERGER; CUKIER, 2022).

Lotfi, Srinivasan, Ertz e Latrous (2023) explicam que, as empresas orientadas por dados utilizam a análise *do big data* para coletar novas informações e identificar padrões desconhecidos de um conjunto de dados massivos, com o objetivo de aprimorar as operações, fornecer um atendimento personalizado ao cliente, gerar campanhas de *marketing* direcionadas e adotar práticas que o diferenciem dos concorrentes em um determinado segmento, visando aumentar a receita e a lucratividade da organização a longo prazo (LOTFI et al, 2023).

Esta ideia dialoga com a percepção dos autores seminais Davenport e Dyché (2013), que no estudo “*Big data in big companies: International Institute for Analytics*”, apresentaram as características das empresas que incorporaram os dados como elementos fundamentais para embasar suas decisões de negócios. Para os autores, essas empresas usam os dados em suas decisões, interações e processos como uma ferramenta estratégica para tomar melhores decisões.

Além disso, a prática organizacional orientada por dados, promove a melhoria contínua para gerar informações de valor em tempo real, aumentando a sua vantagem competitiva em um mercado em constante transformação (DAVENPORT; DYCHÉ, 2013). A junção destes fatores atrelados a crescente necessidade de cooperação e compartilhamento de dados entre as organizações de um mesmo nicho para obter *insights* valiosos, é um dos fatores que torna *o big data* um diferencial competitivo para as organizações de diferentes áreas de atuação (LOTFI et al, 2023).

2.4. O uso estratégico do *big data* no setor de saúde

No setor de saúde, o uso estratégico do *big data* pode ser compreendido como a capacidade de uma instituição obter novas informações por meio da análise dos dados, a fim de gerar conhecimentos úteis e com valor significativo para a sociedade (SAYANTAN et al, 2020). Sob esta ótica, a verdadeira revolução não está somente nas máquinas que processam grandes quantidades de dados médicos, mas sim na maneira como as instituições gerenciam e extraem valor do *big data* (MAYER-SCHONBERGER; CUKIER, 2022).

A pesquisa conduzida por Wang et al (2018) abordou o uso estratégico do *big data* na saúde por meio dos impactos da adoção da tecnologia em quatro frentes: (1) qualidade dos serviços em saúde; (2) práticas médicas; (3) atividades científicas e de pesquisa; e (4) gestão das instituições de saúde. Neste estudo, foram citados como fatores de contribuição para a melhoria da qualidade dos serviços em saúde a difusão da prática médica orientada por dados e os recentes avanços no diagnóstico e detecção precoce de doenças, potencializados pela medicina de precisão e o uso do *big data* na composição de conhecimentos médicos.

Adicionalmente, Bainbridge (2019) explica que, uma das maneiras de alcançar estas melhorias é utilizar o *big data* como um vasto acervo de amostras de dados populacionais, sendo esta, uma forma de subsidiar informações relevantes para as instituições promoverem estratégias de saúde inteligentes. Os dados de saúde são extraídos dos registros dos pacientes e as pesquisas populacionais fornecem dados fundamentais para compreender o perfil das pessoas que vivem em uma determinada região e suas condições de vida (KARPATHAKIS et al, 2021).

A integração destes dados permite determinar com maior precisão os tratamentos mais eficientes para os pacientes, respeitando as diferenças sociais e biológicas que tornam cada indivíduo único (BAINBRIDGE, 2019).

Segundo esta mesma linha de pensamento, os autores Paveena e MK Rao (2019) entendem que o *big data* na saúde se apresenta como uma ferramenta robusta para o armazenamento dos dados de saúde coletados, e seus benefícios podem ser potencializados pela fusão com outras tecnologias, como a Inteligência Artificial (IA) e o *Machine Learning*, utilizadas para maximizar a capacidade do cruzamento dos dados.

No entanto, Long et al (2021) ressaltam que, no setor de saúde os benefícios do *big data* vão além da criação de um grande repositório de dados médicos e reforçam esta afirmação ao lembrar que, durante o período mais crítico da pandemia do COVID-19, a integração de tecnologias como a IA e o *big data* foram ativos fundamentais para auxiliar a área médica a compreender o comportamento do vírus para conter os avanços da doença. O rastreio e a análise

do *big data* foram fundamentais para gerar novos conhecimentos sobre o vírus e, diante deste acontecimento, não se pode ignorar o fato de que os *insights* obtidos pela análise do *big data* apoiaram as decisões médicas e orientaram as práticas dos profissionais de saúde, que atuaram na linha de frente do combate à pandemia global (CARDONA, 2022).

Batko e Slezak (2022) e Zheng et al (2022) ressaltam que, a área de pesquisa em saúde também é apoiada pelo *big data*. Com a integração dos dados médicos, é possível reunir no *big data* os prontuários dos pacientes das redes públicas e privadas, viabilizando a realização de estudos direcionados para atender os pacientes com diferentes realidades e necessidades.

Em paralelo, o uso de ferramentas tecnológicas desenvolvidas para o cruzamento de grandes volumes de dados clínicos, registros médicos e biométricos capturados por sensores em tempo real, potencializam a velocidade com que os dados coletados são analisados e, consequentemente, na área de pesquisa em saúde, os conhecimentos extraídos do *big data* são utilizados para fomentar os estudos voltados para o tratamento de doenças complexas em diferentes cenários (ZHENG et al, 2022).

Outra área fortemente impactada pelo uso estratégico do *big data* é a de negócios e gestão de saúde, neste sentido é preciso ter em mente que a natureza do *big data* se transforma de acordo com o contexto em que os dados estão inseridos (PAIGE et al, 2021). No setor público, por exemplo, por meio dos resultados extraídos do *big data* é possível identificar as atividades e os procedimentos médicos duplicados ou desnecessários, diminuindo custos e mitigando o risco de desperdício de materiais, melhorando, assim, a gestão da verba pública destinada para as instituições de saúde (D'AGOSTINO et al, 2021).

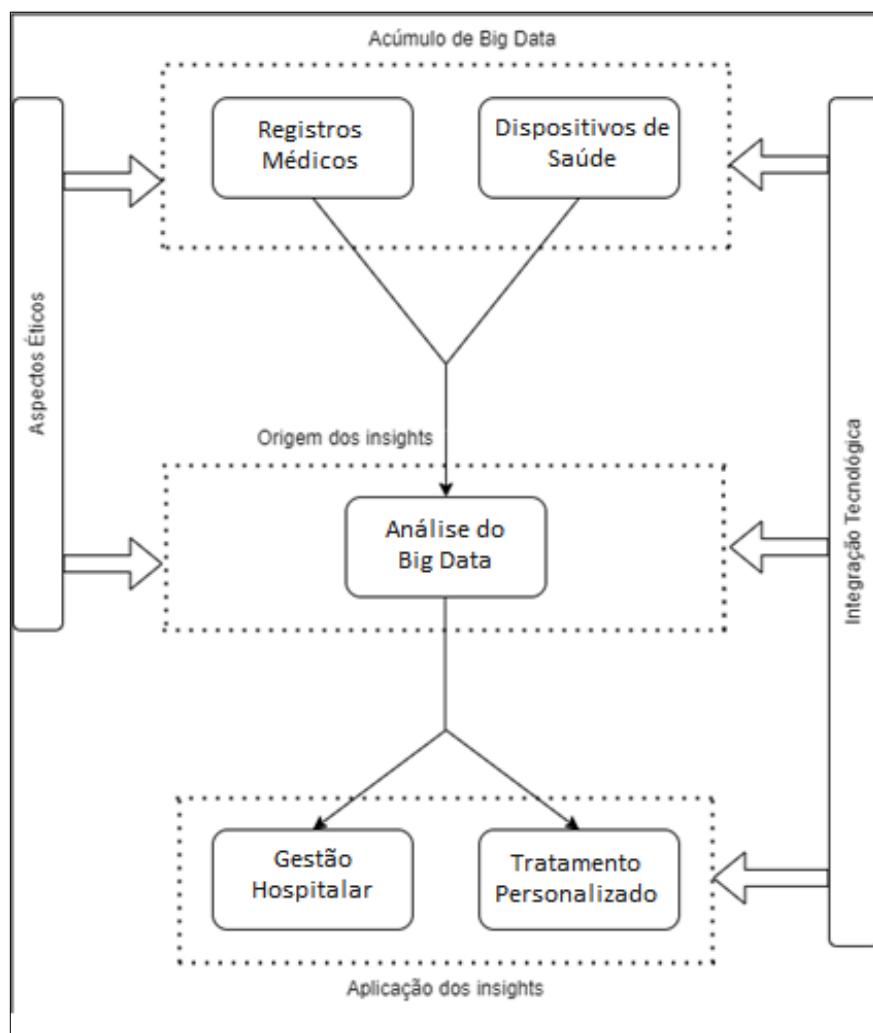
Na saúde privada, sob a perspectiva do aumento da lucratividade, os resultados extraídos da análise do *big data* ajudam a detectar pacientes cujo tratamentos geram altos custos para a instituição, além dos métodos de trabalho que requerem um maior aporte financeiro (JUDOO et al, 2021). Desse modo, as instituições podem lançar mão destas informações para oferecer aos profissionais e paciente soluções personalizadas, que beneficiem ambas as partes, diminuindo os custos e aumentando a lucratividade dos estabelecimentos de saúde privada (WESSON et al, 2022).

Em concordância com estas duas ideias, Wang et al (2018) complementam ao afirmar que, quando o dado bruto do *big data* é combinado com tecnologias de alta performance, informações de valor podem ser extraídas dos dados. Estes conhecimentos, podem ser aplicados pela gestão em diferentes tarefas relacionadas aos negócios de saúde, aumentando a eficiência e a qualidade das decisões.

De forma prática, os autores citaram como exemplo de aplicabilidade destes *insights* o desenvolvimento das capacidades dos profissionais da área, a alocação eficiente de recursos e o aumento do lucro das organizações (WANG et al, 2018). Estes fatores atrelados ao uso de tecnologias inteligentes baseadas em nuvem com alto poder de processamento, oferecem uma maior velocidade e escalabilidade para análise do *big data*, potencializando sua capacidade de gerar novas informações e *insights* de valor para os gestores das instituições de saúde (SAYANTAN et al, 2020).

Partindo do pressuposto de que a integração das tecnologias desempenha um papel fundamental na conversão das informações extraídas do *big data* em ações tangíveis na área da saúde, os autores Sayantan et al (2020) desenvolveram uma estrutura abrangente (Figura 3) para apresentar uma visão sistematizada do uso estratégico do *big data Analytics* na saúde

Figura 3- Estrutura para aplicação do *big data Analytics* na saúde.



Fonte: Elaboração própria (2023), adaptado de Sayantan et al (2020).

Sob a perspectiva de Sayantan et al (2020), a origem do *big data* tem como fonte os registros médicos e os dados médicos disponíveis em formatos eletrônicos, gerados pela troca de informações em tempo real e potencializados pela integração dos dispositivos de saúde, no qual os autores denominaram como dados de sensor. Neste sentido, os registros médicos atuam como blocos de construção de dados históricos no setor de saúde, principalmente para obter dados dos pacientes a partir dos relatórios de diagnóstico, registros hospitalares e histórico de atendimento dos pacientes (SAYANTAN et al, 2020; TARIQ et al, 2020; SILVENNOINEN, 2020).

Do ponto de vista dos aspectos éticos, uma das questões levantadas nas pesquisas realizadas por Xia et al (2020), Wu e Chen (2022) é a preocupação com a privacidade e segurança na manipulação dos dados dos pacientes. De acordo com os autores, é fundamental que as organizações de saúde que acumulam *big data* garantam que as diretrizes legais e éticas preservem a integridade dos dados sensíveis nas etapas de coleta, armazenamento e compartilhamento de informações de cunho pessoal (XIA et al, 2020; WU; CHEN, 2022). Desta forma, o gerenciamento do *big data* precisa estar em conformidade com a lei que rege a segurança dos dados pessoais de cada país (WONG et al, 2022), que, no caso do Brasil é Lei Geral de Proteção de Dados de 14 de agosto de 2018 (LGPD, 2018).

Na percepção de Rojas et al (2018), Jindal et al (2018) e Sayantan et al (2020), a integração de tecnologias e a análise dos dados, têm potencial para gerar *insights* confiáveis para auxiliar os gestores na alocação dinâmica dos recursos de saúde, fator fundamental para a adoção de práticas administrativas eficientes e que refletem positivamente no posicionamento estratégico das instituições de saúde, na esfera pública e privada (JINDAL et al, 2018).

Sob a ótica do serviço público, as instituições de saúde que utilizam os dados para nortear suas ações estratégicas, criam políticas mais assertivas e condizentes com a realidade em que está inserida (WERUTSKY et al, 2021; YUNITA et al, 2022). Desta forma, os gestores das organizações públicas de saúde podem utilizar as informações extraídas do *big data* para compreender as condições de vida e saúde dos indivíduos, tendo em vista que o direito à saúde passa pelas diferenciações sociais, respeitando a pluralidade regional, para que assim, seja possível garantir ao cidadão serviços de saúde que respeitam o princípio da equidade (CARDONA, 2022).

2.5. Desafios para implementação do *big data* na saúde

O setor de saúde sempre gerou grandes quantidades de dados e isso está relacionado, entre outros fatores, com a necessidade de armazenamento de prontuários do paciente. No entanto, seu armazenamento não é unificado em uma única base e nem segue um padrão de registros das informações (LIANG, 2021).

Shubham et al (2021), afirmam que um dos principais desafios analíticos para implementação do *big data* em uma instituição de saúde, se refere a qualidade dos dados que são originados de diferentes fontes, como os registros eletrônicos de saúde e dispositivos médicos interconectados. Ainda de acordo com os autores, estas fontes de dados frequentemente apresentam problemas de precisão e consistência dos registros, fazendo com que os dados apresentem inconsistências capazes de distorcer os resultados das análises e levando a conclusões equivocadas por parte dos profissionais de saúde (SHUBHAM et al, 2021). Desse modo, a falta de padronização e preparação dos dados de saúde coletados pelas instituições, são etapas críticas para que os profissionais de saúde possam obter resultados confiáveis e significativos da análise do *big data* (TALWAR et al, 2020).

Outro ponto de atenção levantado pelos autores estudados, está na necessidade de seleção apropriada de algoritmos e técnicas analíticas condizentes com os objetivos estratégicos da instituição. Al Mayahi (2019) explica que é fundamental que os profissionais responsáveis por gerenciar os dados de saúde, escolham algoritmos que sejam compatíveis com os tipos de dados e as questões clínicas específicas que estão sendo tratadas pela instituição. Além disso, a compreensão e interpretação dos modelos é crucial no contexto médico, pois os profissionais de saúde precisam entender como as decisões são tomadas para confiar nos resultados e aplicá-los de maneira adequada (YUNITTA, 2022).

A validação dos resultados da análise dos dados pelos profissionais de saúde é outro desafio que não pode ser subestimado. Embora os *insights* obtidos por meio das análises do *big data* sejam promissores, é fundamental que esses *insights* sejam validados pelos profissionais de saúde, para que desta forma seja possível afirmar que os conhecimentos extraídos do *big data*, de fato são aplicáveis no mundo real (SCHMARZO, 2018).

Também faz parte desta lista, a necessidade de um investimento inicial maciço para aquisição dos recursos (tecnológicos e humanos), necessários para viabilizar e potencializar os benefícios do *big data* no setor de saúde (AL MAYAHI, 2019). Autores como Shubham et al (2021) ressaltam as nuances deste processo no setor privado e público e afirmam que, ambos os contextos requerem uma análise cuidadosa dos custos e benefícios da adoção do *big data* no

setor de saúde, sempre considerando os objetivos estratégicos que a instituição espera alcançar com a adoção desta tecnologia (SHUBHAM et al, 2021).

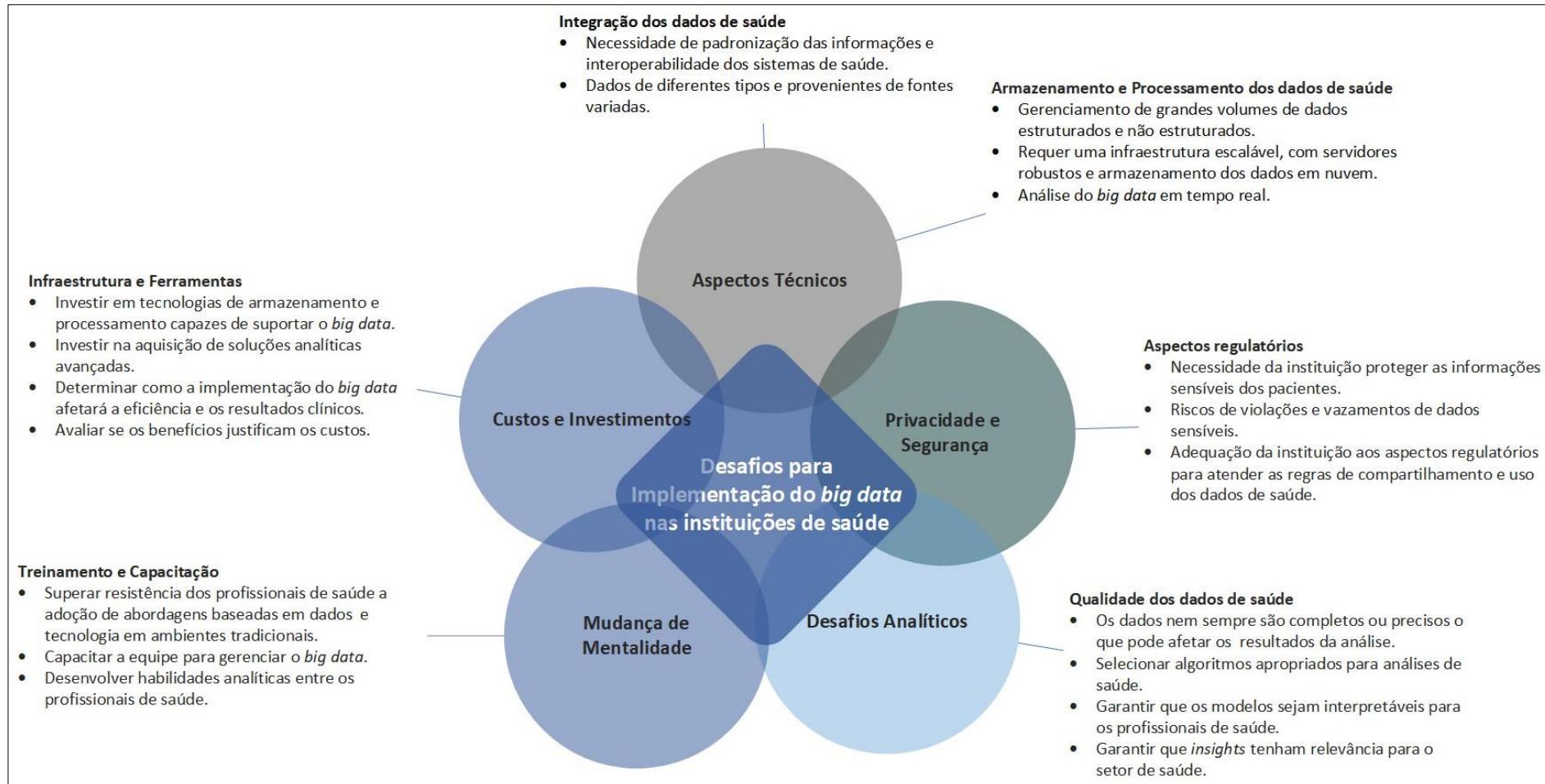
As instituições de saúde do setor privado possuem uma mentalidade orientada para o lucro, fator que incentiva a adoção de inovações para melhorar a eficiência operacional e aumentar a atração de pacientes para os seus estabelecimentos (WERUTSKY et al, 2021). Além disso, por vezes, as instituições privadas de saúde possuem uma maior flexibilidade dos recursos financeiros, o que permite a aquisição de uma infraestrutura de armazenamento, processamento e análise de dados robusta, sendo este um pré-requisito para o sucesso na implementação do *big data* (SHUBHAM et al, 2021).

No entanto, o setor privado não está isento de desafios para implementação do *big data*. Sob a perspectiva do retorno sobre o investimento, lógica predominante no setor privado de saúde, os custos associados à implementação, treinamento de pessoal e manutenção da infraestrutura precisam estar diretamente associados aos benefícios tangíveis e intangíveis que poderão ser alcançados, como a melhoria na experiência do paciente e diferenciação no mercado, de modo a justificar os altos investimentos para aquisição das tecnologia necessárias para implementação do *big data* na instituição (CARDONA, 2022).

Ao contrário do setor privado, o foco principal das instituições de saúde pública não é o lucro, mas sim o bem-estar da população (LIANG, 2021). Isso significa que os investimentos são frequentemente justificados pelo potencial do uso das tecnologias para melhorar os serviços de saúde prestados à população em grande escala, reduzindo os custos dos tratamentos a longo prazo e otimizando os recursos disponíveis (SAYANTAN, 2020). No entanto, a obtenção de financiamento contínuo pode ser um desafio para as instituições do setor público de saúde, exigindo a demonstração contínua dos benefícios gerados pelo uso do *big data* na instituição (SHUBHAM et al, 2021).

Estudiosos como Sayantan (2020), Liang (2021), e Yunitta (2022) alertam sobre a necessidade de cuidados com a segurança e privacidade dos dados dos pacientes, especificamente com relação a extração, manipulação e compartilhamento não autorizado de dados sensíveis. A implementação bem-sucedida do *big data* na saúde requer um compromisso contínuo da instituição com a privacidade e segurança dos dados coletados, de modo a garantir que os avanços tecnológicos sejam alcançados sem comprometer a confidencialidade dos dados dos pacientes (LIANG, 2021). A Figura 4 buscou apresentar de forma sistematizada o que na visão dos autores que compõem este tópico, podem ser considerados como os principais desafios para implementação do *big data* nas instituições de saúde.

Figura 4 - Desafios para a Implementação do *big data* na Saúde.



Fonte: Elaboração própria (2024), com base em Sayantan (2020), Liang (2021), Shubham et al, 2021 e Yunitta (2022).

Devido a sua faceta disruptiva, o entendimento sobre o potencial do *big data* pode não ficar claro para os profissionais da saúde, e esta falta de clareza, pode gerar resistências quanto a sua aceitação (TALWAR et al, 2020; SAYANTAN, 2020). Desta forma, é fundamental que os setores de tecnologia, gestão hospitalar e demais alas administrativas, se esforcem mutuamente para conscientizar os profissionais sobre como o uso estratégico do *big data* pode agregar valor para o setor de saúde, sendo esta, uma ação tão necessária quanto preparar a infraestrutura para dar suporte operacional para sua implementação (LIANG, 2021).

2.6. Projeções futuras para o uso do *big data*

Com base nos estudos selecionados neste trabalho foram identificadas duas oportunidades auspiciosas para potencializar os benefícios da adoção do *big data* no setor de saúde: *Machine Learning* (ML) e Inteligência Artificial (IA), que compreendem métodos que permitem que os computadores executem tarefas geralmente associadas à inteligência humana, de forma automatizada (MANREET et al, 2021). Os conceitos de ML e IA são facilmente confundidos, devido à similaridade dos termos. No entanto, para perceber os impactos do seu uso no setor de saúde, é importante delimitar seus conceitos (WERUTSKY, 2021).

Tecnicamente, o *Machine Learning* é considerado uma subcategoria da Inteligência Artificial, mas apesar disso, conceitualmente é compreendido como um conjunto de técnicas desenvolvidas para aprender por meio do processamento de uma grande massa de dados, de maneira automatizada e em constante evolução (MANYAZEWAL et al, 2021). Em contrapartida, Inteligência Artificial pode ser definida como o ramo da ciência da computação que desenvolve sistemas que simulam a capacidade de aprendizagem humana para a resolução de um problema (SUPRIYA, 2021).

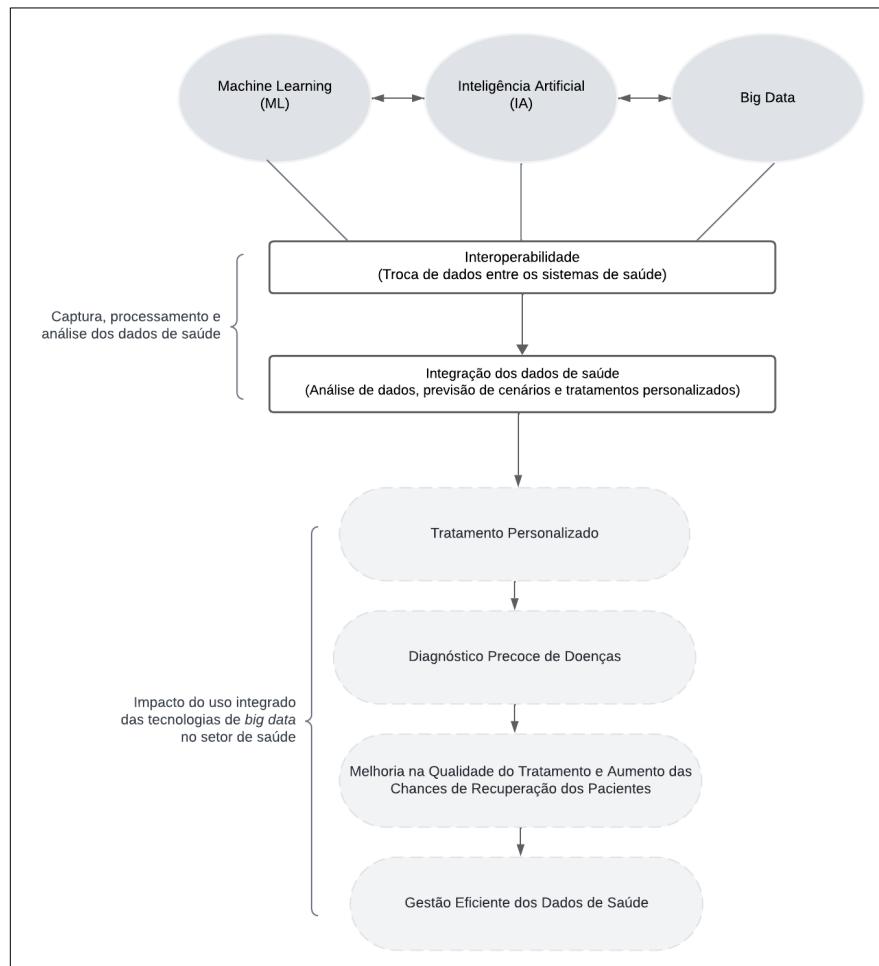
Quando aplicados no setor de saúde, a Inteligência Artificial e o *Machine Learning* utilizam as máquinas para definir os algoritmos adequados para a análise do *big data*, com o objetivo de identificar os componentes do problema, suas possíveis soluções e ao final, gerar *insights* e tomar decisões com precisão, sem a necessidade de intervenção humana (KIM; CHUNG, 2019). Estes recursos tecnológicos tem como principais características a flexibilidade e interoperabilidade, que no setor de saúde significa a troca de dados simultânea e em alta velocidade entre os diferentes sistemas de saúde, gerando mais informações e relações significativas para o cuidado do paciente ((MANREET et al, 2021; (CHIEN-LUNG; CHI-CHANG, 2020).

No estudo intitulado “*Spatial Clusters of Cancer Mortality in Brazil: A Machine Learning Modelling Approach*”, os autores Teixeira et al (2021) utilizaram o modelo de *Machine Learning* para identificar o percentual de pacientes diagnosticados com câncer em um município do Rio Grande do Sul, com o objetivo de apresentar dados estatísticos atualizados e as possíveis causas para o aumento do número de mortes por câncer na região. A partir da análise preditiva e do cruzamento dos dados clínicos e demográficos dos pacientes diagnosticados com a doença, foi possível identificar os bairros que registraram um índice de mortalidade acima do esperado, além da identificação dos dez tipos mais frequentes de câncer por localidade.

A implicação prática da pesquisa de Teixeira (2021) e seus colaboradores foi a identificação de possíveis causas para o aumento do número de mortes dos pacientes com câncer de esôfago na região oeste do Rio Grande do Sul, enquanto na região norte o câncer de estômago foi o mais nocivo, com um aumento do número de mortes maior do que na região oeste. Do ponto de vista da gestão pública de saúde, estas informações são relevantes para a elaboração do projeto de política pública da saúde local (TEIXEIRA et al, 2021).

Diante dos impactos das inovações tecnológicas na área da saúde, autores como Paul et al (2022) e Werutsky et al (2021) acreditam que a adoção do *Machine Learning* e da Inteligência Artificial combinadas com o uso estratégico do *big data* são capazes de gerar conhecimentos relevantes para a área da saúde, uma vez que, por meio deste conjunto de tecnologias é possível prever a propagação de doenças, promover tratamentos personalizados, aumentar a precisão dos diagnósticos e, principalmente, aumentar a qualidade do tratamento e as chances de recuperação dos pacientes (WONG et al, 2022), conforme apresentado na Figura 5.

Figura 5 - Integração do *machine learning*, inteligência artificial e *big data* e seus benefícios para as práticas em saúde.



Fonte: Elaboração própria (2024) com base em Werutsky et al (2021), Paul et al (2022) e Wong et al (2022).

2.7. Possíveis caminhos para implementação do *big data* em uma instituição de saúde

A adoção do *big data* em uma instituição de saúde requer, entre outros fatores, ter um conjunto organizado de diretrizes e ações planejadas para implementar a tecnologia no setor (DAUGHTON E PAUL, 2021). Esse plano envolve a definição de objetivos claros, a identificação dos recursos necessários, o estabelecimento de metas mensuráveis e a criação de um plano de ação para guiar o processo de implementação do *big data* na instituição (DREDZE et al, 2019).

De acordo com os autores Adler-Milstein e Jha (2017), Brownstein e Freifeld (2018), Dredze et al (2019), Kahn (2020) e Daughton e Paul (2021), a elaboração de um plano estratégico para implementação bem sucedida do *big data* em uma instituição de saúde inclui as seguintes etapas:

1. Definir objetivos e metas condizentes com o plano estratégico da instituição: O primeiro passo para implementação do *big data* no setor de saúde é fornecer uma direção sólida para o planejamento e execução do projeto de implementação do *big data* na instituição, tendo como base os objetivos estratégicos definidos a médio e longo prazo (DAUGHTON e PAUL, 2021). Neste sentido, o plano de ação deverá ter como meta aprimorar a tomada de decisões clínicas e organizacionais do setor de saúde, otimizar a eficiência operacional da instituição, melhorar a qualidade do serviço prestado ao paciente e impulsionar a pesquisa médica (BROWNSTEIN e FREIFELD, 2018).
2. Avaliar a infraestrutura tecnológica da instituição: É importante que os profissionais envolvidos no projeto de implementação do *big data*, avaliem a infraestrutura tecnológica existente na instituição. Isso inclui verificar a capacidade de armazenamento, processamento e integração de dados, bem como a interoperabilidade dos sistemas de informação de saúde (ADLER-MILSTEIN e JHA, 2017). Complementando esta ideia, Topol (2019) reforça a importância de identificar os equipamentos obsoletos que compõem o parque tecnológico para que, a partir disso, seja possível prover uma renovação tecnológica com base em um planejamento estratégico, sendo este um fator crucial para que a instituição esteja preparada para lidar com o volume e a complexidade do *big data* (TOPOL, 2019).

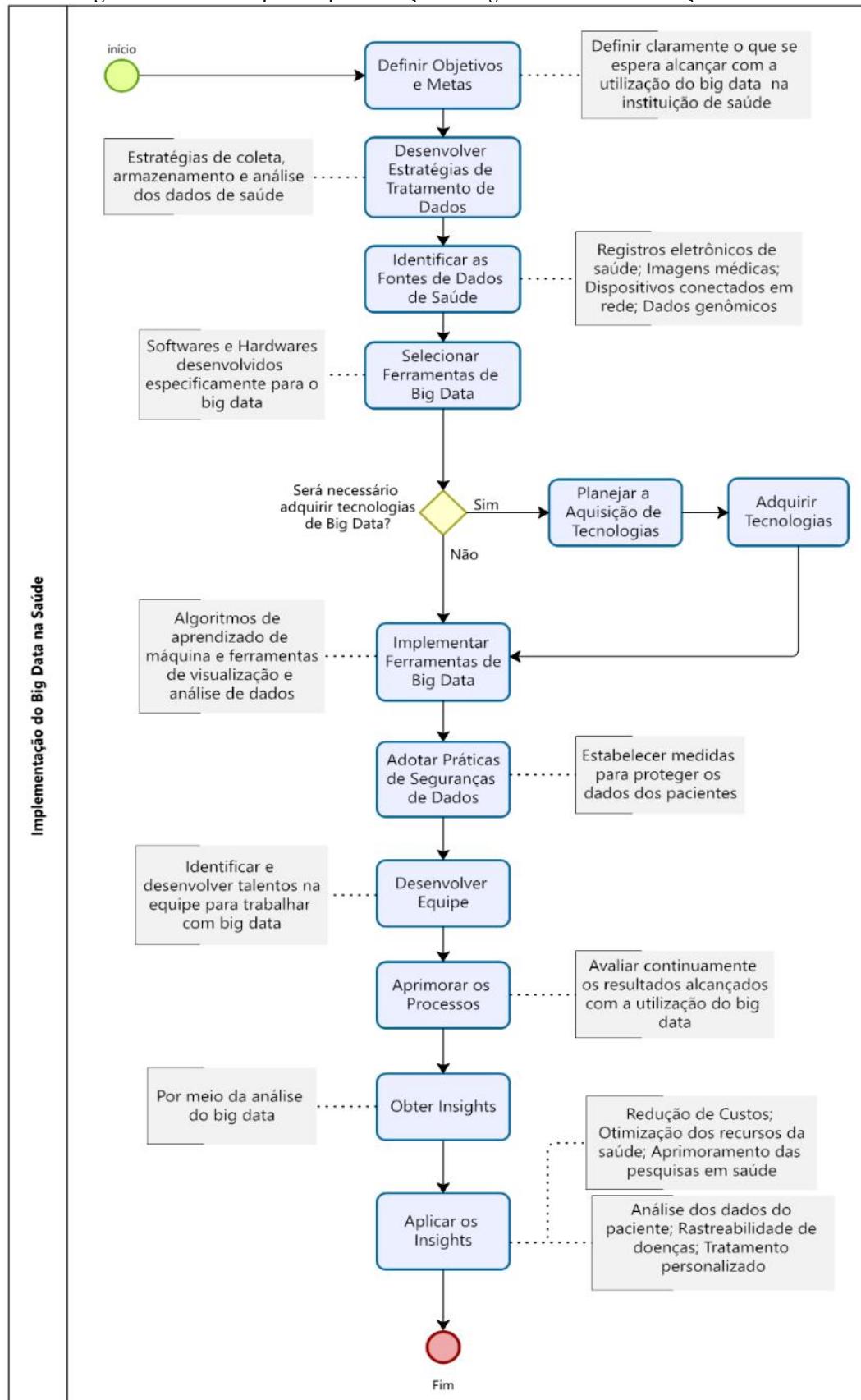
3. Coletar e integrar os dados de saúde: É fundamental que a instituição que está se preparando para a implementação do *big data*, tenha as suas fontes de dados de saúde bem definidas (DAUGHTON e PAUL, 2021). Para criar essa estrutura, é necessário identificar as fontes de dados disponíveis na instituição, como registros eletrônicos de saúde, dados de dispositivos médicos, informações genômicas e dados de pesquisa clínica (SAYANTAN et al, 2020). Em seguida, será necessário desenvolver um processo minucioso de integração dos dados de saúde, de modo a criar um ambiente propício para a análise do *big data* (LOTFI, 2023).
4. Garantir a qualidade dos dados de saúde coletados: A qualidade dos dados é fundamental para obter insights confiáveis e significativos do *big data*. Sendo assim, será necessário que a instituição de saúde implemente processos e protocolos rigorosos para garantir a integridade, precisão e consistência dos dados coletados e esse processo envolve padronizar os formatos de registro dos dados do paciente, eliminar as informações redundantes ou inconsistentes e aplicar técnicas de higienização e validação dos dados de saúde (KHAN, 2020).
5. Aplicar técnicas de análise do *big data*: A aplicação de técnicas avançadas de análise e mineração dos dados é uma etapa importante para extrair conhecimentos do *big data* (SENTHILKUMAR et al, 2018). Isso inclui o uso de algoritmos de aprendizado de máquina, inteligência artificial e análise preditiva para identificar padrões, tendências e criar insights relevantes para a área da saúde (SAYANTAN, 2020; LOTFI, 2023).
6. Adotar práticas de segurança dos dados sensíveis: Garantir a privacidade e segurança dos dados dos pacientes é um aspecto crítico na implementação do *big data* no setor de saúde. Isso requer implementar práticas adequadas para o tratamento dos dados sensíveis, além do comprometimento da instituição em implementar mecanismos de segurança dos dados, com o objetivo de minimizar o risco de pessoas não autorizadas obterem acesso aos dados do paciente (LYANG, 2021).
7. Criar grupos multidisciplinares para elaboração e manutenção do projeto de implementação do *big data*: Uma forma de minimizar o risco de resistência dos

profissionais de saúde quanto ao uso do *big data* no setor é criar comitês ou grupos de trabalho multidisciplinares que envolvam profissionais da saúde, especialistas em tecnologia, pesquisadores e administradores (SUPRIYA 2021). Esta iniciativa, desempenha um papel fundamental no sucesso da implementação, uma vez que permite que a tecnologia seja percebida sobre diferentes perspectivas. Além disso, a troca de experiências entre os profissionais, pode facilitar a identificação de desafios e a busca por soluções para os empecilhos que surgirem ao longo do processo de implementação (BUCKERIDGE et al, 2018).

8. Aprimorar o processo continuamente: A implementação do *big data* em uma instituição de saúde é um processo contínuo. É importante realizar avaliações regulares para medir o impacto das iniciativas implementadas, identificar oportunidades de melhoria e ajustar a estratégia conforme necessário (KHAN, 2020). A coleta de feedback dos profissionais de saúde e dos usuários finais é essencial para entender os impactos do uso do *big data* e direcionar de forma assertiva os próximos passos (SINGH et al, 2021).

Buscando sintetizar os caminhos que podem ser adotados por uma instituição de saúde para implementação do *big data*, a Figura 6 apresenta as etapas deste processo, de acordo com a perspectiva dos autores que compõem esta sessão.

Figura 6 - Caminhos para implementação do *big data* em uma instituição de saúde.



Fonte: Elaboração própria (2024).

No contexto da saúde pública, os autores Daughton e Paul (2020) explicam que a implementação do *big data* pode envolver algumas etapas adicionais e necessidades de ajustes específicos em comparação a uma instituição de saúde privada. Estes autores afirmam que, as instituições de saúde pública geralmente trabalham em estreita colaboração com outras entidades, como agências governamentais, organizações não governamentais e instituições de pesquisa, sendo necessário criar políticas de incentivo ao compartilhamento dos dados, conhecimentos e recursos tecnológicos, sendo estes desafios que precisam ser superados para que a adoção do *big data* tenha sucesso no setor público de saúde (DAUGHTON e PAUL, 2020).

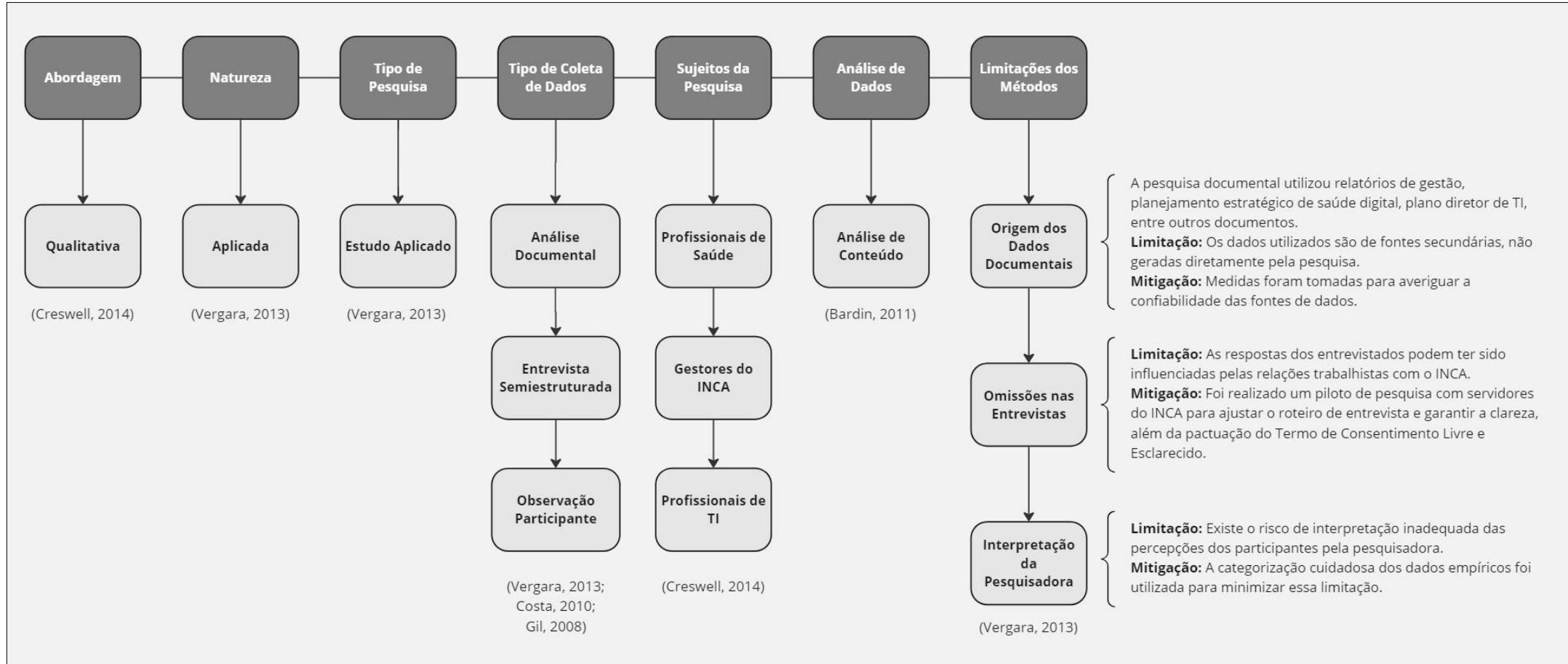
Dredze et al (2019), alertam que, obter uma visão abrangente da saúde da população requer o uso de dados provenientes de diversas fontes como os registros de saúde, dados demográficos, informações de vigilância epidemiológica e dados de saúde ambiental. Dessa forma, a integração e a interoperabilidade dos sistemas de saúde se tornam ainda mais críticas em uma instituição pública de saúde que pretende implementar o *big data* para auxiliar na tomada de decisões estratégicas (DREDZE et al, 2019).

Por fim, Khan (2020) e Lotfi (2023) explicam que a implementação do *big data* em uma instituição de saúde pública não é um processo estático, mas sim contínuo e iterativo, sendo necessário realizar avaliações regulares dos resultados obtidos pelo uso do *big data* para garantir que as metas e objetivos estratégicos da instituição estejam sendo alcançados (LOTFI, 2023). Dessa forma, o processo de implementação do *big data* deve ser constantemente revisado e ajustado à medida que novos desafios surgem, as necessidades do setor de saúde se modificam e a tecnologia avança (KHAN, 2020).

3. METODOLOGIA

Este capítulo aborda o delineamento adotado neste estudo, os sujeitos e critérios de seleção de atores para a entrevista, o método adotado para coleta e análise dos dados e as limitações dos métodos escolhidos. A Figura 7 apresenta as etapas metodológicas adotadas nesta pesquisa aplicada.

Figura 7 - Diagrama metodológico da pesquisa.



Fonte: Elaboração própria (2024).

3.1 Caracterização da Pesquisa

Este projeto se constitui em uma pesquisa aplicada, visando a geração de conhecimentos práticos para a elaboração de um relatório técnico conclusivo com recomendações estratégicas para implementação do *big data* no INCA.

A abordagem do problema é realizada a partir de uma pesquisa qualitativa, baseada na premissa de que as pesquisas qualitativas utilizam estruturas que abarcam as percepções que os indivíduos possuem sobre determinado problema, tendo em suas conclusões as concepções dos participantes e do pesquisador, a interpretação do problema e a proposta de mudança (CRESWELL, 2014). Por meio de uma pesquisa exploratória, buscou-se adquirir familiaridade com o problema para entendê-lo e explicitá-lo de maneira clara e facilitar o alcance dos objetivos.

Quanto aos meios, foi realizada uma pesquisa bibliográfica por meio de um mapeamento sistemático da literatura para o levantamento de artigos científicos nas bases de periódicas *Scopus*, *Web of Science* e *IEEE Xplore*, bem como pesquisas em sites e demais fontes que pudessem embasar o tema.

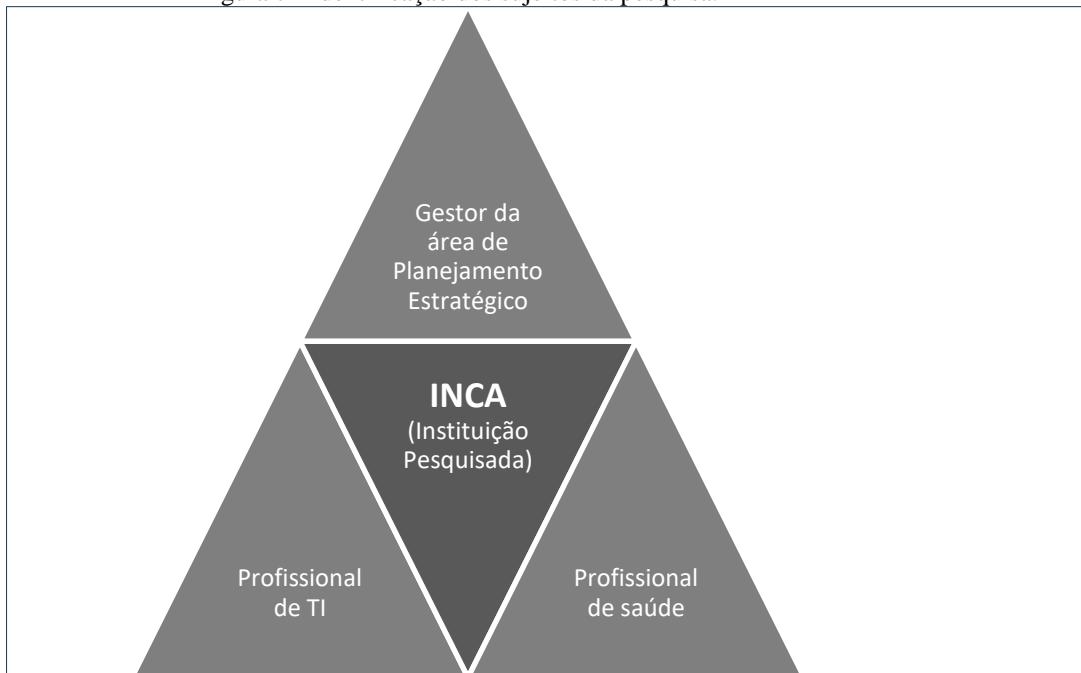
Como procedimento metodológico, optou-se pelo estudo aplicado. Esta escolha se deu pela necessidade de gerar conhecimentos e soluções práticas para o mundo real. Essa abordagem de pesquisa visa direcionar os esforços para a aplicação imediata dos resultados obtidos, com o objetivo de resolver questões específicas ou melhorar determinados processos ou práticas (VERGARA, 2013).

3.2 Seleção dos Sujeitos da Pesquisa

Para a seleção dos sujeitos da investigação foi adotado o critério da acessibilidade, uma vez que a pesquisadora conta com o apoio do Setor de Tecnologia do INCA (SETI) para a elaboração da pesquisa, o que facilitou a identificação e seleção de indivíduos habilitados a fornecer as informações necessárias para a avaliação e elaboração do diagnóstico para atender o objetivo da pesquisa.

Outro critério adotado na seleção dos sujeitos da pesquisa foi a técnica do “*snowball*” que permite a seleção de novos respondentes por meio das redes de referência dos participantes (VINUTO, 2014), de modo que os profissionais entrevistados pudessem atender os seguintes requisitos: (1) ser gestor de áreas responsáveis pelo planejamento estratégico do INCA; (2) ser profissional atuante da área da saúde; (3) ser profissional da área de tecnologia da informação.

Figura 7 - Identificação dos sujeitos da pesquisa.



Fonte: Elaboração própria (2024).

3.3 Coleta de Dados e Plano Piloto

As técnicas para coleta de dados utilizadas na pesquisa seguiram as recomendações referentes ao perfil de levantamento de dados propostas por Gil (2008), que menciona a utilização de mais de uma técnica de coleta em estudos de caso. Desta forma, a coleta de dados foi feita por meio do levantamento bibliográfico, para uma melhor compreensão do *corpus* em estudo, bem como da análise documental da Instituição, a qual a pesquisadora possui livre acesso.

Como instrumento de coleta de dados (ICD) foi elaborado um roteiro de entrevista semiestruturado, que permite flexibilidade ao entrevistador no momento de sua execução, pois, com um roteiro de perguntas pré-determinadas, dependendo do rumo que a entrevista tome, novos questionamentos podem ser realizados, mesmo que não previstos no roteiro (FLICK, 2009). As observações por parte da pesquisadora foram realizadas de maneira participante durante as entrevistas. Tal escolha permitiu o acesso às informações necessárias de forma diferenciada, além de permitir a captação da realidade por meio da perspectiva dos profissionais internos da instituição (YIN, 2015).

O ICD foi criado com base na Fundamentação Teórica, com o objetivo de

compreender a percepção dos entrevistados sobre as estratégias que podem ser adotadas pelo Instituto Nacional de Câncer (INCA) para a implementação do *big data* na instituição. O roteiro de entrevista inicialmente desenhado está descrito no Quadro 3.

Quadro 3 - Apresentação da versão inicial do roteiro de entrevista semiestruturado.

1. Apresentação da Pesquisa, da pesquisadora e o objetivo da entrevista.
2. Qual atividade você desempenha no INCA e qual a sua experiência na área da saúde?
3. De maneira geral, como o *big data* pode ser usado para melhorar a qualidade dos cuidados de saúde no INCA?
4. Na sua percepção, quais são os principais desafios e oportunidades para a implementação do *big data* no INCA?
5. De maneira geral, como você avalia os impactos que a adoção do *big data* no INCA teria para o desenvolvimento do seu trabalho?
6. Na sua percepção, existe uma padronização para registrar eletronicamente os dados dos pacientes no INCA?
7. Poderia citar as fontes de dados dos pacientes disponíveis no INCA?
8. Na sua percepção, o INCA possuí os requisitos tecnológicos necessários para implementar com sucesso o uso do *big data* no instituto? Em caso de resposta negativa, poderia por favor citar os motivos que te fizeram chegar a essa conclusão?
9. De maneira geral, você considera que a implementação do *big data* no instituto traria desafios éticos e legais quanto a coleta de dados dos pacientes? Caso a resposta seja positiva, poderia apontar os principais desafios e como eles poderiam ser superados?
10. Tendo em vista que o INCA é uma instituição que atua no desenvolvimento e coordenação de ações integradas para a prevenção e controle do câncer no

Brasil, na sua percepção, quais seriam os impactos que a implementação *do big data* traria para o Instituto?

11. De maneira geral, você considera que a adoção do *big data* no INCA poderia auxiliar na tomada de decisões estratégicas da instituição? Caso a resposta seja positiva, poderia por favor, explicar os motivos que te fizeram chegar a essa conclusão?
12. De maneira geral, você considera que a adoção do *big data* no INCA poderia auxiliar na tomada de decisões clínicas dos profissionais de saúde? Caso a resposta seja positiva, poderia por favor, explicar os motivos que te fizeram chegar a essa conclusão?

*Tempo médio para elaboração da entrevista: Estima-se o período de 45 minutos por entrevista.

Fonte: Elaboração própria (2024).

Com o propósito de garantir resultados assertivos no desenvolvimento da pesquisa de campo foi feita a validação deste instrumento através de um plano piloto. Os potenciais sujeitos participantes da pesquisa foram convidados via e-mail institucional para participar da entrevista. Cada solicitação foi precedida pelo envio do Termo de Consentimento Livre e Esclarecido (TCLE), elaborado de acordo com a Resolução 466/12, capítulo IV do Conselho Nacional de Saúde. O roteiro de entrevista detalhado, encontra-se no Apêndice B deste estudo, enquanto o TCLE, pode ser visualizado no Anexo A.

O teste piloto foi aplicado a três profissionais atuantes da área da saúde, sendo um gestor da área de Planejamento Estratégico, um profissional da área de Tecnologia da Informação e um profissional da área da saúde, com o objetivo de auxiliar na validação do instrumento de pesquisa desenhado, de modo a corrigir possíveis falhas e ambiguidades no roteiro, antes de aplicá-lo aos sujeitos selecionados para a pesquisa de campo, conforme descrito no tópico 3.2 desta seção.

Das percepções dos entrevistados sobre os instrumentos de pesquisa foram sugeridas as seguintes adequações para o roteiro inicialmente elaborado:

Quadro 4 - Adequações sugeridas pelos avaliadores do plano piloto.

Sugestão do avaliador	Parecer do Pesquisador
<p>De acordo com um dos avaliadores, seria oportuno incluir uma pergunta de aprofundamento após a discussão dos desafios e oportunidades para a implementação do <i>big data</i> no INCA, permitindo, assim, explorar com um maior detalhamento as percepções dos entrevistados sobre as possíveis soluções e estratégias que podem ser adotadas pela instituição para superar os desafios identificados.</p>	<p>A adição da pergunta de aprofundamento após a discussão dos desafios e oportunidades é pontual, e irá permitir uma análise mais completa das estratégias que podem ser adotadas pelo INCA para superá-los.</p> <p>Desta forma, será incluída a seguinte pergunta, após a pergunta de número 3 da versão inicial do roteiro de entrevista:</p> <p>“Considerando os desafios mencionados, como você sugere que o INCA possa superá-los, aumentando assim as chances de uma implementação bem-sucedida do <i>big data</i> no instituto? Por favor, você poderia compartilhar alguma experiência anterior ou exemplos de melhores práticas que poderiam ser aplicadas neste contexto?”</p>
<p>Um dos avaliadores sugeriu a inclusão de uma pergunta sobre o processo de capacitação e treinamento dos profissionais do INCA, para lidar com o uso do <i>big data</i> na rotina da instituição, uma vez que, na sua percepção “este ponto será importante para garantir que a equipe esteja adequadamente preparada para utilizar as ferramentas e técnicas de análise de dados de forma eficaz.”</p>	<p>Diante da observação do avaliador, a seguinte pergunta será incluída no roteiro de entrevista:</p> <p>“Como você percebe o processo de capacitação e treinamento dos profissionais do INCA para lidar com o uso do <i>big data</i>? Você acredita que há recursos e programas adequados para preparar a equipe para utilizar as ferramentas e técnicas de análise de dados de forma eficaz?”</p> <p>A pergunta acima, será incluída após a pergunta sobre os requisitos tecnológicos (pergunta 8) do instrumento inicial, pois está relacionada à preparação da equipe para lidar com as ferramentas e tecnologias necessárias para a implementação do <i>big data</i> na instituição.</p>
<p>Um dos entrevistados sugeriu aprofundar a discussão sobre os</p>	<p>Diante deste parecer, cabe ressaltar que os impactos éticos e legais da implementação do <i>big data</i> no INCA já havia</p>

Sugestão do avaliador	Parecer do Pesquisador
questionamentos éticos e legais que podem surgir com a implementação do <i>big data</i> no INCA, uma vez que os dados sensíveis dos pacientes serão manipulados em grande escala.	sido contemplado no roteiro proposto, através da pergunta sobre os desafios éticos e legais relacionados à coleta de dados dos pacientes (pergunta 10). Essa pergunta aborda diretamente a preocupação com questões éticas e legais, permitindo que os entrevistados compartilhem suas percepções sobre os desafios nesse aspecto e como eles poderiam ser superados.

Fonte: Elaboração própria (2024).

Após analisar as sugestões dos sujeitos participantes do plano piloto, o instrumento desenhado para coleta de dados foi melhorado, constituindo a versão adotada na pesquisa campo e descrita no Quadro 5.

Quadro 5 - Apresentação do roteiro de entrevista semiestruturado após a aplicação do plano piloto.

1. Qual atividade você desempenha no INCA e qual a sua experiência na área da saúde?
2. De maneira geral, como o *big data* pode ser usado para melhorar a qualidade dos cuidados de saúde no INCA?
3. Na sua percepção, quais são os principais desafios e oportunidades para a implementação do *big data* no INCA?
4. Considerando os desafios mencionados, como você sugere que o INCA possa superá-los, aumentando assim as chances de uma implementação bem-sucedida do *big data* no instituto? Por favor, você poderia compartilhar alguma experiência anterior ou exemplos de melhores práticas que poderiam ser aplicadas neste contexto?
5. De maneira geral, como você avalia os impactos que a adoção do *big data* no INCA teria para o desenvolvimento do seu trabalho?
6. Na sua percepção, existe uma padronização para registrar eletronicamente os dados dos pacientes no INCA?

7. Poderia citar as fontes de dados dos pacientes disponíveis no INCA?
8. Na sua percepção, o INCA possuí os requisitos tecnológicos necessários para implementar com sucesso o uso do *big data* no instituto? Em caso de resposta negativa, poderia por favor citar os motivos que te fizeram chegar a essa conclusão?
9. Como você percebe o processo de capacitação e treinamento dos profissionais do INCA para lidar com o uso do *big data*? Você acredita que há recursos e programas adequados para preparar a equipe para utilizar as ferramentas e técnicas de análise de dados de forma eficaz?
10. De maneira geral, você considera que a implementação do *big data* no instituto traria desafios éticos e legais quanto a coleta de dados dos pacientes? Caso a resposta seja positiva, poderia apontar os principais desafios e como eles poderiam ser superados?
11. Tendo em vista que o INCA é uma instituição que atua no desenvolvimento e coordenação de ações integradas para a prevenção e controle do câncer no Brasil, na sua percepção, quais seriam os impactos que a implementação do *big data* traria para o Instituto?
12. De maneira geral, você considera que a adoção do *big data* no INCA poderia auxiliar na tomada de decisões estratégicas da instituição? Caso a resposta seja positiva, poderia por favor, explicar os motivos que te fizeram chegar a essa conclusão?
13. De maneira geral, você considera que a adoção do *big data* no INCA poderia auxiliar na tomada de decisões clínicas dos profissionais de saúde? Caso a resposta seja positiva, poderia por favor, explicar os motivos que te fizeram chegar a essa conclusão?

*Tempo médio para elaboração da entrevista: Estima-se o período de 60 minutos por entrevista.

Fonte: Elaboração própria (2024).

Destaca-se que o estudo de campo foi realizado após a aprovação do projeto de

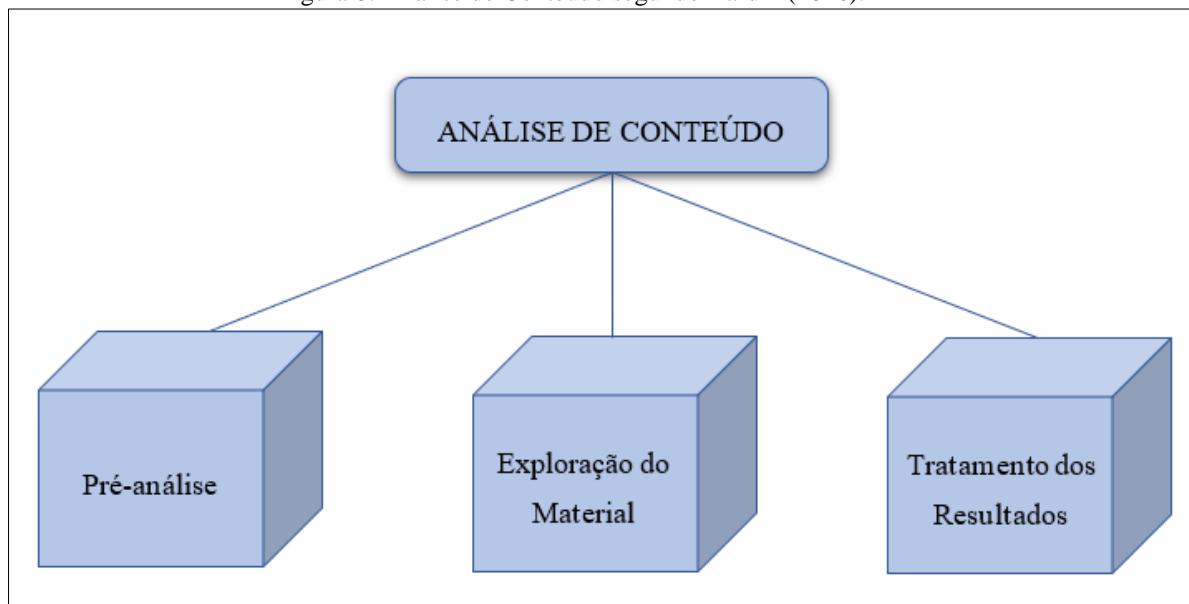
pesquisa pelo Comitê de Ética em Pesquisa da Plataforma Brasil, a qual ocorreu em novembro de 2023. As entrevistas estão protocoladas na CONEP (Comissão Nacional de Ética em Pesquisa) através do seguinte número CAAE (Certificado de Apresentação de Apreciação Ética): 71597923.3.0000.0311.

O Termo de Anuência, foi encaminhado ao gestor do setor de Tecnologia da Informação do INCA para assinatura e o documento assinado encontra-se no Anexo B deste projeto.

3.4 Análise dos Dados

Na análise de dados foram utilizados a análise documental e a análise do conteúdo pautada nos apontamentos defendidos por Bardin (2016, p.125). Quanto ao método e a organização da análise, estas tiveram como premissas três pontos básicos: (1) pré-análise, (2) exploração do material e; (3) categorização e tratamento dos resultados, conforme apresentado na Figura 8:

Figura 8: Análise de Conteúdo segundo Bardin (2016).



Fonte: Elaboração própria (2024), baseada em Bardin (2016).

Etapa 1 - Pré-Análise

Na etapa de pré-análise, os dados foram preparados e contextualizados por meio da revisão da literatura relevante para o tema de pesquisa. Além disso, foi realizada a análise dos documentos institucionais, como relatórios de gestão do INCA e o Plano

Estratégico do INCA, visando-se obter uma visão abrangente da instituição estudada.

Durante a revisão da fundamentação teórica, foram identificados conceitos e temas que permitiram uma compreensão mais aprofundada das questões relacionadas a adoção de tecnologias como o *big data* em instituições de saúde. A análise dos documentos institucionais, proporcionaram a obtenção de uma base contextual e acesso a informações adicionais importantes sobre o planejamento estratégico do INCA para o período de 2020 à 2023.

Etapa 2 - Exploração do Material

Nesta etapa, foi realizada a leitura iterativa e imersiva dos dados coletados nas entrevistas realizadas com os sujeitos de pesquisa no teste piloto. As entrevistas foram gravadas e transcritas em texto, para que, posteriormente fosse possível explorar o material coletado, identificando trechos relevantes para responder o problema de pesquisa proposto. As primeiras impressões e ideias iniciais foram registradas para orientar a análise posterior.

Etapa 3 - Categorização e Tratamento dos Resultados

Nesta fase, as declarações dos entrevistados foram organizadas em categorias com base na frequência de repetição de termos comuns e temas emergentes. Além disso, a estruturação das categorias foi fundamentada no referencial teórico, utilizando como referência principal o estudo de Sayanta et al. (2020), “*Big data analytics in healthcare: a systematic literature review*”, que serviu como base para identificar temas relevantes para a discussão em estudo. As categorias listadas no Quadro 6, foram ajustadas de acordo com as conexões conceituais identificadas na pesquisa, que utilizou uma abordagem analítica detalhada para mapear as declarações dos entrevistados às categorias conceituais.

Quadro 6 - Categorização dos resultados.

Categoria	Título	Descrição
1	Desafios e oportunidades para implementação do <i>big data</i> no INCA.	Esse tema teve como objetivo identificar quais serão os possíveis desafios para implementação do <i>big data</i> no INCA e as

		oportunidades que o uso do <i>big data</i> poderá gerar para a instituição.
2	Benefícios potenciais da adoção do <i>big data</i> no INCA.	Aqui, buscou-se compreender os benefícios específicos que o uso do <i>big data</i> poderá proporcionar para o INCA, enquanto instituição, e para a atividade realizada diretamente pelo entrevistado no instituto.
3	Preocupações éticas relacionadas à implementação do <i>big data</i> no INCA.	Neste tópico o objetivo foi compreender a percepção dos entrevistados sobre as questões éticas e de privacidade que podem surgir quando dados sensíveis de saúde são coletados, armazenados e analisados em larga escala.
4	O uso do <i>big data</i> para tomada de decisões estratégicas no INCA.	Este tema teve como objetivo identificar as fontes de dados disponíveis no instituto capazes de alimentar o <i>big data</i> . Outro ponto importante, está na observação de como os entrevistados enxergam a capacidade do INCA de extrair informações úteis do <i>big data</i> para apoiar o processo de tomada de decisões.

Fonte: Elaboração própria (2024).

Os dados coletados na pesquisa foram gravados e transcritos em texto, tendo seus conteúdos analisados indutivamente (THOMAS, 2006). A análise dos dados foi qualitativa, de acordo com os levantamentos realizados na etapa de coleta dos dados (VERGARA, 2015).

3.5 Limitações do Método de Pesquisa

Como é próprio do processo de seleção, ao se escolher determinados métodos ou técnicas de pesquisa, reconhece-se que os métodos de pesquisa selecionados, pela sua própria natureza, possuem limitações intrínsecas na sua aplicação, tanto na fase de coleta, quanto na fase de análise de dados. Entretanto, cuidados foram tomados visando superar as limitações do método, como descrito a seguir. A pesquisa documental, que envolveu relatórios de gestão, planejamento estratégico de saúde digital, plano diretor de TI, dentre outros documentos, teve como limitação a origem dos dados, uma vez que são de fontes secundárias, ou seja, não geradas durante a pesquisa. Tal limitação foram superadas a partir da averiguação da confiabilidade das fontes responsáveis pelo seu fornecimento.

A entrevista individual com roteiro semiestruturado apresentou limitações, como a influência das relações trabalhistas em suas respostas e a inteligibilidade das perguntas realizadas e das respostas dadas. Para atenuar este risco, foi realizado o piloto de pesquisa com um grupo de servidores do INCA, para corrigir pontos dúbios e validar o roteiro de entrevista, com a pactuação do Termo de Consentimento Livre e Esclarecido.

As análises indutivas e interpretativas enfrentaram a limitação de uma possível interpretação inadequada das percepções dos participantes pela pesquisadora. No entanto, essa questão foi mitigada por meio de uma categorização cuidadosa dos dados empíricos. Por outro lado, não foi possível prolongar a etapa de coleta, uma vez que os profissionais estavam sobrecarregados com as demandas de suas funções. Porém, isso não resultou a prejuízo para a pesquisa.

4. ESTUDO APLICADO

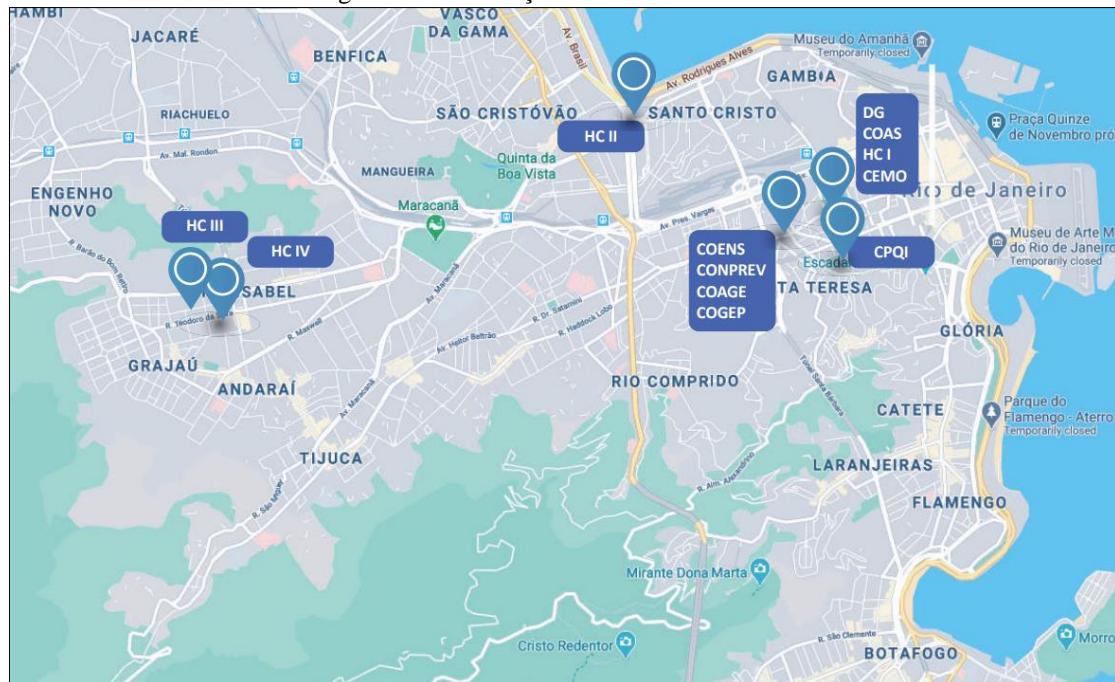
Neste tópico é feita a caracterização do Instituto Nacional de Câncer (INCA) e apresentado o produto tecnológico que será elaborado neste estudo.

4.1 Caracterização da unidade de estudo

A organização escolhida para a realização deste estudo foi o Instituto Nacional de Câncer (INCA), instituição pública de saúde sob a administração direta do Ministério da Saúde, e vinculado à Secretaria de Atenção à Saúde (SAS), com unidades distribuídas em diferentes bairros da cidade do Rio de Janeiro (INCA, 2021).

A representação cartográfica da Figura 9, ilustra a localização geográfica da Direção- Geral (DG), de suas coordenações (Administração-Geral, Assistência, Ensino, Gestão de Pessoas, Pesquisa e Inovação e Prevenção e Vigilância) e das unidades hospitalares (Hospital do Câncer I, Hospital do Câncer II, Hospital do Câncer III, Hospital do Câncer IV), além do Centro de Transplante de Medula Óssea (CEMO) que compõem o INCA (INCA, 2022).

Figura 9 - Distribuição das unidades do INCA.



Fonte: Relatório de Gestão do INCA (2024).

O INCA atua na área de prestação de serviços de saúde à sociedade, valendo-se de sua condição como órgão de pesquisa, onde desempenha um importante papel na disseminação do conhecimento e da prática oncológica no país. A pluralidade de atuação do INCA como prestador de assistência médico-hospitalar especializada e sua efetiva participação na formulação da Política Nacional de Prevenção, Diagnóstico e Tratamento do Câncer (INCA,2019) ampliam a sua importância no cenário da saúde do Brasil. Compete ao Instituto Nacional de Câncer (INCA, 2022):

- I. Participar da formulação da política nacional de prevenção, diagnóstico e tratamento do câncer;
- II. Planejar, organizar, executar, dirigir, controlar e supervisionar planos, programas, projetos e atividades, em âmbito nacional, relacionados com prevenção, diagnóstico e tratamento das neoplasias malignas e das afecções correlatas;
- III. Exercer atividades de formação, treinamento e aperfeiçoamento de recursos humanos, em todos os níveis, na área de cancerologia;
- IV. Coordenar, programar e realizar pesquisas clínicas, epidemiológicas e experimentais em cancerologia; e
- V. Prestar serviços médico-assistenciais aos portadores de neoplasias malignas e afecções correlatas.

O INCA definiu três objetivos estratégicos para o período de 2020-2023: 1) Alavancar a integração institucional para potencializar os resultados do INCA, com vistas ao fortalecimento do SUS e do protagonismo nacional; 2) Fortalecer a política de desenvolvimento de pessoas com foco na gestão por competência e; 3) Impulsionar a eficiência dos processos de trabalho, contribuindo para o alcance de uma gestão de excelência no SUS.

Os objetivos foram definidos com base em indicadores, que na perspectiva do INCA são capazes de impulsionar a criação de valor a curto, médio e longo prazo para os serviços de saúde prestados pelo Instituto à população (INCA, 2022). Na perspectiva do instituto, sua posição no cenário da saúde nacional será fortalecida por meio da integração sistêmica entre as áreas organizacionais e pelo desenvolvimento de soluções estratégicas para impulsionar suas ações nos segmentos de atuação do INCA (INCA, 2021).

Para alcançar o objetivo de alavancar a integração institucional para potencializar os resultados do INCA visando o fortalecimento do SUS, foram definidas as seguintes metas:

- a) Desenvolver a integração sistêmica entre as áreas: Prevenção e Vigilância, Assistência, Ensino e Pesquisa.
- b) Aprimorar a rede de forma multidisciplinar para assegurar a qualidade sistêmica de controle do câncer.
- c) Promover a integração institucional.
- d) Aprimorar ações integradas entre as áreas.
- e) Aumentar a produção do conhecimento em Oncologia (produção científica, políticas de saúde assertivas e formação de profissionais).
- f) Contribuir para a otimização da rede de controle de câncer.
- g) Fortalecer a imagem institucional, por meio de ações integradas do modelo técnico-científico.
- h) Ratificar o papel do INCA como centro especializado em incorporação tecnológica em Oncologia para o SUS.

Para alcançar o objetivo de fortalecer a política de desenvolvimento de pessoas, com foco na gestão por competência, foram definidas as seguintes metas:

- a) Aprimorar a alocação assertiva de servidores, baseada em competências e posições aderentes aos seus perfis.
- b) Elevar a eficiência operacional de processos críticos e de alto impacto para a operação do INCA.
- c) Promover capacitação profissional.
- d) Mapear as competências do INCA, para viabilizar a distribuição de pessoal (por competências) de maneira equilibrada entre as áreas.

Por fim, para alcançar o objetivo de impulsionar a eficiência dos processos de trabalho, contribuindo para o alcance de uma gestão de excelência no SUS, foram definidos os seguintes direcionadores:

- a) Assegurar a eficiência operacional, por meio de mapeamento, melhoria, capacitação e sustentação de processos críticos para a operação do INCA.
- b) Garantir a aplicação eficiente dos recursos institucionais.

A Figura 10 apresenta os três objetivos estratégicos definidos pelo INCA para o triênio 2020-2023.

Figura 10 - Objetivos estratégicos do INCA (2020-2023).



Fonte: Elaboração própria (2023), adaptado do Painel de Indicadores (INCA, 2021).

Este projeto de pesquisa, se alinha com o seguinte objetivo do Planejamento Estratégico Institucional do INCA proposto para o ciclo 2020-2023: “**Impulsionar a eficiência dos processos de trabalho, contribuindo para o alcance de uma gestão de excelência no SUS**”, uma vez que a implementação de inovação e incorporação de novas tecnologias é um dos temas considerados essenciais pela instituição, como ação capaz de impactar diretamente no alcance das metas estabelecidas no planejamento estratégico da instituição (INCA, 2021).

Um dos temas compreendidos como essenciais pela gestão do INCA para geração do valor dos serviços prestados pelo instituto, passa pela gestão de recursos orçamentários, inovação e incorporação de novas tecnologias e capacitação de profissionais e colaboradores, como ações necessárias para a prestação de serviços de saúde com qualidade para a população (INCA, 2021), neste cenário o *big data* de apresenta como um ativo de valor para que o instituto alcance seus objetivos.

Outro ponto relevante, está no fato de que para a elaboração do Planejamento Estratégico, o INCA considerou como instrumento de partida o Plano Nacional de Saúde (PNS). Os objetivos expressos no PNS incluem o fomento à produção do conhecimento científico, destinado a promover o acesso da população às tecnologias e informações em saúde de forma equitativa, igualitária, progressiva e sustentável (INCA,2020).

Neste sentido, estima-se que a popularização do acesso aos serviços digitais de saúde contribuirá para o aumento do volume dos dados de saúde gerados em tempo real, o que torna o uso de tecnologias, como *o big data*, um ativo de valor para o gerenciamento dos dados de saúde coletados por meio do uso das plataformas digitais de saúde (JANETT e YERACARIS, 2020), especialmente no caso do INCA.

No ano de 2020 o Setor de Tecnologia da Informação (SETI), em conjunto com os gestores das áreas de planejamento estratégico do INCA, apresentaram um projeto para a implementação da área de Ciência de Dados e Gestão do Conhecimento no instituto (INCA,2021).

As áreas envolvidas no projeto, justificaram a relevância do projeto ao afirmar que o INCA é produtor de grandes volumes de dados de saúde, estruturados e não estruturados, originados de diferentes fontes, e que a criação de uma área para análise descritiva e preditiva desta base de dados de saúde, poderá propiciar a geração de conhecimentos relevantes para a área da oncologia.

Estas informações, poderão ser utilizadas para apoiar à tomada de decisões estratégicas da instituição, tendo em vista a necessidade existente de mapear, armazenar e compartilhar as experiências e conhecimento produzidos pelo INCA com outras instituições de saúde do país (INCA, 2021). O escopo do projeto prevê:

1. A seleção de servidores com expertise em tecnologias como *o big data* e Inteligência Artificial.
2. Estruturação tecnológica da instituição, com a aquisição e implantação do *softwares* e *Hardwares* desenvolvidos especificamente para suportar o gerenciamento de um volumoso conjunto de dados de saúde.
3. Estruturação física da Instituição.

4. Capacitação dos profissionais de saúde em tecnologias de *big data* e inteligência artificial.

Paralelamente, tomando como base os estudos analisados ao longo desta pesquisa, observou-se que as instituições de saúde orientadas por dados aplicaram terapias mais assertivas, controlaram seus ativos com maior eficiência e beneficiaram os pacientes com tratamentos personalizados, fatores que melhoraram a qualidade do tratamento e aumentaram as chances de sobrevivência dos pacientes com câncer (SCHLICK et al, 2018; KARPATHAKIS et al, 2021).

Portanto, pressupõem-se que, por se tratar de uma instituição pública de saúde, o INCA não está alheio a esta realidade. Neste sentido, tecnologias como o *big data* podem se apresentar como uma ferramenta estratégica de valor para as instituições de saúde que buscam alcançar seus objetivos estratégicos a médio e longo prazo (CARDONA, 2022).

4.2 Produto Tecnológico

A partir da pesquisa conduzida, foi elaborado um plano estratégico com recomendações práticas para a implementação do *big data* no Instituto Nacional de Câncer (INCA). Além disso, foi desenvolvido um plano de ação complementar, na forma de um Produto Técnico-Tecnológico (PTT), que aborda os principais desafios que poderão surgir durante o processo de implementação do *big data* e as estratégias que poderão ser adotadas pela instituição para mitigá-los. Este processo compreendeu uma análise minuciosa dos dados coletados durante a pesquisa aplicada, juntamente com uma revisão das práticas e políticas vigentes no Instituto.

É relevante destacar a escassez de estudos que delineiam os caminhos para a efetiva implementação e uso do *big data* nas instituições de saúde pública brasileiras, conferindo a este relatório uma contribuição substancial para o avanço do conhecimento nesta área.

Em suma, espera-se que o produto técnico elaborado nesta dissertação não apenas forneça orientações práticas para o INCA, mas também represente uma importante contribuição para o avanço dos conhecimentos relacionados aos caminhos que podem ser adotados por uma instituição de saúde pública para a

implementação de tecnologias como o *big data*.

O próximo passo deste estudo será a elaboração de um plano estratégico com o objetivo de apresentar diagnósticos, resultados encontrados e propostas recomendadas (CAPES, 2019) para o contexto do INCA. A seção que apresentará o produto tecnológico resultante deste estudo será discutida no Capítulo 6 desta dissertação. Abaixo estão os subcapítulos delineados para constituí-lo:

a) Classificação e Caracterização do Produto

Neste subcapítulo (sessão 6.1), serão apresentadas as informações gerais desta dissertação, como título, identificação do docente orientador, identificação da discente, linha de pesquisa vinculada à proposta da produção, impacto da produção tecnológica e outros dados relevantes para o estudo, seguindo as diretrizes estabelecidas pela CAPES (2019).

b) Relatório Técnico Conclusivo

O relatório técnico conclusivo tem como propósito apresentar recomendações estratégicas para viabilizar a implementação do *big data* no Instituto Nacional de Câncer (INCA). Sua estrutura é dividida em quatro partes.

- Parte 1: irá apresentar brevemente a proposta do estudo, especificando seu contexto e finalidade.
- Parte 2: com base nos resultados obtidos na coleta de dados, esta etapa terá como objetivo apresentar as ações práticas que poderão ser adotadas pelo INCA para implementação do *big data* na instituição.
- Parte 3: considerações finais acerca do produto tecnológico elaborado.

Por fim, será elaborado um plano de ação complementar, na forma de um Produto Técnico-Tecnológico (PTT), para abordar os principais desafios que, na percepção dos sujeitos entrevistados, poderão surgir durante o processo de implementação do *big data* no Instituto e as estratégias que poderão ser adotadas pelo INCA para mitigá-los.

5. RESULTADOS

A análise dos resultados desta pesquisa qualitativa, foi conduzida por meio de entrevistas semiestruturadas, com três perfis de entrevistados: (1) gestores responsáveis pelo planejamento estratégico do INCA; (2) profissionais da área da saúde; e (3) profissionais de tecnologia da informação do Instituto Nacional de Câncer (INCA).

O objetivo principal foi explorar as percepções desses profissionais sobre as estratégias que podem ser adotadas pelo INCA para implementação do *big data* no instituto. Os resultados foram organizados em categorias e correlacionados com o referencial teórico discutido nesta dissertação.

5.1 Descrição dos Participantes

Esta sessão foi composta por um universo de 11 entrevistados, sendo 4 profissionais de saúde, incluindo médicos e profissionais da assistência, 3 gestores da área de planejamento estratégico do INCA e 4 profissionais de TI atuantes na instituição. Estes participantes apresentaram potencial para abranger um amplo *know-how* no que tange aos conhecimentos práticos e organizacionais, tendo em vista que em média possuem entre 7 e 26 anos de atuação no instituto, além do conhecimento teórico sobre tecnologias com potencial para melhorar a prática em saúde, como é o caso do *big data*. No que se refere aos perfis dos trabalhadores entrevistados nesta sessão, seguem abaixo as seguintes informações:

Quadro 7 - Dados referentes aos perfis dos entrevistados.

ID do Entrevistado	Expertise	Tempo de Serviço no INCA (anos)
Entrevistado 1	Especialista em TI com foco na implantação de soluções estratégicas para qualidade e desempenho dos serviços de TI do INCA.	26
Entrevistado 2	Gestor de TI especializado em governança e segurança de dados.	25
Entrevistado 3	Analista de Infraestrutura de TI com vasta experiência na manutenção e integração de sistemas tecnológicos de alta criticidade.	15
Entrevistado 4	Gestor de TI, enfermeiro de formação com especialização na área da saúde e experiência na assistência ao paciente com câncer.	18
Entrevistada 5	Médica com atuação na assistência ao paciente	7

ID do Entrevistado	Expertise	Tempo de Serviço no INCA (anos)
	oncológico.	
Entrevistado 6	Médico com expertise no tratamento dos pacientes oncológicos e atuação direta no diagnóstico e tratamento de tumores (câncer).	23
Entrevistado 7	Especialista em TI com participação no desenvolvimento de sistemas de saúde no INCA.	20
Entrevistado 8	Médico com experiência em assistência a pacientes com câncer e atuação direta no diagnóstico e tratamento da doença.	18
Entrevistado 9	Enfermeira com atuação direta na assistência aos pacientes oncológicos do setor de pediatria.	5
Entrevistado 10	Analista em Ciência e Tecnologia, Mestre na área de tecnologia da informação e com atuação direta nas decisões estratégicas do Serviço de Tecnologia da Informação (SETI).	22
Entrevistado 11	Analista em Ciência e Tecnologia com atuação em pesquisa sobre o câncer e bioinformática.	7

Fonte: Elaboração própria, com base na coleta de dados (2024).

5.2 Categorias de Análise

Durante a análise dos dados coletados, além das categorias principais predefinidas na sessão 3.4 desta dissertação, emergiram temas adicionais significativos para o tema em estudo, chamados de categorias emergentes. Essas categorias complementares foram identificadas com base nas informações fornecidas pelos entrevistados e representam áreas de interesse adicionais que surgiram durante o processo de análise.

As categorias principais listadas no Quadro 8, foram ajustadas de acordo com as conexões conceituais identificadas, que utilizou como referência principal o estudo “*Big data analytics in healthcare: a systematic literature review*” de Sayantan et al (2020), para definir, com base na literatura, temas relevantes para análise do objeto em estudo.

Quadro 8 - Categorias principais utilizadas na etapa de análise dos resultados.

Identificação	Descrição	Perguntas do Roteiro de Entrevista
Categoria Principal 1 (CP1)	Esta categoria buscou explorar a percepção dos entrevistados sobre os desafios e oportunidades inerentes à integração do <i>big data</i>	<ul style="list-style-type: none"> • De maneira geral, como o <i>big data</i> pode ser usado para melhorar a qualidade dos cuidados de saúde no INCA?

Identificação	Descrição	Perguntas do Roteiro de Entrevista
<p>Desafios e oportunidades para a implementação do <i>big data</i> no INCA.</p>	<p>no Instituto Nacional de Câncer (INCA). Incluindo questões relacionadas à infraestrutura tecnológica, capacitação de recursos humanos, financiamento para aquisição de tecnologias e a implementação de políticas robustas de segurança dos dados de saúde.</p>	<ul style="list-style-type: none"> • Na sua percepção, existe uma padronização para registrar eletronicamente os dados dos pacientes no INCA? • Poderia citar as fontes de dados dos pacientes disponíveis no INCA? • Considerando os desafios mencionados, como você sugere que o INCA possa superá-los, aumentando assim as chances de uma implementação bem-sucedida do <i>big data</i> no instituto? Por favor, você poderia compartilhar alguma experiência anterior ou exemplos de melhores práticas que poderiam ser aplicadas neste contexto? • Na sua percepção, o INCA possui os requisitos tecnológicos necessários para implementar com sucesso o uso do <i>big data</i> no instituto? Em caso de resposta negativa, poderia por favor citar os motivos que te fizeram chegar a essa conclusão?
<p>Categoria Principal 2 (CP2)</p> <p>Benefícios potenciais da adoção do <i>big data</i> no INCA.</p>	<p>Esta categoria buscou explorar a percepção dos entrevistados sobre os possíveis benefícios e impactos positivos que a adoção do <i>big data</i> poderá gerar para o INCA.</p>	<ul style="list-style-type: none"> • Na sua percepção, quais são os principais desafios e oportunidades para a implementação do <i>big data</i> no INCA? • De maneira geral, como você avalia os impactos que a adoção do <i>big data</i> no INCA teria para o desenvolvimento do seu trabalho?

Identificação	Descrição	Perguntas do Roteiro de Entrevista
		<ul style="list-style-type: none"> • Tendo em vista que o INCA é uma instituição que atua no desenvolvimento e coordenação de ações integradas para a prevenção e controle do câncer no Brasil, na sua percepção, quais seriam os impactos que a implementação do <i>big data</i> traria para o Instituto?
Categoria Principal 3 (CP3) Preocupações éticas relacionadas à implementação do <i>big data</i> no INCA.	<p>Esta categoria buscou explorar as questões éticas, legais e sociais associadas à utilização do <i>big data</i> no contexto do INCA. Isso inclui preocupações com a privacidade e segurança dos dados dos pacientes, o consentimento informado para o uso de informações pessoais, o potencial de discriminação e estigmatização, a equidade no acesso aos serviços de saúde, e a transparência dos responsáveis pela gestão dos dados.</p>	<ul style="list-style-type: none"> • De maneira geral, você considera que a implementação do <i>big data</i> no instituto traria desafios éticos e legais quanto a coleta de dados dos pacientes? Caso a resposta seja positiva, poderia apontar os principais desafios e como eles poderiam ser superados?
Categoria Principal 4 (CP4) O uso do <i>big data</i> para tomada de decisões estratégicas no INCA.	<p>Esta categoria procurou explorar a percepção dos entrevistados sobre o potencial do <i>big data</i> para aprimorar as decisões estratégicas do INCA.</p>	<ul style="list-style-type: none"> • De maneira geral, você considera que a adoção do <i>big data</i> no INCA poderia auxiliar na tomada de decisões estratégicas da instituição? Caso a resposta seja positiva, poderia por favor, explicar os motivos que te fizeram chegar a essa conclusão? • De maneira geral, você considera que a adoção do <i>big data</i> no INCA poderia

Identificação	Descrição	Perguntas do Roteiro de Entrevista
		auxiliar na tomada de decisões clínicas dos profissionais de saúde? Caso a resposta seja positiva, poderia por favor, explicar os motivos que te fizeram chegar a essa conclusão?

Fonte: Elaboração própria (2024).

O Quadro 8 foi elaborado com o objetivo de apresentar as categorias emergentes identificadas na etapa de análise dos dados, cada uma acompanhada por uma breve descrição que destaca sua relevância para este estudo.

Quadro 8 - Categorias emergentes que complementam a análise dos resultados.

Identificação	Descrição
Categoria Emergente 1 (CE1) Governança de dados.	Esta categoria abordou os desafios relacionados à gestão e organização dos dados no contexto do INCA, incluindo a implementação de políticas, procedimentos e controles para garantir a qualidade, segurança, privacidade e integridade dos dados de saúde no Instituto.
Categoria Emergente 2 (CE2) Manutenção da qualidade dos dados gerenciados pelo INCA.	Esta categoria emergiu como um tema complementar à Governança de Dados, devido à frequência com que os entrevistados destacaram os desafios relacionados à manutenção da qualidade, integridade, segurança e acessibilidade dos dados de saúde utilizados pela instituição, a preparação da infraestrutura necessária para viabilizar a implementação do <i>big data</i> e a necessidade de capacitação dos profissionais de saúde para lidar com tecnologias de <i>big data</i> no Instituto.
Categoria Emergente 3 (CE3)	Esta categoria teve por objetivo explorar a percepção dos entrevistados sobre a necessidade

Identificação	Descrição
Desafios para Integração e Interoperabilidade dos Sistemas Ativos no INCA.	de integração dos dados de saúde originados de diferentes sistemas e plataformas ativas no INCA. Neste sentido, os participantes destacaram os desafios envolvidos na integração e interoperabilidade entre os diferentes sistemas ativos no INCA e os desafios envolvidos no processo de tratamento e manipulação das informações geradas por estas aplicações.

Fonte: Elaboração própria (2024).

As subseções seguintes, foram elaboradas com o objetivo de apresentar as percepções dos entrevistados sobre o tema em estudo, utilizando como base de análise as categorias descritas acima.

5.3 CP 1: Desafios e Oportunidades para Implementação do *Big Data* no INCA

Esta categoria de análise, teve por objetivo explorar a percepção dos entrevistados sobre os desafios e as oportunidades relacionados à implementação do *big data* no contexto do INCA. Os principais desafios identificados pelos participantes incluem: governança dos dados de saúde; garantia da qualidade dos dados capturados e armazenados nas bases do INCA; integração dos sistemas ativos no Instituto; privacidade e segurança dos dados sensíveis e; obtenção de investimentos em treinamento e capacitação ou na contratação de profissionais multidisciplinares com *expertise* para lidar com *big data*.

Além disso, os dados revelaram o *big data* apresenta potencial para contribuir de forma significativa para melhoria dos serviços e processos de saúde do INCA, tais como: otimização do processo de identificação de ineficiências nos processos administrativos da instituição, visando melhorar a eficiência operacional e reduzir custos; melhoria na previsão das demandas e no gerenciamento de estoques, como medidas para garantir que os recursos necessários estejam disponíveis para uso da instituição; utilização dos dados para auxiliar na detecção precoce de câncer e no monitoramento contínuo da saúde dos pacientes; uso de dados nas decisões clínicas para prever a progressão da doença e sugerir

intervenções preventivas e; apoio a descobertas e inovações na pesquisa oncológica.

5.3.1.1 CE 1: Governança dos dados

A governança de dados foi apontada pelos entrevistados como um dos principais desafios para a implementação eficaz do *big data* no Instituto. Neste contexto, os participantes destacaram a importância de fortalecer as políticas, procedimentos e normas de gestão dos dados de saúde, para garantir a qualidade, segurança e privacidade dos dados gerados e gerenciados pelo INCA.

O entrevistado 2, gestor de TI com 25 anos de experiência no INCA e atuação no Setor de Tecnologia da Informação do INCA (SETI), cujo a função atual está diretamente relacionada com a implantação de soluções estratégicas de governança e segurança de dados no instituto, contextualizou que no setor de saúde estabelecer políticas de governança de dados para proteger a privacidade dos pacientes é uma premissa para assegurar a conformidade da instituição com as diretrizes estabelecidas pela LGPD.

“No INCA, estamos passando por um processo de amadurecimento das políticas de governança de dados, o foco desta estruturação não está apenas em mitigar os riscos relacionados a segurança dos dados, mas também na definição de métodos para a garantia da integridade dos dados dos pacientes que serão manipulados e compartilhados ao longo dos nossos processos e atividades” (Entrevistado 2).

O entrevistado 1, especialista em TI com 26 anos de atuação no instituto, cujo atividade está diretamente relacionada ao SETI e alinhada com a implantação de soluções estratégicas com o objetivo de fornecer qualidade e desempenho aos serviços disponibilizados pelo INCA, afirmou que do ponto de vista estratégico, este não é um processo simples e envolve a participação ativa e colaborativa entre diferentes esferas.

“Este é um processo orquestrado de médio a longo prazo, que requer um planejamento estratégico envolvendo todas as lideranças do Instituto. O sucesso no planejamento e implementação das políticas de governança de dados no INCA, será um reflexo da integração entre pessoas, processos e tecnologia” (Entrevistado 1).

O tema levantado pelos entrevistados sobre os desafios relacionados a integração dos processos coexistentes no setor de saúde, remeteu a Sayantan et al (2020), que no estudo “*Big data analytics in healthcare: a systematic literature review*”, sugeriu a adoção de um *framework* estruturado como uma ação eficaz para implementação de políticas e diretrizes de governança de dados no setor de saúde. Existe um consenso entre

os entrevistados de que, no contexto do INCA, o gerenciamento adequado e estruturado dos dados de saúde, é um fator importante para evitar inconsistências e ruídos capazes de comprometer a exatidão dos resultados obtidos por meios das análises realizadas pelo Instituto.

O entrevistado 8, profissional de saúde com 18 anos de experiência na assistência ao paciente com câncer, reforçou a relevância da governança dos dados para o INCA ao explicar que, a adoção de políticas para um melhor gerenciamento dos dados dos pacientes, é um fator importante para as expectativas futuras do Instituto, no que tange a integração entre os sistemas do INCA com tecnologias avançadas, como o *Machine Learning* e Inteligência Artificial, para apoiar a prática médica, a tomada de decisões estratégicas, aprimorar a análise dos dados de saúde gerados pela instituição e gerar conhecimentos relevantes para a pesquisa oncológica na área médica.

Com base nos dados foi possível observar que, as percepções dos entrevistados não apenas estão alinhadas com os conhecimentos teóricos apresentados no referencial teórico desta dissertação, mas também estão alinhadas com a visão estratégica do INCA a médio e longo prazo, tendo em vista que em seu Planejamento Estratégico para o ciclo de 2024 à 2027, assim como no planejamento do ciclo 2020-2023, o Instituto destacou a gestão da qualidade e segurança dos dados de saúde como fatores importantes para consolidar o INCA como uma instituição de referência nas ações integradas para o controle do câncer no Brasil.

Ainda com base na percepção dos entrevistados, foi possível identificar os pilares para implementação de um *framework* de governança de dados de saúde robusto e que esteja alinhado com os objetivos estratégicos da instituição. Entre eles:

1. Implementar padrões e melhores práticas de governança de dados para proteger os dados do instituto contra acessos não autorizados e ataques cibernéticos. Este processo, poderá envolver o uso de tecnologias de criptografia e mecanismos robustos de autenticação, como a autenticação em duas etapas (multifator) nos sistemas e aplicações que compõem a estrutura tecnológica da instituição.
2. Implementar e disseminar normas de privacidade para o gerenciamento e

manipulação dos dados sensíveis, com o objetivo de garantir que os dados dos pacientes sejam utilizados de maneira transparente e em conformidade com as leis de proteção de dados vigentes no Instituto.

3. Implementar métodos para assegurar a confiabilidade e precisão dos dados dos pacientes. Este processo poderá incluir a implementação de *softwares* ou algoritmos² projetados para realizar tarefas automatizadas de limpeza e verificação dos dados de saúde, desenvolvidos especificamente para identificar e corrigir erros, inconsistências e duplicatas das informações, economizando tempo e recursos institucionais.

Os pilares apresentados pelos entrevistados, alinham-se aos pontos fundamentais levantados pelos autores Abrahan et al. (2019) no estudo “*Data governance: A conceptual framework, structured review, and research agenda*”, no que se refere aos processos de planejamento e implementação de políticas de governança de dados nas instituições de saúde. Neste estudo, os autores também destacaram a importância de implementar programas abrangentes de capacitação para os profissionais de saúde, com ênfase na promoção de melhores práticas de governança dos dados.

Nesse contexto, a entrevistada 5, médica oncologista com 7 anos de atuação no instituto, destacou que seria interessante que os programas de treinamento para capacitação dos profissionais de saúde no INCA, além de abordarem questões específicas como boas práticas para a manutenção da privacidade do paciente, também incluíssem orientações para a manipulação segura dos dados em todas as etapas da jornada dos pacientes na instituição. Isso abrangeeria desde a análise e diagnóstico, passando pelo tratamento e assistência, até o período de pós-alta, considerando que, durante cinco anos, o paciente será acompanhado regularmente pelo INCA para monitorar a evolução da doença e a resposta ao tratamento.

“Não estamos falando de algo simples ou que faça parte do nosso cotidiano. Portanto, será necessário reforçar e disseminar este conhecimento para que

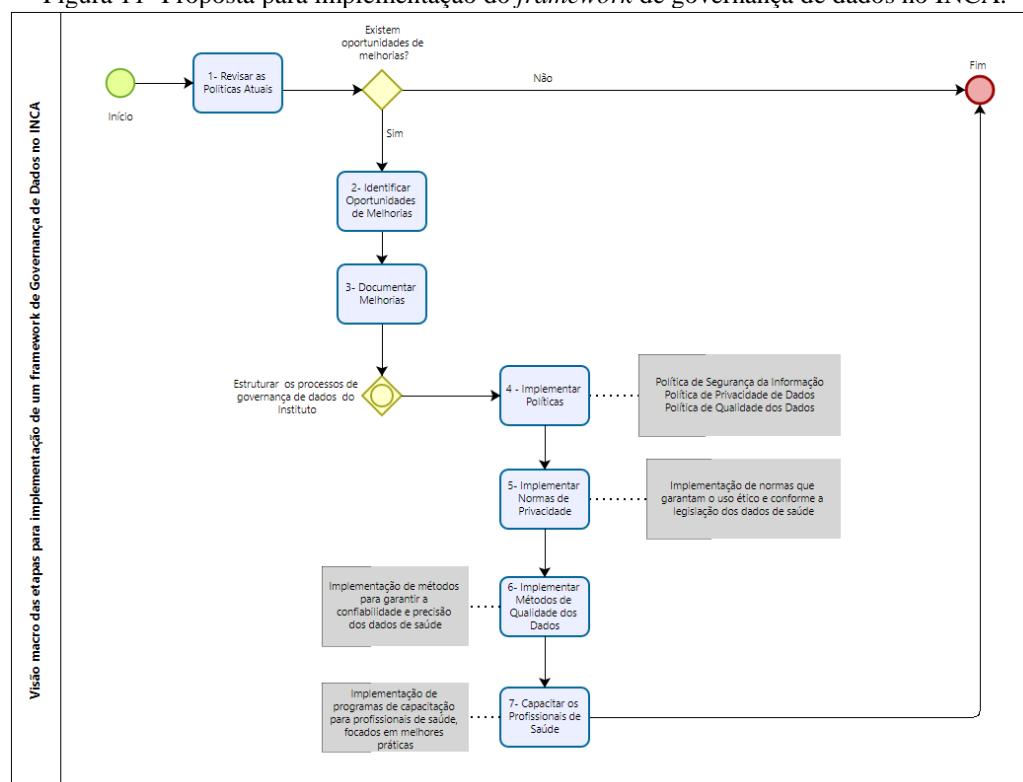
² Algoritmo de verificação de dados é um conjunto de instruções lógicas e computacionais projetadas para analisar, validar e garantir a qualidade e integridade dos dados. Esses algoritmos são desenvolvidos para realizar verificações automáticas em conjuntos de dados, identificando possíveis erros, inconsistências ou anomalias. Além disso, os algoritmos podem ser programados para verificar a precisão dos dados, garantir que estejam dentro de limites aceitáveis, detectar valores anômalos ou outliers e validar a conformidade com os padrões estabelecidos (CHRISTEN, 2018).

nós, profissionais atuantes do INCA, estejamos familiarizados com essas metodologias, e consequentemente, mais bem preparados para enfrentar os desafios e responsabilidades relacionados à gestão dos dados na instituição” (Entrevistado 9).

Complementando esta ideia, o entrevistado 4 ressaltou a importância da atualização regular dos treinamentos de capacitação dos profissionais do INCA, para gestão eficiente e segura dos dados, e como forma de garantir que o conteúdo sempre esteja alinhado com as mudanças nas tecnologias e nas políticas de segurança da informação vigentes na instituição.

O diagrama BPMN apresentado na Figura 11, foi elaborado com o objetivo de sistematizar os caminhos que podem ser adotados pelo INCA, para a implementação de um *framework* de melhores práticas de governança de dados no Instituto. Espera-se que este *framework* seja capaz de apoiar no processo atual do Instituto de implementação de uma base de dados robusta para apoiar decisões estratégicas e operacionais, fortalecendo as iniciativas coordenadas pelo INCA para prevenção, diagnóstico e tratamento do câncer.

Figura 11- Proposta para implementação do *framework* de governança de dados no INCA.



Fonte: Elaboração própria, com base nos dados coletados no INCA (2024).

Os dados sugerem que, no caso do INCA este é um processo dinâmico e que requer melhorias contínuas, onde integração entre processo, pessoas e tecnologia serão a base para a implementação bem-sucedida das diretrizes e políticas de governança de dados no Instituto.

Neste contexto, os processos serão os responsáveis por estabelecer as diretrizes e fluxos de trabalho necessários para gerenciar os dados de saúde gerados pelo INCA com eficiência. Por sua vez, as pessoas serão responsáveis por executar e monitorar esses processos e a tecnologia será o meio para que o instituto esteja preparado com ferramentas e infraestrutura adequadas para suportar e automatizar esses processos. Assim, as chances de que todas as partes do sistema estejam alinhadas e trabalhando de maneira coordenada aumentam, mitigando os riscos de falha de comunicação ou resistência às mudanças processuais por parte dos atores envolvidos nesta frente.

Por fim, os dados sugerem que, fomentar o conhecimento das políticas e diretrizes de melhores práticas para gerenciamento dos dados na instituição e promover a conscientização e o treinamento contínuo dos colaboradores, serão fatores importantes para aumentar a eficácia, a sustentabilidade e a adaptabilidade das práticas de governança de dados no Instituto, gerando um ambiente propício para adoção de tecnologias de ponta.

5.3.1.2 CE 2: Manutenção da qualidade dos dados gerenciados pelo INCA

A manutenção da qualidade dos dados de saúde capturados, armazenados, manipulados e gerenciados pelo INCA, emergiu como um tema complementar à categoria de Governança de Dados. Os entrevistados destacaram a importância de ter acesso a dados de qualidade para garantir a precisão e a confiabilidade das informações que serão utilizadas pela instituição para apoiar a tomada de decisões e apoiar nas ações de combate e prevenção do câncer coordenadas pelo Instituto.

Para o entrevistado 1, ao abordar a qualidade dos dados, estamos tratando da confiabilidade, consistência e valor das informações que serão utilizadas pelo INCA na implementação de políticas públicas de saúde no Brasil, com ênfase na promoção da saúde e na prevenção da doença. “Esses dados são fundamentais para a estruturação das Redes de Atenção à Saúde Pública e para a melhoria contínua dos serviços de saúde oferecidos à população”. O entrevistado 2 complementou:

“Estamos lidando com vidas humanas e garantir a confiabilidade dos dados é fundamental para que a adoção de inovações em tecnologia da informação faça sentido para a instituição. Do que adianta ter acesso a informações que não são confiáveis? É como construir uma casa sem uma estrutura sólida, ou seja, se essa base não for firme, toda a estrutura pode desmoronar. Neste sentido, a governança dos dados de saúde é a base sólida sobre a qual construímos nossas iniciativas de inovação em TIC e assegurar a qualidade e a confiabilidade dos dados é primordial para que essa estrutura se mantenha firme” (Entrevistado 2).

Os profissionais de TI entrevistados destacaram que existem projetos em andamento no INCA cujo sucesso está intrinsecamente relacionado à garantia da qualidade dos dados de saúde gerenciados pela instituição. Um exemplo citado foi a iniciativa liderada pelo Serviço de Tecnologia da Informação (SETI) para a implementação de um Portal de Dados Abertos. O objetivo deste portal é democratizar o acesso aos dados de saúde gerados pelo INCA e por meio do compartilhamento destes dados fomentar a pesquisa na área da Oncologia, apoiar a tomada de decisões de saúde das instituições com base em evidências e promover a inovação no setor de saúde.

Sobre o Portal de Dados Abertos do INCA, os profissionais entrevistados, que estão diretamente envolvidos nesta iniciativa, explicaram que o projeto está na fase de planejamento e que de forma macro, pode ser dividido em três etapas. A primeira prevê a implementação de um piloto, onde os dados gerados e capturados dos sistemas de informação em saúde do Instituto e que serão armazenados em um grande repositório de dados da instituição, serão tratados e disponibilizados no Portal, permitindo que a Coordenação de Ensino (COENS) e a Pesquisa Clínica do INCA, tenham acesso a essas informações de maneira estruturada e centralizada. A segunda etapa prevê a ampliação dos setores da instituição com permissão de acesso aos dados de saúde publicados na plataforma. Na terceira fase, espera-se que o projeto esteja estruturado de forma a permitir que os dados sejam disponibilizados, não apenas no ambiente do INCA, mas também para as instituições de saúde a nível nacional.

O entrevistado 7, especialista em TI, com participação ativa no processo de desenvolvimento dos sistemas do INCA, acredita que um dos principais desafios que o INCA enfrentará para manter a qualidade dos dados que alimentarão essas bases, está relacionado à necessidade de higienizar os dados gerados pelos sistemas legados da instituição, e explicou que “por vezes esses sistemas não apresentam críticas internas

suficientes para validação e consistência dos dados, o que pode resultar em registros imprecisos ou incorretos”.

Consequentemente, isso aumenta a probabilidade de que existam bases de dados antigas poluídas e que necessitam passar por um rigoroso processo de tratamento e validação das informações.

O entrevistado 4, profissional de saúde, que atuou na assistência do INCA e que atualmente ocupa o cargo de Gestor de TI no setor de tecnologia do Instituto, reforçou essa visão, destacando que “do ponto de vista estratégico, não existe política pública sem dados de qualidade para embasar essas decisões” e complementou:

“No INCA, temos acesso a uma ampla variedade de dados de diferentes sistemas internos, a exemplo do Absolute, PACS, Matrix, Anatomia Patológica, Radiologia, Banco Nacional de Tumores (BNT), Registro Hospitalar de Câncer (RHC), entre outros. Esses dados são extremamente valiosos, pois incluem informações detalhadas sobre os pacientes, características dos tumores, resultados clínicos e de imagem dessas lesões. Ao integrar e analisar esses dados de forma estruturada, dentro de um padrão lógico, as chances de gerar conhecimento relevantes para o campo da oncologia aumentam significativamente”
 (Entrevistado 4).

Seguindo nesta mesma linha, o entrevistado 1, utilizou o princípio de '*Garbage In, Garbage Out*', para explicar como a falta de padronização no registro das informações dos pacientes nos sistemas institucionais, podem influenciar diretamente na qualidade das informações que serão utilizadas pelo INCA para nortear as suas decisões estratégicas, e complementou:

“Para reduzir os riscos de que as boas práticas no registro de informações dos pacientes não sejam seguidas no Instituto, a gestão precisa ir além de métricas e diretrizes. É essencial que exista uma mudança de mentalidade da equipe, para que entendam a garantia da qualidade dos dados como um investimento estratégico para o INCA. No entanto, para que os profissionais da linha de frente comprem essa ideia, será necessário que a gestão esteja próxima do operacional, não engessando os processos, mas criando um ambiente propício para que os registros sejam realizados com precisão e apoiando estes profissionais ao longo deste processo de aprendizagem e adaptação”
 (Entrevistado 1).

É consenso entre os sujeitos participantes que, manter a qualidade dos dados no setor de saúde é algo complexo e requer um esforço coordenado que envolve pessoas, processos, tecnologias e capacitação, para que a instituição esteja mais bem preparada para aplicar estas boas práticas em sua rotina de trabalho e apoiar a mudança de *mindset*

dos profissionais quanto à valorização do uso dos dados na rotina da instituição.

O Quadro 7 busca apresentar de forma sistematizada, os principais desafios apontados pelos entrevistados em relação à manutenção da qualidade dos dados de saúde gerenciados pelo INCA, que podem vir a impactar no processo de implementação do *big data* no Instituto.

Quadro 9 - Percepção dos entrevistados sobre os principais desafios que podem ser enfrentados pelo INCA, para garantir a qualidade dos dados de saúde gerenciados pela instituição.

Desafio	Descrição
Higienização dos Dados de Sistemas Legados.	Os sistemas legados do INCA, muitas vezes não apresentam críticas suficientes sobre os registros realizados pelos usuários em suas interfaces, o que pode resultar no armazenamento de informações imprecisas ou incorretas. Neste sentido, existe uma preocupação latente dos profissionais de TI entrevistados, em relação a necessidade de tratamento dessas bases de dados antigas para remover dados poluídos. Essa ação não é vista como algo simples, uma vez que envolve processos complexos de limpeza, validação e integração de dados, além de requerer recursos específicos e <i>expertise</i> técnica para garantir a qualidade e integridade das informações.
Confiabilidade das Informações.	Garantir a precisão e a consistência dos dados de saúde que serão utilizados nos processos da instituição, é percebido pelos entrevistados como um fator importante para aumentar as chances de sucesso das iniciativas capitaneadas pelo Instituto para promoção do controle do câncer e ações nacionais integradas em prevenção, assistência, ensino e pesquisa na área da oncologia. Além disso, gerar dados confiáveis prepara o Instituto para adoção de tecnologias robustas como o <i>big data</i> , a inteligência artificial e ferramentas integradas que viabilizam a análise preditiva dos dados de saúde gerados pelo INCA,

Desafio	Descrição
	<p>para que assim, o Instituto possa otimizar a sua capacidade de gerar novos conhecimentos que sejam pertinentes à realidade do Instituto, o que, por sua vez, impacta positivamente na qualidade dos serviços prestados pela instituição aos pacientes com câncer.</p>
Integração e Análise dos Dados de Saúde no ambiente tecnológico do INCA.	<p>Os entrevistados apresentaram preocupação em relação à falta de padronização entre os dados originados pelos sistemas ativos na instituição. Neste sentido, eles explicaram que, parte desses sistemas são geridos por empresas externas e utilizam linguagens e estruturas de dados distintas dos sistemas nativos do INCA, além disso, existem cenários em que os sistemas de terceiros apresentam interfaces licenciadas, o que dificulta a integração eficaz entre os diferentes sistemas. E ressaltaram que, cada sistema foi desenvolvido para atender às necessidades específicas de seu período de implementação, mas, devido aos avanços tecnológicos, muitos deles não se alinham mais às exigências atuais do instituto. Isso resultou em dados estruturados de maneiras variadas, dificultando a garantia da qualidade dos registros das informações de maneira unificada.</p> <p>Existe um consenso sobre a falta de padronização nestes registros, o que pode comprometer a consistência e a integridade dos dados legados, tornando mais complexa a migração, integração e análise sistemática desses dados. Outro ponto destacado pelos sujeitos entrevistados, está relacionado ao fato de que parte dos sistemas antigos eram instalados localmente e não contavam com servidores dedicados para o armazenamento das informações. Isso significava que as bases de dados poderiam ficar armazenadas localmente nas máquinas, gerando um risco significativo de perda de dados em caso de falha do</p>

Desafio	Descrição
	<p><i>hardware</i>, como a quebra da máquina ou problemas no disco rígido.</p> <p>Esses fatores, podem gerar lapsos temporais na análise dos registros, o que pode impactar negativamente na capacidade do INCA de extrair informações relevantes do <i>big data</i>. Assim, a confiabilidade dos dados é percebida como um desafio significativo, que precisa ser abordado com atenção para garantir decisões estratégicas baseadas em informações precisas e fidedignas.</p>
Apoio à Pesquisa e Inovação	<p>A democratização do acesso aos dados de saúde gerados pelo INCA, por meio de iniciativas como o Portal de Dados Abertos, que tem a intenção de fomentar a pesquisa na área de oncologia e promover a inovação no setor de saúde, apoiando a reformulação de políticas públicas de combate e prevenção do câncer no Brasil está diretamente relacionada a confiabilidade das informações de saúde geradas pela instituição. Para que estas iniciativas sejam exitosas, é necessário que o Instituto tenha uma atenção especial em relação a necessidade de promoção de melhores práticas para garantia da qualidade e confiabilidade dos dados de saúde que serão utilizados pela instituição nestas frentes.</p>
Investimento em Infraestrutura para o Tratamento Adequado dos Dados de Saúde Gerenciados pelo INCA.	<p>Na percepção dos entrevistados, existe a necessidade de investimentos em infraestrutura por parte do INCA, para criar um ambiente tecnológico que permita o tratamento adequado do <i>big data</i>, uma vez que os recursos tecnológicos existentes, são reservados a manutenção e suporte do ambiente de TI atual.</p> <p>Os entrevistados também descreveram como um desafio significativo para implementação bem-sucedida do <i>big data</i> no INCA, a necessidade de investimentos em capacitação</p>

Desafio	Descrição
	dos profissionais para lidar de maneira com as tecnologias complementares a análise do <i>big data</i> , mitigando o risco de subutilização dessas tecnologias.

Fonte: Elaboração própria, com base na análise dos dados coletados no INCA (2024).

Em relação a aquisição de tecnologias para viabilizar o tratamento adequado dos dados de saúde gerados pelo INCA, o entrevistado 10, analista em Ciência e Tecnologia no Instituto Nacional de Câncer (INCA/MS), atualmente na função de Chefe do Serviço de Tecnologia da Informação (SETI), explicou que o investimento em infraestrutura é um ponto a ser observado, concorda com a percepção dos entrevistados de que a infraestrutura ainda é inadequada para armazenamento e processamento de grandes volumes de dados, porém o INCA possui a previsão de aquisição de equipamentos até o final do ano de 2024 que permitirão a criação da infraestrutura necessária para implementação de tecnologias para este fim.

Complementando esta ideia, o entrevistado 7, explicou que o INCA está avançando na modernização de sua infraestrutura tecnológica e citou como exemplo o projeto para implementação de uma solução de infraestrutura hiperconvergente no INCA. Este projeto estratégico visa integrar de forma eficiente os recursos de computação, armazenamento e rede da instituição em uma plataforma unificada, simplificando a gestão dos dados e sistemas críticos de saúde.

A adoção da hiperconvergência não só poderá otimizar o desempenho e a eficiência operacional, mas também facilitar a escalabilidade da infraestrutura à medida que a demanda por serviços de saúde cresce no Instituto. Com um gerenciamento centralizado e uma abordagem simplificada para manutenção dos dados, o INCA poderá reduzir custos e melhorar a capacidade de resposta aos desafios emergentes no setor.

“O projeto de hiperconvergência não está sendo proposto com essa finalidade (implementação do *big data* no INCA). No entanto, ele pode estabelecer uma base que facilite a adoção dessa tecnologia no futuro, contribuindo para o aumento da segurança, confiabilidade e integridade dos dados manipulados pelo Instituto” (Entrevistado 7).

Para superar os desafios relacionados à manutenção da qualidade dos dados de saúde, os entrevistados sugeriram a implementação de processos automatizados que

abrangem desde a extração automática de grandes volumes de dados até a integração dos sistemas de informação em saúde, passando pela validação e limpeza automática dos dados capturados, resultando na geração de relatórios detalhados em ferramentas de *Business Intelligence* (BI).

Além de promoverem a precisão e a consistência dos dados, tais medidas visam reduzir a carga de trabalho manual dos profissionais de saúde, permitindo que eles se concentrem mais em atividades que demandam julgamento clínico e *expertise*, ao invés de tarefas operacionais repetitivas, resultando em uma redução potencial da incidência de erros na manipulação das informações de saúde registradas nos sistemas da instituição.

No entanto, o entrevistado 1, alertou que, embora seja importante contar com o suporte das automações e *softwares* inteligentes para o tratamento adequado dos dados, não se pode ignorar o fato de que sua efetividade depende da integração com os sistemas usados diariamente pelos profissionais da instituição, ou seja, da integração dos dados entre os diferentes sistemas ativos no ambiente do INCA.

Ainda de acordo com o entrevistado, assegurar que essas informações sejam acessíveis e de fácil visualização pelos profissionais de saúde não é uma proposta simples e depende de iniciativas coordenadas a nível organizacional, visando aumentar a probabilidade de uso efetivo dos dados.

“Um dos principais objetivos da ciência de dados no INCA é apoiar as decisões estratégicas do Instituto de forma assertiva, mas do que adianta elaborar um BI, com informações estruturadas e de fácil acesso, se os nossos recursos humanos não fizerem uso dessa ferramenta de forma estratégica? Se as partes interessadas não tiverem o interesse de se debruçar sobre esses dados para extrair conhecimentos relevantes para apoiar as suas decisões, o uso dessas aplicações perde o sentido” (Entrevistado 1).

Com base nos dados, identificou-se a necessidade de apresentar uma metodologia estruturada, com estratégias que podem ser adotadas pelo INCA para superar os desafios relacionados à manutenção da qualidade dos dados de saúde capturados, armazenados, manipulados e gerenciados pela instituição. Neste sentido, propõe-se a aplicação do ciclo PDCA (Plan-Do-Check-Act) para apoiar neste processo.

A fase de Planejamento (*Plan*) prevê a definição dos objetivos e metas de qualidade dos dados específicos para o INCA. Envolve a identificação dos requisitos de qualidade dos dados, mapeamento e documentação dos processos atuais de coleta,

armazenamento e uso dos dados institucionais. Além disso, desenvolvem-se políticas e procedimentos para garantir a qualidade dos dados, bem como métricas e indicadores de desempenho para monitorar a qualidade dos dados de saúde gerados pelo Instituto.

A fase de Execução (*Do*) prevê a implementação das políticas, procedimentos e estratégias desenvolvidos na fase de planejamento. Será necessário treinar os funcionários sobre as novas políticas e procedimentos de governança dos dados vigentes no Instituto, e implementar sistemas de coleta e armazenamento de dados que atendam aos requisitos de qualidade do INCA. Será fundamental que a coleta e registro dos dados sigam os procedimentos definidos, garantindo a precisão e integridade dos dados processados.

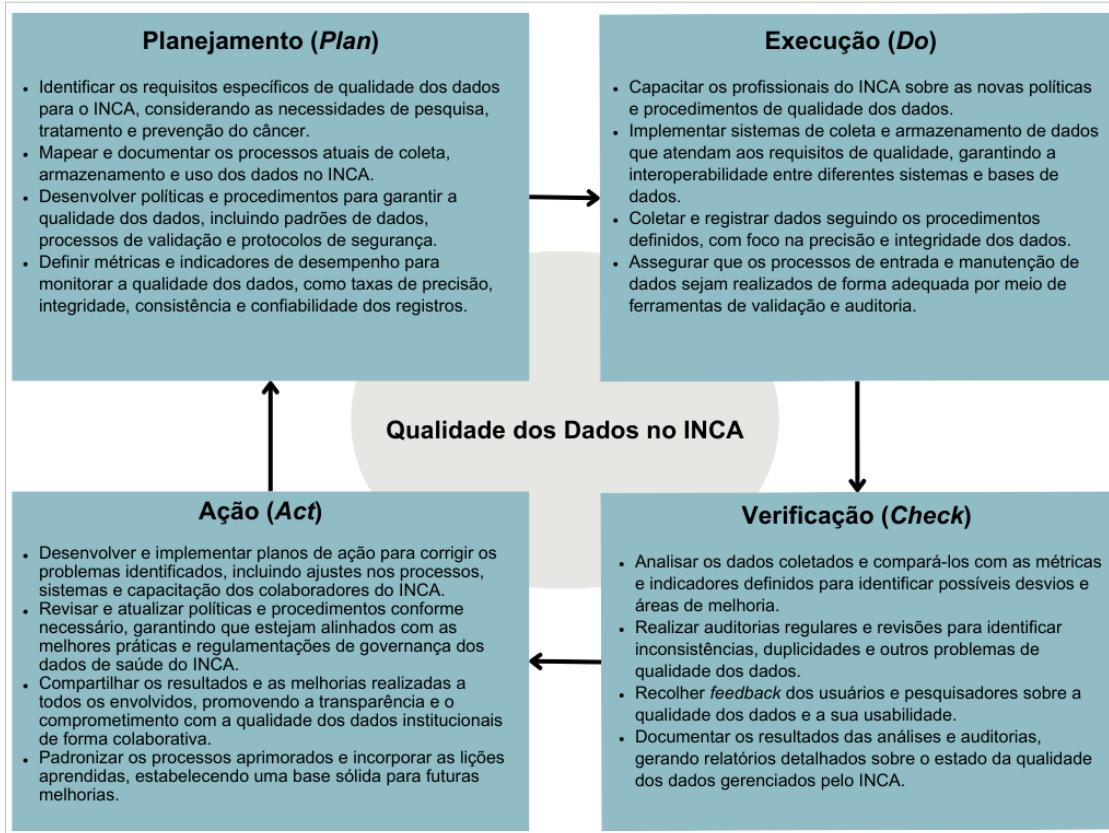
A fase de Verificação (*Check*) prevê que o INCA monitore e avalie o desempenho dos processos de qualidade dos dados capturados na etapa de execução. Isso inclui analisar os dados coletados e compará-los com as métricas e indicadores definidos, realizar auditorias e revisões regulares para identificar inconsistências, e recolher *feedback* dos usuários sobre a qualidade dos dados utilizados na instituição. Por fim, os resultados das análises e auditorias serão documentados para orientar ações futuras de governança dos dados.

Por fim, com base nos resultados da fase de verificação, implementam-se ações corretivas e de melhoria contínua para governança dos dados na instituição. Para isso, o INCA poderá elaborar planos de ação para corrigir os problemas identificados, revisar e atualizar as políticas e procedimentos de gestão dos dados conforme necessário, e aplicar novos treinamentos para os colaboradores com foco na correção das práticas inadequadas identificadas ao longo deste processo.

É uma boa prática que, as melhorias realizadas e os resultados alcançados sejam compartilhados com as partes envolvidas neste processo, visando a promoção da transparência e do compromisso com a qualidade dos dados institucionais de maneira colaborativa. Após refinar e padronizar os processos, será importante documentar as lições aprendidas para que estas sejam incorporadas as boas práticas de governança dos dados da instituição, estabelecendo uma base sólida para melhorias futuras.

A Figura 12 tem por objetivo apresentar as etapas do ciclo PDCA e as ações sugeridas para cada fase, adaptadas ao contexto e necessidades do INCA.

Figura 12 - Ciclo PDCA para manutenção da qualidade dos dados de saúde no INCA.



Fonte: Elaboração própria, com base na análise dos dados coletados no INCA (2024).

No caso do INCA, a implementação de estratégias de planejamento, execução e melhoria contínua para o gerenciamento do *big data*, se apresentam como um fator importante para preparar o ambiente do INCA estrategicamente para a adoção de tecnologias de ponta, aumentando assim, as chances de sucesso na adoção do *big data* nas práticas de saúde da instituição.

5.3.1.3 CE 3: Integração e Interoperabilidade dos Sistemas Ativos no INCA.

Os dados revelaram que, ao viabilizar a análise de grandes volumes de dados provenientes de diferentes fontes sistêmicas, o *big data* apresenta potencial para contribuir de forma significativa para melhoria das análises clínicas, otimização do

processo de identificação de padrões, tendências e biomarcadores tumorais³, além de aprimorar a precisão e personalização dos diagnósticos e tratamentos do câncer prestados pelo INCA. No entanto, para colher os benefícios do uso dessa tecnologia, será necessário que o Instituto enfrente dois desafios complexos: a integração e a interoperabilidade dos sistemas de saúde ativos na instituição.

O INCA possui uma variedade de sistemas que gerenciam os processos assistenciais e administrativos do Instituto. Parte desses sistemas foi desenvolvido para atender a necessidades específicas dos setores da instituição, resultando em uma arquitetura de TI heterogênea, onde nem sempre estas aplicações conversam entre si ou estão integrados com outros sistemas de saúde nacionais.

Com base nos dados foi possível identificar que, do ponto de vista da gestão, a interoperabilidade entre os diferentes sistemas de saúde não é apenas desejável, mas essencial para garantir o direito de acesso a saúde para a população. Neste sentido, o entrevistado 4 descreveu que muitos pacientes do INCA são assistidos por outros hospitais, pois fazem parte da rede do SUS e que a troca de informações sobre o estado de saúde dos pacientes entre diferentes aplicações de saúde é fundamental para viabilizar a continuidade do tratamento. No entanto, essas informações são compartilhadas entre diferentes sistemas de saúde, que nem sempre são integrados entre si, o que pode dificultar a integração com os sistemas do INCA devido à falta de compatibilidade entre as tecnologias utilizadas em cada ambiente.

Do ponto de vista tecnológico, em um ambiente complexo como o de saúde, esta incompatibilidade sistêmica e a falta de padronização no compartilhamento dos dados, pode ser causada por diferentes fatores, incluindo diferenças nas tecnologias utilizadas, como linguagens de programação e estruturas de banco de dados distintas, até questões relacionadas a governança dos dados, uma vez que essas diretrizes podem se modificar a depender do cenário em que os sistemas de saúde estão inseridos.

Diante disso, não necessariamente as instituições de saúde de uma mesma região

³ Biomarcadores tumorais são proteínas encontradas em tumores, no sangue ou em outros líquidos biológicos, que funcionam como indicadores da presença de câncer. Além de possibilitar o diagnóstico precoce, são capazes de identificar o órgão afetado e a extensão da doença no paciente, aprimorando as pesquisas e análises clínicas no campo da oncologia (ALMEIDA et al, 2024).

terão o mesmo padrão tecnológico e processual. É este contexto heterogêneo, com diferentes modelos de dados, padrões de formatação e decodificação dos registros de saúde, que fazem com que a interoperabilidade dos sistemas seja algo complexo não apenas para INCA, mas para o sistema de saúde como um todo.

“A falta de informações complementares na rede de saúde pode criar lacunas na qualidade do tratamento de pacientes com câncer, o que pode afetar de forma crítica os resultados de saúde desse indivíduo” (Entrevistado 3).

Outro ponto destacado pelo entrevistado 3, foi a ausência de interfaces padronizadas, desenvolvidas para viabilizar a comunicação entre os sistemas legados do INCA e as plataformas externas de saúde, a exemplo do Ministério da Saúde, sendo este um fator que pode intensificar os desafios para interoperabilidade dos sistemas de saúde ativos no Instituto. No entanto, os profissionais de TI entrevistados, ressaltaram que está em processo de desenvolvimento uma API (*Application Programming Interface*, ou em português Interface de Programação de Aplicação) de comunicação entre os ambientes do INCA e do MS, com o objetivo de otimizar o processo de compartilhamento dos dados de saúde de forma integrada entre estes dois contextos.

“A expectativa é de que, ao entrar em produção, o processo de transferência dos dados de saúde gerados pelo Instituto para as bases do MS seja aprimorado, aumentando a eficiência e segurança deste processo, para que as informações sensíveis dos pacientes sejam compartilhadas de forma assíncrona, padronizada e dentro das legislações de governança dos dados vigentes no setor de saúde” (Entrevistado 7).

Além disso, a incompatibilidade de versões de *software* e *hardware* entre os sistemas legados e os sistemas que possuem atualizações periódicas ativos no INCA, as restrições de acesso a alguns módulos de sistemas licenciados que foram desenvolvidos por empresas externas e hoje estão descontinuados, as vulnerabilidades que podem ser geradas pelo uso de tecnologias obsoletas e cujo período de sustentação dessas aplicações pelo fornecedor está expirada, são fatores apontados pelos entrevistados como pontos que merecem atenção do INCA no que tange a necessidade crescente de interoperabilidade no setor de saúde.

Corroborando com esta ideia, o entrevistado 1 explicou que, criar um ambiente de TI integrado é um fator importante para a expectativa futura do Instituto de implementar tecnologias de *big data* na instituição, e explicou que a criação de interfaces capazes de

viabilizar a comunicação e o compartilhamento de dados entre as diferentes tecnologias que compõem a infraestrutura de TI do INCA e do setor de saúde como um todo, é um processo capaz de aumentar ou diminuir as chances de que informações de valor sejam extraídas do *big data*.

“Sem uma integração adequada entre os sistemas de saúde, os dados permaneceriam isolados em ilhas de conhecimento, dificultando a análise abrangente necessária para iniciativas de *big data* no INCA” (Entrevistado 1).

Sob a ótica de gestão dos dados de saúde, a integração sistêmica não apenas facilitaria a organização dos dados gerenciados pelo INCA, mas também contribuiria para a diminuição de redundâncias e inconsistências dos registros de saúde, promovendo uma base de dados mais robusta e confiável, sendo este um fator primordial para extração de informações de valor do *big data* na saúde.

O entrevistado 11, Analista em Ciência e Tecnologia, há 7 anos atuando no Instituto, com passagens pela gestão nos setores de recursos tecnológicos e desenvolvimento de sistemas, atualmente alocado no setor de pesquisa sobre câncer/bioinformática, destacou que, uma vez que os sistemas do INCA estejam integrados, o próximo passo é viabilizar a interoperabilidade entre os sistemas integrados de saúde e que, na sua percepção, a interoperabilidade transcende a conexão entre os sistemas institucionais; ela permite que os dados sejam trocados e compreendidos de maneira eficiente e precisa entre diferentes plataformas de saúde. Em suas palavras, “no contexto do INCA, isso significa que as informações clínicas e administrativas da instituição poderão ser interpretadas, compartilhadas e utilizadas por diferentes sistemas conectados, independentemente de suas diferenças tecnológicas”.

Alinhando-se a essa perspectiva, o entrevistado 4 ressaltou que a interoperabilidade é fundamental para o avanço dos estudos oncológicos, pois facilita a consolidação de informações clínicas e laboratoriais provenientes de sistemas especializados diversos. Essa integração amplia significativamente as possibilidades de análise e correlação de dados, que são essenciais para a realização de pesquisas genéticas e epidemiológicas conduzidas pelo INCA. Contudo, o entrevistado 11 alertou que, por vezes, os dados clínicos estão espalhados entre diversos sistemas da instituição, nem sempre interligados entre si, e que essa falta de padronização sistêmica faz com que “parte

dos pesquisadores do Instituto relatam que, por diversas vezes, precisam consultar as fichas físicas em papel para buscar dados relevantes para a pesquisa em andamento”.

O relato do entrevistado 11, corrobora com os estudos que relatam que a crescente digitalização dos serviços de saúde levou à proliferação de sistemas de informação independentes, cada um com suas próprias especificidades e estrutura de dados no setor de saúde. Além disso, cabe destacar que, as instituições de saúde são compreendidas como um setor complexo para implementação de tecnologias de *big data*, uma vez que é composto por diferentes sistemas, de diferentes versões, desenvolvidos por diferentes empresas e que por vezes não se comunicam, mas coexistem num mesmo ambiente. Consequentemente, esse conjunto de fatores pode limitar a capacidade do INCA de consolidar informações originadas por diferentes fontes e dificultar a execução de análises abrangentes e precisas dos dados oncológicos.

Na percepção dos entrevistados, para superar os obstáculos relacionados à integração e interoperabilidade dos sistemas de saúde no INCA, será necessário que a instituição adote abordagens estratégicas que incluam o mapeamento dos sistemas existentes na instituição, o desenvolvimento de ferramentas tecnológicas capazes de apoiar no processo de integração e promover a troca padronizada e estruturada dos dados de saúde entre os sistemas integrados do Instituto. O Quadro 10 foi elaborado com o objetivo de apresentar essas percepções de maneira sintetizada.

Quadro 10 - Caminhos que podem ser adotados pelo INCA para viabilizar a integração e interoperabilidade dos sistemas de saúde ativos na instituição.

ID	Etapa	Descrição
1	Mapeamento e análise dos sistemas existentes no ambiente do INCA.	Identificação detalhada de todos os sistemas ativos no INCA, suas funcionalidades específicas e os tipos de dados que gerenciam. Esta análise, tem por objetivo proporcionar uma compreensão abrangente do ecossistema de TI do INCA, permitindo a identificação de pontos críticos de integração e possíveis gargalos no processo de interoperabilidade dos sistemas de saúde.

ID	Etapa	Descrição
2	Desenvolvimento de interfaces de integração entre os sistemas de saúde.	Criação de soluções de TI que permitam a comunicação entre os sistemas ativos no INCA, garantindo a troca de dados de forma segura, eficiente e em conformidade com os padrões de governança dos dados estabelecidos pelo Instituto. Sugere-se como uma boa prática, que estas interfaces sejam escaláveis para suportar futuras atualizações dos sistemas.
3	Adoção de padrões de interoperabilidade.	Implementação de normas amplamente aceitas para a troca eficientes de informações em saúde, como por exemplo a HL7 (<i>Health Level Seven</i>) e FHIR (<i>Fast Healthcare Interoperability Resources</i>). A adoção desses padrões, aumentam a probabilidade de que os dados sejam interpretados de maneira consistente e precisa entre sistemas com diferentes tecnologias, promovendo uma linguagem comum que facilita a interoperabilidade e a integração eficiente das informações.
4	Treinamento e capacitação dos profissionais de saúde do INCA.	Na percepção dos entrevistados, será necessário que o INCA invista em programas de formação para capacitar os seus profissionais com as habilidades necessárias para operar e manter os sistemas integrados e interoperáveis no ambiente tecnológico do INCA de forma segura e padronizada. Além disso, fomentar a conscientização e a aprendizagem colaborativa entre os profissionais que atuam nestas frentes, será importante para mitigar os riscos de rejeição quanto ao uso de tecnologias disruptivas na instituição.

Fonte: Elaboração própria, com base na análise dos dados coletados no INCA (2024).

Os dados revelam que, enfrentar de forma estratégica os desafios de integração e interoperabilidade dos sistemas de saúde, aumenta a probabilidade do INCA ser percebido como uma referência na manutenção do ecossistema de dados de saúde pública, onde as instituições trabalham de forma integrada e colaborativa para promover uma compreensão real e abrangente das condições de saúde da população, buscando aprimorar os seus processos e serviços com base nos conhecimentos gerados pela análise destes dados.

Adotar uma abordagem proativa e bem planejada para superar esses desafios, pode ser compreendido não apenas como uma necessidade técnica para o INCA, mas uma estratégia para que o instituto possa extrair valor do uso do *big data* em seus processos e práticas, aumentando as chances de que a adoção de inovações tecnológicas na instituição impactem positivamente na melhoria contínua dos serviços de saúde prestados pelo INCA ao paciente com câncer.

Ao melhorar a eficiência e a qualidade do tratamento oncológico, o INCA não apenas eleva o padrão da prática em saúde, mas também contribui para uma saúde pública mais equitativa, alinhada aos valores do SUS de garantir o direito à saúde e o acesso igualitário aos serviços de saúde para toda a população.

5.4 CP 2: Benefícios potenciais da adoção do *big data* no INCA.

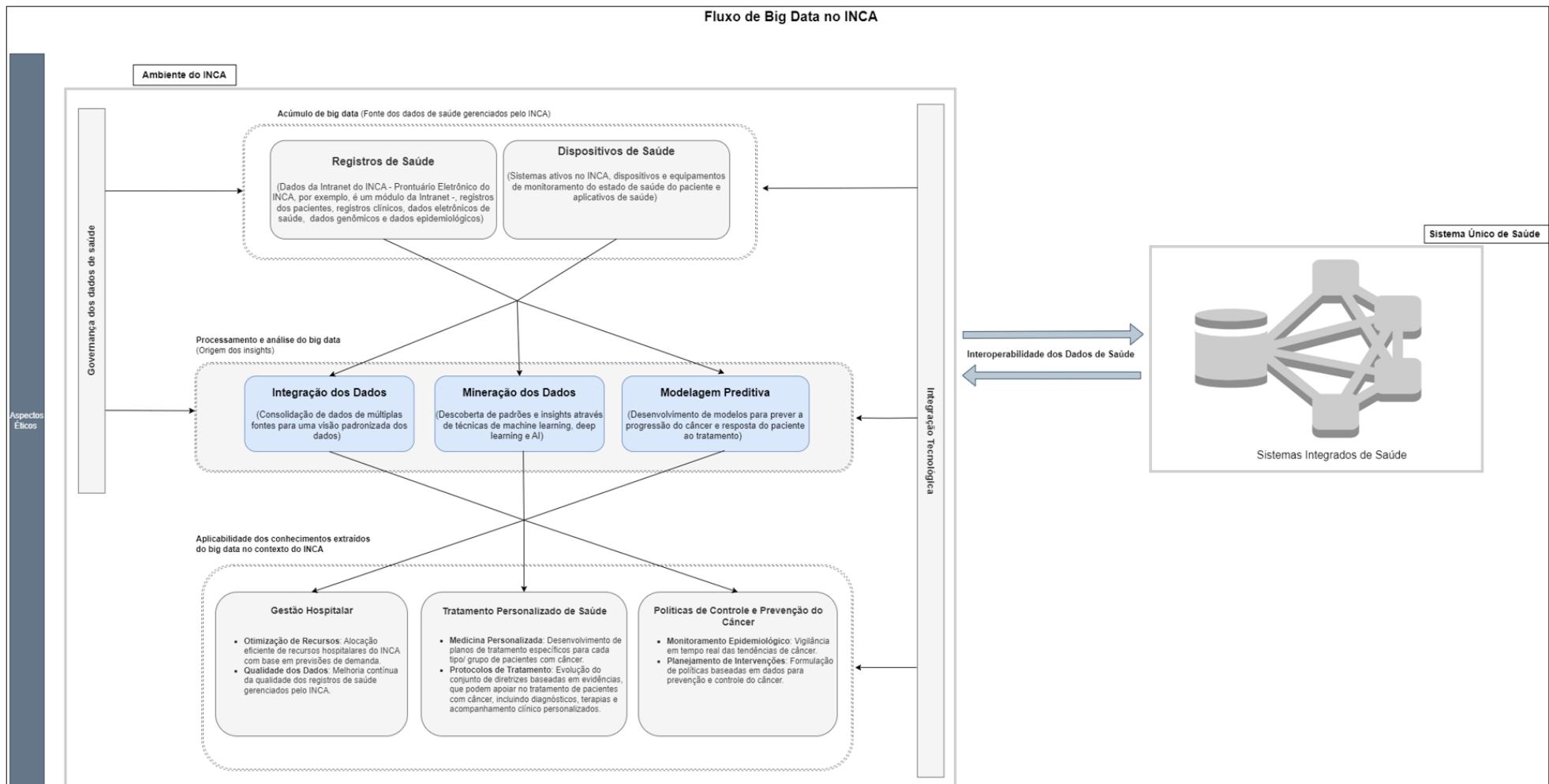
A análise dos dados revelou que, a implementação do *big data* no INCA pode aprimorar os serviços de saúde prestados pela instituição ao paciente com câncer e agregar valor às ações de controle e prevenção do câncer promovidas pelo Instituto. Entre os principais benefícios identificados estão a otimização dos processos administrativos, a detecção precoce do câncer, o monitoramento contínuo dos pacientes, a melhoria na tomada de decisões clínicas e avanços substanciais na pesquisa oncológica.

“Quando falamos em implementação do *big data* no INCA, entendo que não se trata apenas de ter máquinas que processam grandes volumes de dados de saúde; a verdadeira revolução acontece na maneira como o Instituto irá gerenciar e extraír valor desses dados. É a forma como aplicamos essas informações que realmente traz benefícios significativos para a sociedade” (Entrevistado 2).

Os participantes destacaram que a integração de tecnologias desempenha um papel fundamental no processo de conversão das informações extraídas do *big data* em

ações tangíveis para o Instituto, tendo em vista que, quando o dado bruto do *big data* é combinado com tecnologias de alta performance, informações de valor podem ser extraídas por meio da análise do *big data*. Estes conhecimentos, podem ser aplicados pelo INCA em diferentes tarefas relacionadas a prática em saúde, aumentando a eficiência da gestão dos processos institucionais e apoiando nas decisões estratégicas do Instituto.

A Figura 13 foi elaborada com o objetivo de apresentar uma síntese de como os participantes percebem a aplicabilidade prática dos conhecimentos extraídos do *big data* no contexto do INCA.

Figura 13: Fluxo de *Big Data* no INCA.

Fonte: Elaboração própria, com base nos dados coletados no INCA (2024).

No universo dos sujeitos entrevistados, foi possível observar que, especificamente no contexto do INCA, o *big data* é percebido como um processo dinâmico e integrado, que abrangeia desde a coleta de diferentes tipos de dados pessoais, clínicos e laboratoriais dos pacientes até a transformação desses dados em insights aplicáveis à realidade do Instituto. Esse processo se iniciaria com a coleta de dados de saúde provenientes de diversas fontes, como sistemas de registro de pacientes, sistemas de gestão de dados clínicos e gerenciais, incluindo a Intranet, registros eletrônicos de saúde, sistemas de imagem e sistemas legados. Após a coleta, os dados seriam armazenados em plataformas escaláveis e seguras, permitindo o gerenciamento eficiente de grandes volumes e tipos variados de dados.

Em seguida, esses dados passariam por um processo de tratamento e integração, consolidando informações de múltiplas fontes para criar uma visão padronizada dos dados. O próximo passo envolveria a aplicação de metodologias e ferramentas analíticas avançadas, como aprendizado de máquina e modelagem preditiva, para identificar padrões e tendências relevantes para o INCA. Os resultados dessas análises seriam então visualizados por meio de dashboards e relatórios inteligentes, o que permitiria aos profissionais de saúde explorar os dados de maneira intuitiva e tomar decisões informadas.

Finalmente, os *insights* obtidos seriam utilizados para melhorar o atendimento, otimizar recursos e desenvolver novas estratégias de tratamento e prevenção do câncer, em um ciclo contínuo de aprimoramento que ajustaria e refinaria práticas e processos com base nas novas informações geradas. Ao longo de todo esse processo, as políticas internas de governança de dados do INCA orientariam a coleta, o armazenamento e o compartilhamento das informações sensíveis dos pacientes, garantindo que essas práticas estariam em conformidade com a LGPD e preservariam a integridade e os direitos dos pacientes.

O Quadro 9 buscou apresentar uma síntese da percepção dos entrevistados sobre como o uso do *big data* pode vir a gerar benefícios práticos para o INCA em diversas áreas, promovendo uma gestão mais eficiente e impactando positivamente na qualidade dos serviços de saúde prestados ao paciente.

Quadro 11: Síntese sobre a aplicabilidade dos *insights* extraídos do *big data* no INCA.

Área de Aplicação	Insight do Big Data	Aplicabilidade Prática
Diagnóstico e Tratamento	Análise avançada de dados operacionais e administrativos do INCA.	<p>Otimização de fluxos de trabalho: Utilização de técnicas de análise preditiva e <i>machine learning</i> para otimizar o fluxo de gestão dos pacientes, melhorando a alocação de recursos e reduzindo o tempo de espera do paciente.</p> <p>Gestão eficiente de recursos: Análise dos dados históricos de gestão para prever a demanda de leitos, equipamentos e recursos humanos, permitindo a alocação eficiente dos ativos do INCA.</p> <p>Redução de custos operacionais: Identificação de áreas problemáticas, por meio da análise de dados financeiros e operacionais do Instituto, apoiando na implementação de medidas para identificar oportunidades de redução dos custos operacionais e logísticos, sem comprometer a qualidade dos serviços prestados ao paciente com câncer.</p>
	Monitoramento de indicadores estratégicos e utilização de recursos institucionais.	<p>Desenvolvimento de <i>dashboards</i> dinâmicos: Criação de painéis interativos que exibem indicadores-chave do INCA em tempo real, não apenas para atender as necessidades da alta gestão do Instituto, mas também para melhorar os processos operacionais da instituição, permitindo que os profissionais do INCA explorem os dados de saúde de maneira intuitiva para tomar decisões com base em evidências.</p> <p>Análise de tendências e padrões: Aplicação de análise de séries temporais para identificar padrões sazonais e tendências nos dados de saúde gerados pelo INCA, apoiando na identificação de oportunidade de melhorias operacionais para otimizar o uso de recursos e prestação dos serviços.</p>

Área de Aplicação	Insight do Big Data	Aplicabilidade Prática
		Avaliação e melhoria contínua: Implementação de ciclos de monitoramento da qualidade dos serviços do INCA baseados em dados, para continuamente avaliar e melhorar o desempenho institucional e a utilização dos recursos da instituição.
	Previsão de demanda e planejamento de recursos	<p>Modelagem preditiva avançada: Uso de modelos de previsão que combinam dados históricos, demográficos e sazonais para prever tendências de saúde, como o aumento de incidência de determinado tipo de câncer, ajudando no planejamento estratégico e na preparação de contingências para suportar o aumento da demanda de assistência ao paciente de forma proativa.</p> <p>Ajuste dinâmico de capacidade: Implementação de sistemas que ajustam automaticamente a capacidade e alocação de recursos do INCA, com base em previsões de demanda e análise temporal dos dados gerenciados pela instituição.</p>
Tratamento Personalizado de Saúde	Identificação de padrões em dados clínicos e genéticos dos pacientes com câncer	<p>Medicina Personalizada: Desenvolvimento de algoritmos que analisam dados clínicos e genéticos para criar perfis de risco e controle de resposta ao tratamento do câncer, permitindo a adoção de práticas clínicas personalizadas e focadas em atender as necessidades de saúde de cada indivíduo.</p> <p>Integração de múltiplos tipos de dados de saúde: Combinação de dados clínicos, genéticos e de estilo de vida do indivíduo para obter uma visão abrangente do estado de saúde do paciente, facilitando a personalização do tratamento e a identificação de novas terapias para conter a evolução da doença.</p> <p>Análise de eficácia de tratamentos: Avaliação contínua dos resultados dos tratamentos aplicados pelo INCA,</p>

Área de Aplicação	Insight do Big Data	Aplicabilidade Prática
	<p>Análise das respostas dos pacientes aos tratamentos aplicados pelo INCA</p>	<p>para ajustar as abordagens terapêuticas e identificar novos padrões e relações entre tratamentos aplicados e os resultados obtidos.</p> <p>Monitoramento dos dados de saúde em tempo real: Implementação de sistemas de monitoramento contínuo que capturam e analisam dados em tempo real sobre a resposta dos pacientes aos tratamentos e seus efeitos colaterais. Essa tecnologia permite que os profissionais de saúde do INCA façam ajustes rápidos e precisos nos planos de tratamento, apoiando a prática médica com informações atualizadas e detalhadas.</p> <p>Análise de reações adversas ao tratamento do câncer: Desenvolver módulos de análise inteligente, que utilizem os dados históricos e informações individuais dos pacientes para prever e mitigar efeitos colaterais durante o tratamento da doença. Esses perfis poderão ser criados nos sistemas de análise dos dados clínicos e de imagem do INCA, permitindo que os profissionais de saúde tomem decisões mais assertivas com base em uma análise detalhada do estado de saúde dos pacientes.</p> <p>Apoio nas decisões clínicas com base na análise dos dados de saúde: Utilização de algoritmos de aprendizado de máquina para aprimorar a análise de grandes conjuntos de dados dos pacientes, coletados ao longo da sua jornada no INCA. Neste sentido, a inteligência por trás dos sistemas do INCA fornecerá uma análise prévia do estado de saúde dos pacientes, tomando como base os índices estatísticos de controle e evolução do câncer, para aprimorar os tratamentos aplicados pelos profissionais de saúde ao paciente com câncer, aumentando as chances de cura da doença para este indivíduo.</p>

Área de Aplicação	Insight do Big Data	Aplicabilidade Prática
	Integração dos dados de saúde e histórico do paciente	<p>Registros de saúde integrados e interoperáveis: Viabilizar a integração dos dados de saúde provenientes de fontes sistêmicas ativas no INCA, incluindo histórico clínico, dados gerenciais, resultados laboratoriais e informações pessoais do paciente, proporcionando um maior controle e rastreabilidade do paciente em todas as etapas do tratamento.</p> <p>Prática multidisciplinares de cuidados ao paciente: Facilitar a integração entre diferentes grupos de profissionais com múltiplas especialidades clínicas para prestar assistência ao paciente do INCA de forma colaborativa, onde o uso dos dados tem como foco apoiar a prática médica para aumentar as chances de recuperação do paciente com câncer.</p>
Políticas de Controle e Prevenção do Câncer	Monitoramento de dados epidemiológicos e tendências de incidência	<p>Formulação de políticas de saúde baseadas em dados: Criação de políticas públicas de controle e prevenção do câncer baseadas em análises detalhadas de dados epidemiológicos, identificando tendências de aumento na incidência de determinado tipo de câncer e fatores de risco específicos para a população.</p> <p>Ajuste de estratégias de prevenção: Desenvolvimento de estratégias de prevenção e rastreamento baseadas em dados sobre a incidência e os padrões de câncer, permitindo uma abordagem mais direcionada e eficaz por parte do INCA.</p> <p>Análise espacial e geoespacial: O INCA poderá utilizar os dados geoespaciais para identificar regiões com alta incidência de câncer, para apoiar a formulação de políticas nacionais de prevenção, diagnóstico e tratamento do câncer capazes de atender as necessidades de saúde específicas de cada área.</p>

Área de Aplicação	Insight do Big Data	Aplicabilidade Prática
	Avaliação da eficácia de programas de rastreamento e campanhas de conscientização de prevenção do câncer	Avaliação com base nas evidências extraídas do <i>big data</i> , da eficácia dos programas de rastreamento e campanhas de conscientização de prevenção do câncer capitaneadas pelo INCA, utilizando métricas como taxas de detecção precoce e adesão ao programa, ajustando as abordagens para melhorar os resultados destas iniciativas a médio e longo prazo.
	Análise dos dados de saúde e aumento da qualidade de vida do paciente	<p>Monitoramento dos dados históricos de saúde dos pacientes: Análise contínua dos dados de saúde e qualidade de vida do paciente assistido pelo INCA, para avaliar o impacto das terapias e intervenções aplicadas a longo prazo, ajustando os protocolos de tratamento e suporte conforme necessário.</p> <p>Melhoria dos cuidados ao paciente com câncer: com base nos conhecimentos extraídos da análise dos dados de saúde coletados pelo INCA, a gestão do instituto poderá implementar ações estratégicas para apoiar os programas de suporte e acompanhamento assistencial ao paciente com câncer de forma proativa, aumentando o bem-estar clínico e psicológico do paciente.</p> <p>Avaliação e melhoria dos resultados clínicos: O INCA poderá utilizar o <i>big data</i> para melhorar continuamente os resultados clínicos e a qualidade de vida dos pacientes durante o tratamento da doença, aumentando as chances de sucesso das abordagens terapêuticas aplicadas ao paciente pelos profissionais de saúde da instituição.</p>

Fonte: Elaboração própria, com base nos dados coletados no INCA (2024).

Em relação aos benefícios gerados pelo uso do *big data* para a gestão dos serviços públicos de saúde, os dados das entrevistas sugerem que o uso estratégico do *big data* propicia a criação de um ambiente favorável para a implementação de políticas mais eficazes e alinhadas com a realidade

local. Neste sentido, os gestores do INCA podem utilizar as informações extraídas do *big data* para um melhor entendimento das condições de vida e saúde da população, compreendendo de forma mais abrangente as diferenças sociais e regionais em que os indivíduos estão inseridos.

“Essa compreensão aumenta as chances de que as práticas de saúde do Instituto promovam o princípio da equidade para a população que necessita dos serviços públicos de saúde” (Entrevistado 9).

Os entrevistados também destacaram que, o uso do *big data* apresenta potencial para apoiar o INCA na condução de estudos abrangentes e aprofundados sobre o câncer, devido à capacidade das tecnologias de *big data*, como *machine learning* e *deep learning*, para analisar grandes volumes de dados clínicos, genômicos e epidemiológicos.

“Atualmente no setor de Pesquisa e Bioinformática do INCA, o *big data* e *deep learning* já são uma realidade, tanto pela natureza do trabalho de pesquisa de ponta quanto pelo volume de dados na casa de *petabytes*. Sem as tecnologias que envolvem os conceitos de *big data*, o exercício de pesquisa de ponta seria inviável” (Entrevistado 11).

A importância da análise de dados para apoiar as pesquisas complexas em andamento no INCA, também emergiu como um tema relevante. Especificamente sobre a pesquisa genômica e sua relevância para o tratamento e detecção do câncer, os participantes explicaram que a utilização de técnicas para análise detalhada dos dados genômicos permite ao pesquisador extrair resultados clínicos para mapear com precisão as alterações genéticas associadas ao câncer, fornecendo uma base sólida para o desenvolvimento de tratamentos mais eficazes e direcionados para cada tipo de tumor.

“Na prática, o uso integrado dos dados não só eleva consideravelmente a qualidade do tratamento do câncer e a qualidade de assistência ao paciente, mas também abre novas possibilidades de descoberta na oncologia. Trata-se de uma fusão poderosa entre ciência e inovação, com potencial para ampliar a nossa compreensão sobre o tratamento e a evolução do câncer” (Entrevistada 5).

Dada a relevância da pesquisa genômica para o campo da oncologia, o setor de tecnologia e inovação do INCA tem trabalhado em parceria com a Coordenação de Pesquisa para apoiar as pesquisas em andamento no Instituto. Entre as ações destacadas, estão a criação de *datasets* para disponibilizar dados de saúde gerados no ambiente de produção do INCA aos pesquisadores de forma padronizada, onde tecnologias de *big data* como o *Pentaho* são utilizadas para automatização da extração de dados de pacientes dos sistemas RHC, EMS, Absolute, Anatomia Patológica, Intranet, entre outros sistemas ativos no Instituto, para que após a aplicação do tratamento estes registros sejam armazenados de maneira estruturada nas bases de dados homologadas pela instituição.

Essas medidas visam facilitar o acesso aos dados de saúde gerados pela instituição e promover um ambiente de pesquisa e aprendizagem colaborativa mais integrado e eficiente, acelerando os avanços na luta contra o câncer.

“Assim como a análise de processos internos e operacionais podem beneficiar na tomada de

decisões estratégicas, a análise e tratamento de dados clínicos podem e devem ajudar nas tomadas de decisão que envolvam os pacientes e seus tratamentos” (Entrevistado 3).

Há um consenso entre os entrevistados de que a adaptação ao uso de dados e tecnologias disruptivas, como o *big data*, para apoiar a prática clínica, requer uma mudança de paradigma. Essa transformação envolve a transição das abordagens tradicionais, que valorizam principalmente a experiência médica, para as atuais, que também incorporam dados e evidências obtidas por meio da análise dos dados de saúde dos pacientes.

“Essa transição não é apenas uma adaptação ao uso de novas ferramentas, mas uma transformação conceitual. Os profissionais de saúde estão agora em um cenário onde precisam se familiarizar e analisar criticamente a confiabilidade dos dados para apoiar suas práticas e isso não significa que os métodos tradicionais de análise dos dados são ineficientes, mas sim, que podem ser aprimorados e otimizados pelo uso de novas tecnologias” (Entrevistado 8).

Essa adaptação demanda do INCA um investimento contínuo em capacitação profissional para desenvolver as habilidades analíticas necessárias à interpretação e aplicação dos *insights* extraídos do *big data* na prática clínica. Os entrevistados alertaram que, atualmente, não percebem iniciativas do INCA voltadas para a capacitação e treinamento dos profissionais de saúde no uso do *big data* ou tecnologias complementares.

“Não tenho conhecimento sobre treinamentos direcionados ao uso de ferramentas de *big data* hoje no INCA, ainda que voltadas para apresentação prática do uso dessas ferramentas e seus possíveis benefícios para o Instituto” (Entrevistado 10).

Diferente do observado na literatura de Carter (2011), Ohlhorst (2012), Beyer e Laney (2012), Lupton e Mike (2016), Knapp (2018), Zhang (2018), Gandomi (2018), Oiszak (2018), Gupta et al (2019), Al Mayahi (2019), Agrawal e Prabakaran (2020), Sayantan et al (2020) e Hripcak (2021), onde não há um consenso em relação ao conceito de *big data*; no contexto do INCA, o *big data* é percebido como um conjunto de tecnologias e metodologias avançadas que viabilizam o armazenamento, processamento e análise de grandes volumes de dados complexos. Esses dados são gerados continuamente, tanto internamente, por meio dos processos do próprio INCA, quanto externamente, pela rede de saúde que interage com o instituto. Ao analisar o *big data* e extrair *insights* de valor, o INCA não apenas poderá elevar a qualidade dos cuidados ao paciente com câncer ao aplicar estes conhecimentos, mas também pavimentar um caminho de avanços significativos na pesquisa e na política de saúde pública, solidificando seu papel de destaque no sistema de saúde nacional.

As oportunidades identificadas reforçam o potencial que a adoção de tecnologias de análise inteligente dos dados combinadas com o uso estratégico do *big data*, possuem para apoiar o INCA

no alcance dos seus objetivos estratégicos a médio e longo prazo, aumentando as chances de que a instituição esteja melhor preparada para enfrentar os desafios contemporâneos da saúde, estabelecendo novos paradigmas nas ações de controle, diagnóstico e prevenção do câncer, ampliando as fronteiras de conhecimento no campo da oncologia.

5.5 CP 3: Preocupações éticas relacionadas à implementação do *big data* no INCA.

A implementação do *big data* no INCA levanta diversas preocupações éticas, especialmente no que diz respeito à privacidade dos pacientes. A capacidade de coletar e analisar grandes volumes de dados pode expor os indivíduos a formas invasivas de vigilância e monitoramento dos dados pessoais, especialmente em cenários onde os dados sensíveis são coletados sem o conhecimento ou consentimento explícito dos usuários, aumentando o risco de abusos por parte das organizações.

Os entrevistados destacaram que muitas instituições ainda não implementaram práticas adequadas de consentimento para uso dos dados dos pacientes, como preconiza a LGPD (2018). Essas práticas incluem a adoção de políticas e diretrizes específicas de segurança dos dados individuais, afim de evitar a violação de direitos e a falta de transparência nas práticas de coleta e uso de dados de saúde, que podem aumentar os riscos de que essas informações sejam utilizadas para gerar análises enviesadas no setor de saúde.

“Com a crescente utilização do *big data* na saúde, torna-se essencial que o INCA incorpore princípios éticos à implementação dessas tecnologias na instituição, implementando diretrizes e princípios que orientam a coleta, processamento e análise dos dados de maneira responsável” (Entrevistado 11).

Os dados dos entrevistados indicam a necessidade do INCA adotar abordagens proativas em suas práticas de manipulação e gerenciamento dos dados de saúde, para que o Instituto seja capaz de antecipar e prevenir incidentes de violação de privacidade e acesso indevido aos dados institucionais com maior eficiência.

“Para isso, será necessário que ocorra uma mudança de paradigma, onde as ações não mais deverão ter como foco remediar o fato depois de ocorrido, mas sim em proteger o nosso ambiente tecnológico de forma contínua, considerando o aumento frequente de casos de violação de dados em instituições de saúde públicas relevantes no cenário nacional” (Entrevistada 5).

Essa perspectiva remete a Sayantan et al (2020) que afirmam que no setor de saúde há uma percepção comum de que as diretrizes éticas são obstáculos burocráticos, em vez de guias essenciais para práticas responsáveis em saúde. Na percepção dos autores, parte das instituições de saúde os veem como exigências adicionais que complicam os processos e aumentam os custos, em vez de reconhecerem seu valor na proteção dos direitos individuais. No caso do INCA, foi possível observar que a ética na manipulação de informações sensíveis deve ser incorporada como um valor central, e

não apenas como um requisito regulatório para a gestão dos dados institucionais.

Os entrevistados também destacaram que, a implementação da LGPD no setor de saúde como um todo, representou avanços significativos para a proteção dos dados dos pacientes, uma vez que estas políticas trouxeram conceitos importantes, como o direito ao esquecimento e a portabilidade dos dados, fortalecendo o controle dos indivíduos sobre suas informações pessoais. Também foi possível perceber uma preocupação crescente em relação ao treinamento das ferramentas analíticas para análise dos dados, que a depender da forma como é feita pode gerar vieses em relação aos resultados de saúde extraídos do *big data*.

“Vieses nos dados de saúde podem surgir devido à forma como os dados de saúde são coletados, analisados e interpretados. Se os dados utilizados para treinar algoritmos de big data não forem representativos ou estiverem enviesados, isso pode resultar em conclusões equivocadas ou injustas para parte da população” (Entrevistado 6).

Um outro exemplo é que, se os dados de saúde coletados pelo INCA forem coletados predominantemente de um grupo demográfico específico, as análises resultantes podem não refletir com precisão a realidade de outros grupos, levando a tratamentos desiguais e perpetuando desigualdades existentes no sistema de saúde. Vieses também podem surgir devido a dados incompletos ou de baixa qualidade, onde os resultados da análise podem favorecer um grupo em detrimento de outro.

“Se os registros de saúde utilizados para treinar algoritmos de diagnóstico de câncer tivessem majoritariamente amostras de dados de populações urbanas, as particularidades de saúde das populações rurais seriam deixadas em segundo plano” (Entrevistado 8).

Especificamente no caso do INCA, o uso ético do *big data* poderá incluir a adoção de ações transparentes no processo de manipulação dos dados dos pacientes e nas práticas de saúde da instituição, promoção ao uso seguro dos dados sensíveis e adoção de práticas que fortaleçam a importância do tratamento adequado dos dados para evitar vieses nas análises do Instituto, conforme apresentado no Quadro 12.

Quadro 12 - Boas práticas para o uso ético do *big data* no INCA.

Boa Prática	Descrição
Adoção de práticas transparentes no gerenciamento dos dados de saúde	Será necessário que o INCA adote práticas transparentes de coleta e uso de dados sensíveis dos pacientes, informando claramente aos pacientes como suas informações serão utilizadas. Programas de conscientização sobre a importância dos valores éticos na manipulação dos dados de saúde poderão apoiar os profissionais da instituição neste processo.

Boa Prática	Descrição
Consentimento do paciente para uso dos dados	Será importante que os profissionais do INCA compreendam a importância de solicitar o consentimento dos pacientes para utilização dos seus dados pessoais e de saúde de forma explícita, garantindo que eles compreendam integralmente as implicações do compartilhamento de seus dados com a instituição. Neste contexto, ferramentas de consentimento dinâmico que permitem aos pacientes gerenciar suas preferências de privacidade continuamente podem apoiar nesta frente.
Responsabilidade organizacional	Será necessário que o INCA adote uma abordagem proativa em relação à ética na análise dos dados de saúde, integrando princípios éticos e democráticos em todas as fases do desenvolvimento e operação de tecnologias de <i>big data</i> no instituto. Treinamentos regulares e a criação de comitês de ética internos podem ajudar a promover uma cultura de responsabilidade no Instituto.
Mitigação de vieses na análise dos dados de saúde	É fundamental que o INCA implemente medidas para identificar e mitigar vieses nos dados e análises de <i>big data</i> . Isso poderá incluir a diversificação das fontes de dados, a validação contínua dos algoritmos e a revisão periódica dos resultados para garantir a equidade.

Fonte: Elaboração própria, com base nos dados dos entrevistados (2024).

No caso do INCA, os dados analisados indicam que a implementação de práticas éticas para a manipulação segura dos dados de saúde exige um esforço coordenado entre a alta gestão, os profissionais de saúde e os profissionais de TI do Instituto. Esse esforço conjunto não apenas garante a segurança e a confiabilidade no tratamento das informações de saúde, mas também aumenta a conscientização sobre a importância de as instituições de saúde zelarem pela proteção uniforme dos direitos dos indivíduos em um mundo cada vez mais interconectado.

5.6 CP 4: O uso do *big data* para tomada de decisões estratégicas no INCA.

A adoção do *big data* no INCA é percebida pelos sujeitos entrevistados como um processo complexo que exige uma abordagem multidisciplinar e bem estruturada, envolvendo a criação de um conjunto organizado de diretrizes e ações planejadas. Esse processo abrange desde a modernização da infraestrutura tecnológica até a necessidade de investir na capacitação dos profissionais da

instituição para compreender e lidar com as tecnologias de análise de *big data* de forma eficaz.

Esse processo envolve diferentes etapas, incluindo a avaliação detalhada das necessidades de gestão, operação e logística do INCA, além das definições estratégicas para a adoção do *big data* no Instituto. É fundamental estabelecer objetivos claros e mensuráveis para que sua implementação gere valor para as práticas de saúde da instituição.

Além disso, a adoção de ações coordenadas entre os diferentes atores envolvidos neste processo é fundamental para que o Instituto seja capaz de escolher e integrar ferramentas e plataformas adequadas para o gerenciamento e análise de grandes volumes de dados de saúde. No caso do INCA, é essencial garantir que essas tecnologias atendam às necessidades de análise e sejam compatíveis com a infraestrutura tecnológica da instituição, evitando a subutilização desses ativos. Superar os desafios relacionados à governança dos dados e manter a conformidade com as regulamentações e normas de proteção de dados da LGPD (2018) também são percebidos como pilares importantes para que esta iniciativa promova melhorias nos serviços de saúde oferecidos pelo Instituto de maneira segura e eficaz.

Os dados coletados sugerem que, no contexto do INCA, o primeiro passo essencial para a implementação do *big data* é a definição de objetivos e metas estratégicas claras. Esta etapa tem por objetivo fornecer uma direção sólida para o planejamento e execução do projeto, baseando-se nos objetivos estratégicos de médio e longo prazo da instituição. É importante que a estratégia adotada tenha um olhar especial no que tange à melhoria da tomada de decisões clínicas e organizacionais, na otimização da eficiência operacional, no aprimoramento da qualidade do serviço prestado ao paciente e no impulso à pesquisa médica.

De acordo com a perspectiva dos entrevistados, será necessário realizar uma avaliação detalhada da infraestrutura tecnológica existente no INCA, tendo em vista que os recursos atuais são destinados a atender as demandas do Instituto. Esta avaliação poderá incluir a verificação da capacidade de armazenamento, processamento e integração de dados, além de um plano de ação para prover a interoperabilidade dos sistemas de informação de saúde ativos na Instituição. Identificar e atualizar os equipamentos obsoletos também será fundamental para garantir que a instituição esteja preparada para lidar com o volume e a complexidade dos dados gerados.

“Para aproveitar todo o potencial do *big data* no INCA, precisamos preparar nossa infraestrutura tecnológica. Isso inclui, por exemplo, modernizar nossos sistemas legados para garantir que os dados gerados pelo INCA sejam trocados de forma integrada e consistente entre todos os sistemas ativos da instituição. Dessa maneira, evitamos a fragmentação dos registros, que pode dificultar a análise e manipulação dos dados, e reduzimos o risco de falhas no armazenamento dessas informações” (Entrevistado 7).

Os entrevistados também destacaram a importância da coleta e integração dos dados de saúde como etapas fundamentais no processo de implementação do *big data*. Neste sentido, será importante que o INCA defina claramente suas fontes de dados, como registros eletrônicos de saúde, dados de dispositivos médicos, informações genômicas e dados de pesquisa clínica, para criar um ambiente propício à análise do *big data*, para que esta ação seja capaz de gerar *insights* significativos e confiáveis para o Instituto.

A qualidade dos dados de saúde foi outro ponto de atenção identificado pelos entrevistados. Para garantir a integridade, precisão e consistência dos dados coletados, é essencial que o INCA implemente processos rigorosos para padronizar os formatos de registro, eliminar informações redundantes ou inconsistentes e aplicar técnicas de higienização e validação dos dados gerenciados pela instituição.

Com a garantia da qualidade dos dados, a aplicação de técnicas avançadas de análise e mineração de dados, como algoritmos de aprendizado de máquina, inteligência artificial e análise preditiva, torna-se ainda mais eficaz. Essas técnicas foram apontadas como essenciais para que o INCA seja capaz de extrair conhecimentos valiosos do *big data*. Elas podem apoiar a instituição na identificação de padrões e tendências do câncer com maior agilidade, gerando *insights* relevantes para o campo da oncologia e auxiliando na tomada de decisões estratégicas da instituição.

Além disso, a segurança dos dados dos pacientes foi uma preocupação recorrente entre os entrevistados. Garantir a privacidade e segurança dos dados é um aspecto crítico na implementação do *big data* em um setor complexo como o de saúde, por isso, será importante que o INCA adote práticas adequadas para o tratamento de dados sensíveis, além de implementar mecanismos de segurança robustos para minimizar o risco de acesso não autorizado as informações de saúde dos pacientes.

Os dados das entrevistas apontam que, no contexto do INCA, a criação de grupos multidisciplinares é uma estratégia eficaz para minimizar a resistência ao uso do *big data* durante o seu processo de implementação. Para isso, será importante que a alta gestão do Instituto apoie a formação de comitês ou grupos de trabalho que envolvam profissionais de saúde, especialistas em tecnologia, pesquisadores e gestores. Esses grupos trabalharão em conjunto na identificação de desafios e na busca por soluções, promovendo uma visão integrada e colaborativa do projeto. Essa abordagem colaborativa poderá contribuir significativamente para a aceitação e sucesso das iniciativas de *big data* no INCA.

A implementação do *big data* no INCA é percebida pelos profissionais de saúde do INCA como um processo contínuo, onde avaliações regulares são necessárias para medir o impacto das

iniciativas de *big data*, identificar oportunidades de melhoria e ajustar a estratégia conforme necessário. Neste processo, manter uma comunicação ativa entre os diferentes atores envolvidos nesta frente será fundamental para que o Instituto possa compreender com clareza os impactos do uso do *big data* em suas práticas e possa direcionar de forma assertiva os próximos passos.

No contexto de uma instituição de saúde pública como o INCA, as evidências indicam a necessidade de adaptações específicas, que incluem a colaboração entre agências governamentais e instituições da rede de saúde para fortalecer as políticas de incentivo ao compartilhamento de informações e recursos tecnológicos no setor de saúde. Além disso, criar um ambiente propício para a integração e a interoperabilidade dos sistemas pode ser a base para que as instituições obtenham uma visão holística do estado de saúde da população, permitindo uma análise mais completa e precisa das informações e, consequentemente, gerar políticas de saúde pública mais assertivas, com ênfase em atender as necessidades dos pacientes.

Tomando como base a análise dos dados, a Figura 14 buscou apresentar a convergência das ações que poderão apoiar o INCA no processo de implementação do *big data* no instituto.

Figura 14- Etapas para implementação do *big data* no INCA.



Fonte: Elaboração própria, com base nos dados das entrevistas (2024).

Por fim, as evidências sugerem que a implementação do *big data* no INCA será um processo contínuo e colaborativo, que precisará ser constantemente avaliado e adaptado às mudanças tecnológicas e às necessidades dos pacientes e profissionais de saúde da instituição. Além disso, será importante que a gestão do Instituto defina com clareza os objetivos a serem alcançados com a adoção de tecnologias complexas como o *big data* em suas práticas. A partir desse entendimento, o INCA poderá adotar estratégias inteligentes para garantir o sucesso dessa iniciativa, mitigando os riscos de que o projeto não avance além da etapa de planejamento ou que, mesmo implementado, a instituição não consiga aproveitar todo o potencial dessas ferramentas em seus processos e atividades operacionais.

Com base nas percepções discutidas, o próximo capítulo apresentará um plano estratégico com as ações práticas que poderão ser adotadas pelo INCA para a implementação do *big data* na instituição.

6. PLANO ESTRATÉGICO PARA A IMPLEMENTAÇÃO DO BIG DATA NO INCA.

Neste capítulo, será descrito o Plano Estratégico resultante desta pesquisa aplicada, classificado conforme a metodologia de avaliação da Produção Técnica e Tecnológica proposta pela CAPES (2019).

6.1 Apresentação do Produto Tecnológico: Plano Estratégico para Implementação do Big Data no INCA

O Quadro 13 apresenta a classificação e caracterização do Plano Estratégico desenvolvido, com base nas especificações delineadas no Relatório de Produção Técnica e Tecnológica da CAPES (2019).

Quadro 13 - Detalhamento do Produto Tecnológico.

Identificação	Descrição
Organização	Universidade Federal Rural do Rio de Janeiro - UFRRJ
Discente	Luana da Silva Alves
Docente orientador	Prof. Dr. Saulo Barbará de Oliveira
Docente coorientador	Prof. Dr. Sandro Luís Freire de Castro Silva
Programa de pós-graduação vinculado	Programa de Pós-graduação em Gestão e Estratégia (PPGE/UFRRJ)
Linha de pesquisa vinculada à produção	Linha 1- Gestão de Processos, Projetos e Tecnologias
Dissertação vinculada	Plano de Recomendações Estratégicas para a Adoção do Big Data na Saúde Pública: um estudo aplicado no Instituto Nacional de Câncer (INCA)
Data da defesa	09/09/2024
Setor beneficiado com o projeto de pesquisa, realizado no âmbito do programa de mestrado	Serviço de Tecnologia da Informação (SETI) do INCA, coordenações de ensino e pesquisa e demais áreas de gestão e estratégia do INCA.
Descrição da finalidade	O produto tecnológico desenvolvido consiste em um Relatório Técnico Conclusivo, cuja elaboração apresenta duas etapas. A primeira consiste na elaboração de um plano estratégico para apresentar estratégias que podem ser adotadas pelo INCA para apoiar no processo de implementação do <i>big data</i> no Instituto. A segunda tem por objetivo apresentar um plano de ação com

Identificação	Descrição
	<p>sugestões para mitigar os riscos associados a implementação do <i>big data</i> no INCA, tomando como base a análise dos dados coletados, conforme apresentado nesta dissertação.</p> <p>Este produto tecnológico, segundo a classificação CAPES, classifica-se no tipo Processos, Tecnologias, Produtos e Materiais.</p>
Avanços tecnológicos/grau de inovação	<p>A pesquisa científica realizada no Instituto Nacional de Câncer (INCA) permitiu compreender a perspectiva dos gestores, profissionais de saúde e profissionais de TI que atuam na instituição sobre as estratégias que podem ser adotadas para viabilizar a implementação do <i>big data</i> no Instituto.</p> <p>Para isso, foi necessário mapear as principais ações práticas que, na perspectiva dos sujeitos entrevistados, podem ser adotadas pelo INCA para aumentar as chances de uma implementação bem-sucedida do <i>big data</i> na instituição. Esta iniciativa visa não apenas a inovação tecnológica na gestão e análise de dados oncológicos, mas também a melhoria contínua dos processos internos do Instituto, cooperando para a criação de um ambiente de dados mais robusto e eficiente para a pesquisa e tratamento do câncer na instituição.</p> <p>Diante do exposto, o plano estratégico desenvolvido pode ser considerado uma produção com médio teor inovativo, uma vez que utilizou conhecimentos preestabelecidos para formular a solução.</p>
Reaplicabilidade da produção tecnológica	<p>Despertar o interesse das instituições de saúde para as recomendações estratégicas apresentadas neste estudo representa um desafio significativo, dado o cenário complexo e diversificado do setor de saúde pública nacional. No entanto, as diretrizes propostas foram elaboradas com base nos princípios de flexibilidade e escalabilidade, para facilitar sua adaptação a contextos organizacionais similares. Além disso, acredita-se que, à medida que mais pesquisadores se dedicarem a esse tema, envolvendo instituições públicas de saúde como objeto de estudo, a disseminação do conhecimento aumentará, aumentando a relevância do tema no campo da Pesquisa em Saúde.</p> <p>Essa expansão poderá reforçar a importância de as instituições adotarem padrões de gestão na saúde pública que priorizem a tomada de decisões estratégicas com base em evidências extraídas da análise dos dados, promovendo práticas que levem a um sistema de saúde mais igualitário, equitativo e justo para a população que necessita de cuidados.</p>
Complexidade da produção tecnológica	<p>O Plano Estratégico desenvolvido apresenta uma complexidade significativa, abrangendo múltiplas áreas de conhecimento, como Administração Pública, Saúde Pública e Tecnologia da Informação e Comunicação (TICs). A</p>

Identificação	Descrição
	necessidade de aprofundar os conhecimentos teóricos sobre as possibilidades de uso do <i>big data</i> em setores complexos como o da saúde e paralelamente considerando o ambiente do INCA para identificar os caminhos que podem viabilizar a implementação de uma infraestrutura de <i>big data</i> no Instituto, adicionou um nível extra de complexidade, exigindo uma pesquisa detalhada e atualizada sobre o uso de tecnologias de <i>big data</i> e seus impactos no setor de saúde nacional. Esse aprofundamento teórico e prático conferiu ao trabalho uma complexidade considerável, destacando sua relevância para a compreensão de como a gestão baseada em dados pode auxiliar o campo da Saúde Pública.
Financiamento	CAPES.
A produção necessita estar no repositório?	Sim, será divulgado internamente na organização examinada para consulta pública.
Documentos anexados (em PFD)	Relatório Técnico Conclusivo.

Fonte: Elaboração Própria, com base no Capes (2024).

6.2 Relatório Técnico Conclusivo: Propostas estratégicas para implementação do *big data* no INCA

6.2.1 Contextualização

Este Plano Estratégico tem como objetivo principal apresentar ações práticas que podem ser adotadas pelo INCA para a implementação do *big data* na instituição. Com base nos dados analisados ao longo deste estudo, identificou-se que, no setor de saúde, o *big data* se apresenta como uma ferramenta de valor, capaz de aprimorar a gestão dos dados de saúde, promover avanços significativos na pesquisa, diagnóstico e tratamento de doenças complexas como o câncer, além de aumentar a eficiência administrativa e operacional das instituições (Sayantan et al., 2020).

Isso ocorre porque o uso de tecnologias robustas de análise, combinadas com o *big data*, otimiza o processamento de grandes volumes de dados, apoiando os gestores de saúde na tomada de decisões. Como consequência, há uma alocação mais eficaz de recursos, identificação de padrões que podem otimizar processos internos, redução de custos e melhoria na qualidade dos serviços de saúde prestados pelas instituições de saúde (Suprya et al., 2021).

Especificamente no caso do INCA, identificou-se, com base nos dados das entrevistas, que o uso integrado das tecnologias de *big data* não apenas tem o potencial de aprimorar a tomada de decisões clínicas e estratégicas do Instituto, mas também de melhorar a qualidade dos serviços de saúde prestados aos pacientes com câncer.

O estudo que embasou a elaboração deste plano estratégico buscou descrever, com base na literatura, a relação entre tecnologia e o processo de evolução tecnológica na área da saúde, a fim de compreender como essas mudanças impactaram o setor e identificar os efeitos do aumento exponencial no volume de dados de pacientes resultante das práticas de saúde. Este estudo também investigou como o uso estratégico do *big data* tem sido aplicado no setor, para, a partir desse entendimento, apresentar um plano com estratégias que possam viabilizar a implementação do *big data* no INCA, por meio de um estudo aplicado.

A relevância deste estudo está em demonstrar como o uso do *big data* pode aprimorar a qualidade do tratamento de pacientes com câncer, fornecendo *insights* a partir da análise de grandes volumes de dados de saúde. Isso não significa que os métodos tradicionais de análise de dados sejam ineficazes, mas sim que podem ser aprimorados pela integração de tecnologias avançadas, como o *machine learning* e o *big data*, e pela *expertise* de profissionais capacitados para utilizar essas ferramentas de maneira eficaz no setor de saúde (LUPTON e MICHAEL, 2016). Essa combinação permite a identificação de padrões e correlações que seriam difíceis de detectar com abordagens convencionais, resultando em diagnósticos mais precisos e ágeis, tratamentos mais eficazes e um

monitoramento mais eficiente da progressão do câncer (PAIGE et al, 2021).

Consequentemente, o uso estratégico do *big data* pode aumentar as chances de sobrevivência dos pacientes e gerar valor para a área oncológica por meio das descobertas que podem ser extraídas desses dados. O impacto potencial do uso estratégico desta tecnologia, está diretamente relacionado à capacidade de utilizar os dados como uma ferramenta de apoio para a prática médica, contribuindo para a melhoria da qualidade de vida e a redução dos riscos de mortalidade dos pacientes atendidos pela instituição.

No entanto, autores como Sayantan et al. (2020) e Teixeira et al. (2021) destacam que o uso dessas tecnologias não substitui os profissionais de saúde, mas sim complementa o conhecimento prático desses profissionais. A integração entre as ferramentas de análise avançada e a *expertise* dos profissionais é essencial para gerar valor nas práticas de saúde, potencializando os benefícios tanto para os pacientes quanto para a instituição (TEIXEIRA et al, 2021).

Como procedimentos metodológicos, foi realizada uma pesquisa aplicada de abordagem qualitativa no Instituto Nacional de Câncer (INCA). Para a coleta de dados, recorreu-se a uma pesquisa documental que precedeu a realização de entrevistas com 11 profissionais do INCA, incluindo médicos, assistentes, gestores de planejamento estratégico e especialistas em tecnologia da informação. Os dados coletados foram analisados utilizando a técnica de análise de conteúdo, o que permitiu uma compreensão aprofundada dos desafios e oportunidades relacionados à implementação do *big data* no INCA.

Diante dessas informações, a próxima seção tem por objetivo apresentar, com base na percepção dos entrevistados, as estratégias que podem ser adotadas pelo INCA para a implementação do *big data* no instituto, utilizando a ferramenta de gestão 5W2H.

Cabe ressaltar que o desenvolvimento deste plano estratégico foi fundamentado nas percepções dos entrevistados e na análise das práticas organizacionais do INCA, apresentadas ao longo da sessão de resultados desta dissertação. Além disso, reitera-se que esta etapa é crucial para a conclusão da dissertação, conforme estabelecido pela Deliberação nº 161 de 12/11/2019 do Programa de Pós-graduação em Gestão e Estratégia, reforçando seu papel como mestrado profissional (UFRRJ, 2019).

Por fim, espera-se que as recomendações apresentadas possam apoiar o INCA em suas iniciativas para aprimorar os processos de gestão de dados, criando um ambiente tecnológico mais robusto e eficiente, e, assim, aumentando as chances de que a instituição seja capaz de extrair todo o potencial do *big data* em suas práticas, o que, por sua vez, fortalecerá o papel do INCA como uma instituição de referência na prevenção, controle, assistência e pesquisa de câncer no Brasil.

Os fluxogramas e recomendações apresentadas neste produto tecnológico, decorrente da pesquisa, serão entregues ao Serviço de Tecnologia da Informação do INCA, cabendo a autoridade competente da instituição decidir por sua institucionalização ou não.

6.2.2 Plano Estratégico para Implementação do Big Data no INCA - Ferramenta 5W2H

Para a elaboração do plano estratégico que visa apresentar os caminhos que poderão ser seguidos pelo INCA para implementação do *big data* no Instituto, foi utilizada a ferramenta de gestão 5W2H. Essa metodologia, baseada nas perguntas *why* (por que), *what* (o que), *where* (onde), *when* (quando), *who* (quem), *how* (como) e *how much* (quanto custa), tem por objetivo auxiliar as instituições na organização das ações estratégicas de maneira clara e objetiva, de modo a garantir que todos os aspectos importantes do processo sejam considerados (JÚNIOR et al., 2021).

O plano estratégico desenvolvido com base na 5W2H apresenta as seguintes ações que podem ser adotadas pelo INCA para a integração do *big data* em suas práticas e processos:

1. **Definição de objetivos específicos para a implementação do *big data*:** Estabelecer metas claras e tangíveis para o uso do *big data*, alinhadas com o planejamento estratégico do INCA, como a melhoria na precisão dos diagnósticos, a otimização dos tratamentos e o aumento da eficiência na gestão dos dados oncológicos gerados pela instituição.
2. **Capacitação e treinamento dos profissionais:** Investir em programas de formação para capacitar os profissionais de saúde do INCA, garantindo que eles estejam aptos a utilizar as novas ferramentas de *big data* de forma eficaz.
3. **Desenvolvimento e integração de infraestrutura tecnológica:** Implementar e integrar os sistemas de saúde ativos no ambiente do INCA, para que estes sejam capazes de lidar com grandes volumes de dados complexos, aumentando as chances de que a infraestrutura tecnológica da instituição seja compatível com as tecnologias de *big data*.
4. **Estabelecimento de políticas de governança de dados:** Criar políticas robustas para a gestão, segurança e privacidade dos dados gerenciados pela instituição, assegurando que todas as práticas estejam em conformidade com as normas regulamentares e éticas que regem o uso dos dados no setor de saúde.
5. **Alocação de recursos financeiros e técnicos:** Destinar recursos específicos para a aquisição de tecnologias, treinamento e suporte contínuo, além de criar um orçamento que cubra os custos de implementação e manutenção das soluções de *big data* no INCA.

6. **Monitoramento e avaliação contínua do processo:** Implementar um sistema de monitoramento para avaliar constantemente os resultados da utilização do *big data* na instituição, permitindo ajustes e melhorias contínuas ao longo do processo.

A seguir, o Quadro 14 apresentar o plano estratégico com as ações práticas que podem ser adotadas pelo INCA para implementação do *big data*.

Quadro 14 - Plano Estratégico para Implementação do Big Data no INCA: Estrutura 5W2H.

	O quê?	Por quê?	Quem?	Quando?	Onde?	Como?	Quanto?
1	Definição de objetivos específicos	Para garantir que o uso do <i>big data</i> esteja alinhado com os objetivos estratégicos do INCA, maximizando o impacto na precisão dos diagnósticos, na eficácia dos tratamentos e na eficiência da gestão de dados oncológicos, assegurando que cada ação contribua diretamente para as metas institucionais estabelecidas a médio e longo prazo.	Equipe de gestão e planejamento do INCA, profissionais de TI e da assistência com participação da alta direção do Instituto.	Durante a fase de planejamento estratégico, para assegurar que os objetivos específicos sejam claramente definidos antes do início da implementação do <i>big data</i> , permitindo que todas as ações subsequentes estejam alinhadas com os objetivos estratégicos do INCA.	Nas áreas estratégicas do INCA, envolvendo os principais setores da instituição envolvidos nesta frente, como a gestão executiva, o setor de tecnologia da informação, as coordenações de ensino e pesquisa e o setor da assistência, para garantir uma abordagem integrada e alinhada com as necessidades do INCA.	Por meio de reuniões estratégicas para a definição de metas.	A definir
2	Capacitação e treinamento dos profissionais	Para garantir que os profissionais do INCA possuam as habilidades e conhecimentos necessários para utilizar as ferramentas de <i>big data</i> de forma eficaz,	Pelo setor de Recursos Humanos em parceria com o setor de Tecnologia da Informação do INCA, com o apoio de	Após a definição dos objetivos específicos e antes da implementação das tecnologias de <i>big data</i> no INCA. Será importante que o treinamento inicial seja concluído nos	Treinamento presencial no INCA e por meio de plataformas <i>online</i> .	Programas de formação, <i>workshops</i> e treinamentos contínuos.	A definir

	O quê?	Por quê?	Quem?	Quando?	Onde?	Como?	Quanto?
		maximizando o valor dessas tecnologias na prática clínica e na gestão de dados gerados pela instituição.	consultores externos especializados em <i>big data</i> e <i>analytics</i> . Além disso, será envolvida a liderança das áreas clínicas e administrativas do INCA para garantir que o treinamento seja relevante e adaptado às necessidades específicas dos diferentes setores da instituição.	primeiros meses do projeto, com sessões de atualização e capacitação contínua programadas ao longo da implementação e manutenção das soluções de <i>big data</i> , aumentando as chances de que os profissionais permaneçam atualizados com as melhores práticas e utilização das ferramentas de <i>big data</i> .			
3	Desenvolvimento e integração de infraestrutura	Para assegurar que a infraestrutura do INCA seja compatível com as tecnologias de <i>big data</i> .	Serviço de TI do INCA, com suporte de fornecedores de tecnologia.	Sincronizado com a capacitação dos profissionais.	No INCA, com ênfase nas áreas críticas de TI.	Atualização dos sistemas legados e aquisição de novos equipamentos e sistemas no INCA.	A definir
4	Estabelecimento de políticas de	Para garantir a segurança e privacidade dos dados	Comitê de governança de dados do INCA, com participação do	Paralelamente a etapa de estruturação da infraestrutura	Em todas as unidades do INCA onde os dados são coletados, armazenados,	Elaboração de políticas internas, revisões periódicas e alinhamento	A definir

	O quê?	Por quê?	Quem?	Quando?	Onde?	Como?	Quanto?
	governança de dados	gerenciados pela instituição.	setor de Tecnologia da Informação, gestão estratégica e alta gestão do Instituto.	tecnológica do INCA para implementação do <i>big data</i> .	processados e compartilhados. Isso inclui as áreas de TI, unidades clínicas, administrativas e de pesquisa da instituição.	com regulamentações alinhadas com as regulamentações de proteção dos dados no Brasil.	
5	Alocação de recursos financeiros e técnicos	Para garantir que os recursos necessários estejam disponíveis para a implementação e manutenção do <i>big data</i> no INCA.	Setor de Planejamento e Gestão Estratégica do INCA, em parceria com o setor de TI e alta gestão do Instituto.	Durante a fase de planejamento e execução do projeto de implementação do <i>big data</i> na instituição.	No INCA, envolvendo todas as partes interessadas no projeto.	Elaboração de um planejamento estratégico e financeiro específico para o projeto de <i>big data</i> no INCA.	A definir
6	Monitoramento e avaliação contínua do processo	Para avaliar constantemente os resultados do uso do <i>big data</i> nos processos da instituição e se necessário, realizar ajustes ao longo do tempo.	Setor de TI do INCA.	Em todas as etapas do projeto de implementação do <i>big data</i> no INCA.	No ambiente do INCA.	Implementação de um sistema de monitoramento, com métrica e indicadores que refletem a realidade da instituição e avaliações regulares do uso do <i>big data</i> no INCA.	A definir

Fonte: Elaboração própria, com base na análise dos dados coletados em campo (2024).

O plano estratégico proposto buscou apresentar ações que podem ser adotadas pelo INCA para integrar o uso do *big data* em suas práticas e processos institucionais, considerando as observações feitas durante a pesquisa de campo e as percepções dos entrevistados. Embora as ações propostas tenham como foco principal a implementação do *big data* no INCA, elas também visam garantir que a instituição esteja preparada para explorar ao máximo o potencial dessas ferramentas, promovendo melhorias nos processos clínicos, administrativos, de pesquisa e de gestão em geral.

Para aumentar as chances de que os profissionais do INCA estejam preparados para utilizar as ferramentas de *big data* na rotina da instituição, foram propostas a implementação de iniciativas de capacitação e treinamento contínuos dos colaboradores. Essas ações poderão incluir a realização de cursos específicos sobre análise de dados, uso de ferramentas de *big data*, e interpretação e demonstração dos *insights* gerados por meio da análise dos dados. Os profissionais de TI e das áreas clínicas, em particular, necessitam de uma atenção maior, tendo em vista que estes serão os principais responsáveis pela implementação e operação dessas tecnologias na instituição, além de oferecer suporte técnico e orientação para outros membros da equipe.

A capacitação poderá ser complementada pela disseminação de conhecimento e debates regulares entre os profissionais do INCA envolvidos nesta frente, discutindo as melhores práticas na utilização de *big data* em saúde. Além disso, o INCA, em parceria com outros parceiros de tecnologia, poderá ofertar cursos atualizados, garantindo que os profissionais da instituição estejam sempre alinhados com as inovações e as normas aplicáveis a sua área de atuação.

O desenvolvimento e a integração da infraestrutura tecnológica, inicialmente focados no ambiente do INCA, mas projetados de forma escalável para permitir uma futura integração com os sistemas de saúde de outras instituições da rede pública, são ações fundamentais. Essas ações incluem a atualização dos sistemas legados e a integração desses sistemas com plataformas capazes de processar grandes volumes de dados complexos. Um planejamento detalhado e uma execução cuidadosa dessa integração serão essenciais para garantir que a infraestrutura do INCA seja robusta e compatível com futuras expansões.

A estruturação das políticas de governança de dados é outro fator importante para garantir a segurança, privacidade e integridade dos dados gerenciados pela instituição. Essas políticas serão desenvolvidas em colaboração com as diversas áreas do INCA, incluindo a equipe de TI, para assegurar que todas as necessidades e desafios sejam adequadamente considerados. O objetivo é alinhar essas políticas às regulamentações e melhores práticas do setor de saúde, criando um ambiente de governança que seja eficiente e funcional. Além disso, será implementado um programa de conscientização e treinamento para garantir que todos os profissionais compreendam e sigam as

diretrizes estabelecidas de maneira consistente e dentro dos padrões esperados pela instituição.

A alocação de recursos financeiros e técnicos adequados é um elemento essencial para garantir que o projeto conte com os meios necessários para sua implementação e manutenção. Como o projeto ainda está em fase de planejamento, os custos exatos para cada ação permanecem "A definir", o que permite flexibilidade na alocação de recursos, à medida que as necessidades específicas forem melhor compreendidas pela instituição e suas equipes responsáveis.

A implementação de um sistema de monitoramento e avaliação contínua aumentará a probabilidade de que o INCA acompanhe o progresso da implementação do *big data* e faça os ajustes necessários em tempo real. Esse monitoramento constante permitirá identificar áreas de melhoria e corrigir eventuais desvios, assegurando que as metas estabelecidas sejam alcançadas de maneira eficiente e alinhada aos objetivos institucionais.

O plano estratégico proposto buscou abordar os principais desafios identificados na etapa de coleta de dados, propondo soluções que possam fortalecer a capacidade do INCA de utilizar o *big data* para otimizar suas operações e melhorar o atendimento aos pacientes. Por meio da capacitação contínua, da modernização da infraestrutura, do estabelecimento de políticas de governança e da avaliação constante, espera-se que o INCA consiga aprimorar seus processos e serviços de maneira significativa.

Por fim, cabe ressaltar que a implementação das propostas aqui apresentadas fica a cargo do INCA. No entanto, espera-se que os caminhos apresentados neste estudo sejam capazes de apoiar a instituição no processo de implementação do *big data*, contribuindo para a sua consolidação como uma instituição de referência no uso de tecnologias avançadas no campos da oncologia.

6.2.3 Mapeamento do Processo para Implementação do Big Data no INCA

O mapeamento das etapas para a implementação do big data no Instituto Nacional de Câncer (INCA) é importante para garantir que as ações estratégicas delineadas pela instituição sejam executadas de forma coordenada e eficaz, assegurando que os resultados esperados sejam alcançados ao final do processo (SAYANTAN et al., 2020). Esta seção se propõe a apresentar de forma estruturada as diferentes etapas envolvidas nesse processo, desde o planejamento inicial até a integração das tecnologias de *big data*, por meio de um mapeamento adaptado ao contexto do INCA. A criação deste modelo teve como base a análise dos dados das entrevistas, conforme apresentado ao longo da seção de resultados desta dissertação.

A etapa inicial envolve a elaboração de um planejamento detalhado do projeto, que inclui a definição dos objetivos específicos que orientarão a implementação do *big data* no INCA. Essa fase

exige uma colaboração estreita entre os diferentes setores da instituição, de modo a garantir que as metas estabelecidas estejam alinhadas com os objetivos estratégicos do INCA. A definição desses objetivos constitui a base para que todas as ações subsequentes sejam direcionadas de forma coesa e estratégica.

Uma vez definidos os objetivos, a próxima fase consiste na capacitação dos profissionais que serão responsáveis pela implementação e operação das tecnologias de *big data* no INCA. Esse treinamento deve ser abrangente, abordando desde a análise de dados até o uso das ferramentas de *analytics* e a interpretação e apresentação dos insights gerados (CARDONA et al., 2022). A disseminação do conhecimento sobre as práticas adotadas pelos profissionais de saúde do INCA, bem como os benefícios e desafios do uso do *big data* nas rotinas da instituição, serão pilares fundamentais para garantir que os atores envolvidos neste processo estejam alinhados com as melhores práticas definidas pela instituição para a manipulação do *big data* (ADJEKUM et al., 2018).

Com os profissionais capacitados, a fase seguinte envolve o desenvolvimento e a integração da infraestrutura tecnológica necessária para suportar o *big data* no contexto do INCA. Essa etapa poderá incluir a atualização dos sistemas legados, a integração de novas plataformas de *big data* e a criação de um ambiente escalável, que permita futuras expansões e a integração com outros sistemas de saúde da rede pública. O planejamento e a execução cuidadosa dessa integração são cruciais para assegurar que a infraestrutura do INCA seja robusta e compatível com as demandas atuais e futuras do setor de saúde.

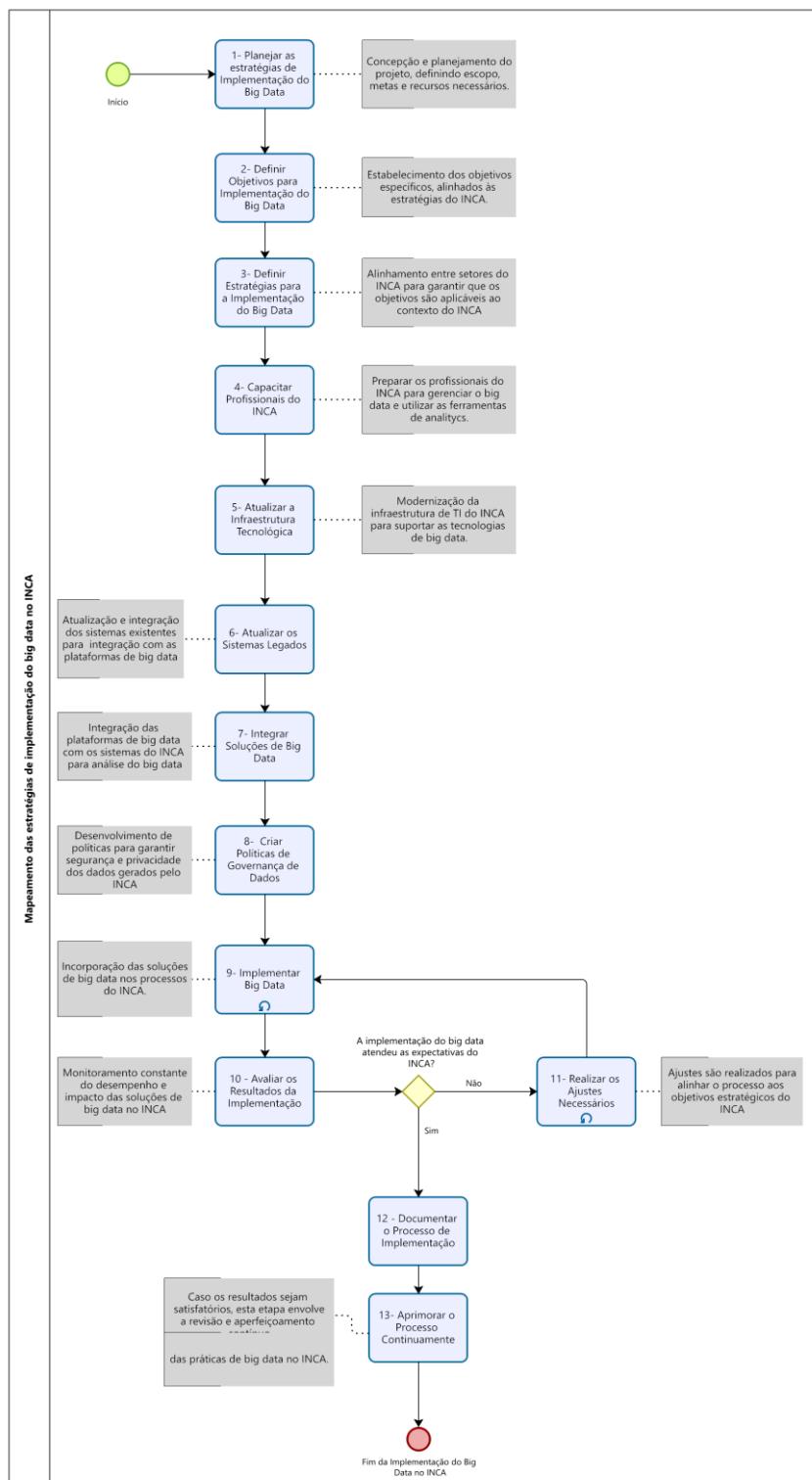
Além do desenvolvimento da infraestrutura, a estruturação das políticas de governança de dados no INCA é outro fator que, na percepção dos entrevistados, é extremamente relevante. Para isso, será necessário que o setor de governança de dados do INCA trabalhe de forma colaborativa com áreas estratégicas da instituição, de modo a garantir que as políticas definidas estejam alinhadas com as regulamentações de segurança da informação vigentes no setor de saúde. A implementação dessas políticas poderá ser acompanhada por programas de conscientização e treinamento, cujo objetivo será aumentar as chances de que os profissionais que atuam na instituição compreendam e sigam as diretrizes estabelecidas pelo Instituto de maneira consistente (CHEUNG, 2020).

Após a modernização da infraestrutura tecnológica e estruturação das políticas de governança da instituição, espera-se que gradualmente o *big data* seja integrado nas operações diárias do INCA. A implementação poderá ser acompanhada por um sistema de monitoramento e avaliação contínua, que permitirá melhorias nos processos em tempo real e a identificação de *gaps* no processo de implementação. Esse monitoramento contínuo é importante para assegurar que as metas estabelecidas pelo INCA sejam alcançadas, e que o uso do *big data* contribua para o aprimoramento dos processos

clínicos, administrativos, de pesquisa e de gestão da instituição (CHIARELLO et al, 2021).

A Figura 15 foi elaborada com o objetivo de apresentar, com base na percepção dos entrevistados, as principais fases do processo de implementação do *big data* no INCA e suas conexões.

Figura 15 - Mapeamento das estratégias de implementação do *big data* no INCA.



Fonte: Elaboração própria, com base nos dados das entrevistas (2024).

Por fim, a revisão periódica do processo de implementação do *big data* no INCA poderá permitir a melhoria contínua das práticas associadas ao uso do big data na instituição. Esta é uma etapa fundamental para que o INCA continue a evoluir e a se adaptar às novas tecnologias e demandas do setor de saúde, fortalecendo sua capacidade de impulsionar avanços no tratamento e na pesquisa do câncer no Brasil.

7. PROPOSTA DE UM PLANO DE AÇÃO PARA MITIGAR OS RISCOS ASSOCIADOS A IMPLEMENTAÇÃO DO BIG DATA NO INCA

O plano de ação apresentado foi elaborado com o objetivo de sugerir ações práticas que o INCA poderá adotar para mitigar os riscos identificados pelos profissionais da instituição durante as entrevistas, aumentando, assim, a probabilidade de uma adoção bem-sucedida dessa tecnologia no Instituto.

Esse plano de ação foi desenvolvido com base na análise dos dados coletados na pesquisa de campo realizada no Instituto Nacional de Câncer (INCA) durante o curso de Mestrado Profissional em Gestão e Estratégia (MPGE) da Universidade Federal Rural do Rio de Janeiro (UFRRJ).

Para realizar a análise de riscos no contexto específico do INCA, foi utilizada uma abordagem qualitativa, envolvendo entrevistas semiestruturadas com 11 profissionais da instituição, entre eles profissionais de saúde, gestores e especialistas em TI, com tempos de atuação que variam entre 7 e 26 anos. As entrevistas foram transcritas e analisadas por meio da técnica de análise de conteúdo, que permitiu identificar temas e padrões emergentes que destacaram os principais desafios que o INCA poderá enfrentar durante a implementação do *big data*, na perspectiva dos sujeitos entrevistados.

A frequência com que esses riscos foram mencionados pelos entrevistados serviu como indicador de prioridade, refletindo o nível de preocupação percebido pelos participantes. Com base nessa análise, foi estruturado um plano de ação que buscou apresentar estratégias para a mitigação dos riscos identificados, com o objetivo de aumentar as chances de sucesso na implementação do *big data* no INCA.

7.1 Plano de Ação: Análise dos Riscos associados a Implementação do Big Data no INCA

Com base nos dados coletados em campo, foi possível identificar que, para os gestores do INCA, a utilização do *big data* poderá oferecer uma visão integrada das informações institucionais, auxiliando na melhoria da tomada de decisões estratégicas do instituto. Paralelamente, as análises preditivas poderão ajudar a gestão a prever demandas, alocar recursos de forma mais eficiente e identificar oportunidades para melhorias contínuas nos serviços prestados pelo Instituto.

Essas práticas podem levar à diminuição dos custos operacionais e à otimização da logística hospitalar, como a gestão de estoques e o planejamento de recursos, aumentando a eficácia dos processos administrativos da instituição (CHEUNG, 2020). Além disso, o processamento de grandes volumes de dados pode facilitar a identificação de tendências epidemiológicas e padrões de evolução do câncer, permitindo ao INCA desenvolver políticas de saúde pública mais eficazes e alinhadas com

as necessidades da população.

Sob a ótica das práticas médicas, os profissionais do INCA acreditam que as ferramentas avançadas de *big data* poderão beneficiar significativamente a tomada de decisões clínicas e melhorar a qualidade do tratamento dos pacientes com câncer. Isso se deve ao fato dessas ferramentas serem percebidas como uma ferramenta de apoio para a prática médica, dado o seu potencial de aprofundar análises de diferentes tipos de dados de saúde, como históricos clínicos, informações genéticas e exames de imagem, incluindo tomografias e ressonâncias magnéticas, permitindo que o profissional de saúde realize uma avaliação mais detalhada do estado de saúde dos pacientes.

Além disso, o *big data* é capaz de integrar dados de biópsias, registros de tratamento e respostas terapêuticas para identificar padrões relacionados ao câncer e prever a progressão da doença (GUPTA et al, 2019). No caso do INCA, ter um ambiente preparado para realizar a análise aprimorada do *big data*, aumentam as chances de que os médicos prevejam complicações potenciais e riscos de evolução da doença com base nas informações individuais de cada paciente, criando planos de tratamento personalizados que considerem o perfil molecular do tumor e as características pessoais de cada indivíduo. Como consequência, o planejamento de estratégias de prevenção e controle do câncer na instituição se torna mais eficaz, com práticas clínicas que deixam de ser centradas na doença e passam a ser centradas no paciente (SENTHILKUMAR et al, 2018).

Na percepção dos profissionais de TI do INCA, a implementação do *big data* é uma oportunidade para aprimorar os conhecimentos e modernizar a infraestrutura tecnológica da instituição, uma vez que, para viabilizar a adoção desta tecnologia será necessário que o INCA adquira tecnologias robustas para o gerenciamento do *big data* e, consequentemente, capacite os seus profissionais com múltiplas habilidades, como a análise preditiva dos dados, *machine learning* e *inteligência artificial*, sendo estes fatores fundamentais para que a instituição esteja preparada para lidar com as novas demandas de *big data* (SUPRIYA, 2021).

Contudo, para que o INCA possa se beneficiar de todas as oportunidades identificadas pelos entrevistados, é importante que a instituição comprehenda que a implementação do *big data* vem acompanhada de riscos que precisam ser cuidadosamente avaliados e gerenciados para que a iniciativa seja bem-sucedida. Conforme Hillson (2016) destaca, a gestão de riscos deve ser uma parte intrínseca de todo projeto, integrando-se ao gerenciamento das etapas de implementação como um componente fundamental. Isso assegura que os riscos sejam não apenas identificados, mas também compreendidos e tratados de forma eficaz de maneira proativa e não reativa. Chapman e Ward (2019) complementam essa visão, argumentando que o objetivo não é eliminar totalmente os riscos, mas sim entendê-los para que assim, a instituição seja capaz de desenvolver estratégias eficazes de mitigação

dos principais riscos identificados.

Projetos de alta criticidade, como a implementação de *big data* em uma instituição pública de saúde plural como o INCA, requerem uma abordagem flexível e adaptável, conforme observado por Williams (2021), que afirma que, no setor de saúde, ter essa adaptabilidade é fundamental para que a instituição esteja preparada para lidar e superar os desafios inesperados que podem surgir ao longo do projeto.

Morris (2021) ressalta ainda a importância de que as ações estratégicas de análise de riscos busquem contemplar todo o ciclo de vida, desde a etapa de planejamento até a fase de implementação, homologação e melhoria contínua do processo. No entanto, Flyvbjerg (2020) alerta que o excesso de otimismo e a falta de entendimento do escopo e dos requisitos técnicos e de negócio frequentemente resultam em atrasos e estouros de orçamento em empreendimentos de alta complexidade. A implementação do *big data* no INCA não foge a essa regra, sendo fundamental que a instituição adote uma postura estratégica e assertiva para evitar cair em tais armadilhas.

Dante do exposto, comprehende-se que no caso do INCA, a identificação dos principais riscos associados à implementação do *big data*, juntamente com as estratégias para mitigá-los, fundamentadas nas percepções dos profissionais-chave da instituição, são fatores importantes para aumentar as chances de sucesso na adoção dessa tecnologia no Instituto.

Para realizar a análise de riscos no contexto específico do INCA, adotou-se uma abordagem qualitativa, utilizando entrevistas semiestruturadas com 11 profissionais da instituição. Entre os participantes, estavam profissionais de saúde, gestores e especialistas em TI, que possuem entre 7 e 26 anos de atuação no instituto. As entrevistas foram transcritas e analisadas por meio da técnica de análise de conteúdo, o que permitiu identificar temas e padrões emergentes que destacaram os principais desafios que podem ser enfrentados pelo INCA durante a implementação do *big data* no instituto. A frequência com que os riscos foram mencionados serviu como um indicador de prioridade, refletindo o nível de preocupação percebido pelos entrevistados.

7.2 Identificação dos riscos

Os principais riscos identificados pelos profissionais do INCA foram classificados em seis categorias: governança dos dados, qualidade dos dados, integração de sistemas, privacidade e segurança de dados, infraestrutura tecnológica e capacitação de profissionais. Cada uma dessas categorias representa um conjunto distinto de desafios que podem vir a comprometer a implementação bem-sucedida do *big data* no INCA, caso a instituição não adote uma estratégia de mitigação eficaz. O Quadro 20 apresenta uma descrição detalhada de cada risco identificado na etapa

de análise dos dados desta dissertação.

Quadro 20 - Riscos Identificados e suas Descrições.

Risco Identificado	Descrição
Governança dos dados	A falta de uma estrutura de governança de dados com políticas claras para a gestão de dados de saúde gerados e gerenciados pelo INCA, pode resultar em falhas de segurança e acessos não autorizados a informações sensíveis dos pacientes, infringindo a Lei Geral de Proteção de Dados (LGPD). e expondo a instituição a riscos legais e éticos.
Qualidade dos dados	A qualidade dos dados foi identificada como um risco crítico. Dados imprecisos ou incompletos podem comprometer a precisão dos diagnósticos e a confiabilidade das análises do INCA, impactando negativamente os serviços de saúde prestados à população.
Privacidade e segurança dos dados	A necessidade de adotar práticas para proteger os dados gerenciados pelo INCA contra violações de segurança e acessos não autorizados, emergiu como um tema central.
Integração dos sistemas ativos no INCA	A integração dos dados gerados pelos sistemas legados com as tecnologias de <i>big data</i> foi identificada como um desafio significativo no processo de implementação desta tecnologia no Instituto. A falta de interoperabilidade entre sistemas pode resultar em silos de dados e ineficiências operacionais.
Infraestrutura tecnológica	Na percepção dos entrevistados a atual infraestrutura tecnológica do INCA ainda é inadequada para suportar as demandas do <i>big data</i> . Essa limitação pode resultar em gargalos significativos no processamento e análise dos dados, impedindo que o INCA possa aproveitar todos os benefícios do <i>big data</i> na saúde.

Risco Identificado	Descrição
Capacitação dos profissionais para manipulação do <i>big data</i>	A falta de recursos humanos com <i>expertise</i> em tecnologias de <i>big data</i> foi percebida como um fator de risco relevante pelos entrevistados. No contexto do INCA, a implementação do <i>big data</i> requer uma equipe com habilidades técnicas e analíticas avançadas.

Fonte: Elaboração própria, com base na análise dos dados coletados no INCA (2024).

7.3 Análise de frequência e impacto potencial dos riscos identificados

Nesta etapa, foi realizada uma análise detalhada da frequência com que os riscos associados à implementação do *big data* no INCA foram mencionados pelos entrevistados, bem como do impacto potencial que cada risco poderá ter sobre a operação e os objetivos estratégicos do Instituto. Acredita-se que a combinação desses dois fatores – frequência e impacto – se apresenta como um ponto de partida capaz de ajudar o INCA na priorização dos riscos que exigem atenção imediata e na definição das estratégias de mitigação que poderão ser adotadas em cada contexto.

7.4 Frequência dos Riscos

A frequência de menção dos riscos observada na etapa de exploração do material coletado, foi utilizada como um indicativo do nível de preocupação dos profissionais do INCA em relação a cada categoria de risco. O Quadro 15 apresenta a frequência de menções dos riscos identificados:

Quadro 15 - Frequência dos Riscos Identificados

Risco Identificado	Frequência de Menção
Governança dos dados	9 menções
Privacidade e segurança de dados	8 menções
Qualidade dos dados	7 menções
Infraestrutura tecnológica	6 menções
Integração dos sistemas ativos no INCA	5 menções

Risco Identificado	Frequência de Menção
Capacitação dos profissionais para manipulação do <i>big data</i>	4 menções

Fonte: Elaboração própria, com base na análise dos dados coletados no INCA (2024).

7.5 Impacto Potencial dos Riscos

Cada risco identificado possui um impacto potencial distinto sobre o INCA, que pode variar de impactos operacionais a estratégicos. Impactos operacionais referem-se aos riscos que afetam diretamente as operações diárias e a eficiência dos processos internos da instituição. Já os impactos estratégicos estão relacionados aos riscos que têm o potencial de comprometer os objetivos de longo prazo, a conformidade com as regulamentações de proteção de dados e a reputação da instituição de saúde (SUPRYA, 2021).

No caso do INCA, na análise do impacto potencial, foram consideradas as percepções dos entrevistados sobre o grau de preocupação com que cada risco poderá afetar os objetivos estratégicos do INCA a médio e longo prazo. O Quadro 16 apresenta o impacto potencial de cada risco identificado no processo de implementação do *big data* no Instituto:

Quadro 16 - Impacto Potencial dos Riscos Identificados.

Risco Identificado	Impacto Potencial	Tipo de Impacto	Justificativa
Governança dos dados	Alto	Estratégico	Pode resultar em não conformidade regulatória, levando a multas e penalidades financeiras, além de comprometer a integridade e a segurança dos dados gerenciados pela instituição.
Privacidade e segurança dos dados	Alto	Estratégico	Incidentes de segurança podem resultar em vazamento de informações sensíveis, afetando a confiança dos pacientes e a imagem da instituição.
Qualidade dos dados	Alto	Operacional	A qualidade inadequada dos dados pode gerar erros em análises clínicas, impactando negativamente a tomada de decisões e a confiança nos resultados das análises realizadas pela instituição.

Risco Identificado	Impacto Potencial	Tipo de Impacto	Justificativa
Infraestrutura tecnológica	Médio	Operacional	A infraestrutura inadequada para lidar com o <i>big data</i> , pode limitar a capacidade do INCA de processar e analisar os dados complexos de saúde em tempo real, reduzindo a eficácia do uso de <i>big data</i> .
Integração dos sistemas ativos no INCA	Médio	Operacional	A falha na integração dos dados gerados pelos sistemas legados do INCA e as tecnologias de <i>big data</i> , pode vir a causar atrasos na etapa de implementação, afetando a eficiência operacional e a continuidade dos serviços da instituição.
Capacitação dos profissionais para manipulação do <i>big data</i>	Médio	Operacional e Estratégico	A falta de capacitação dos profissionais do INCA para lidar com as tecnologias de <i>big data</i> , pode restringir o uso efetivo dessas ferramentas, levando a uma subutilização dos recursos tecnológicos e dificultando o alcance dos objetivos estratégicos da instituição a médio e longo prazo.

Fonte: Elaboração própria, com base na análise dos dados coletados no INCA (2024).

7.6 Considerações sobre os Riscos de Baixo Impacto

Os riscos considerados de baixo impacto e frequência não foram listados nesta análise porque, embora esses riscos possam ter algum efeito sobre a instituição no futuro, considerando que este é um processo dinâmico, hoje, na percepção dos entrevistados eles não representam ameaças significativas ao sucesso do projeto e, portanto, foram priorizados em menor grau. A atenção foi focada nos riscos de maior impacto e frequência, que exigem medidas de mitigação mais imediatas e abrangentes para assegurar o sucesso da implementação.

O Quadro 17 apresenta uma síntese dos riscos considerados de baixo impacto na implementação do *big data* no INCA.

Quadro 17 - Riscos de Baixo Impacto Identificados.

Risco Identificado	Descrição	Impacto	Estratégia de Mitigação
Mudanças nas normas e regulamentações	Alterações futuras em normas regulatórias que afetem a gestão de <i>big data</i> no setor de saúde.	Baixo	Acompanhamento regular de mudanças regulatórias e adaptação proativa das políticas de conformidade.
Rotatividade de Pessoal	Saída de recursos humanos com <i>expertise</i> em <i>big data</i> durante o projeto.	Baixo	Plano de sucessão e treinamento de <i>backup</i> para garantir a continuidade do projeto.
Tempo de Implementação	Atrasos no cronograma de implementação do <i>big data</i> no INCA.	Baixo	Estabelecimento de cronogramas realistas e utilização de técnicas de gerenciamento de tempo e recursos em todas as etapas do projeto de implementação do <i>big data</i> .

Fonte: Elaboração própria, com base na análise dos dados coletados no INCA (2024).

7.7 Estratégia de resposta aos riscos identificados

Especificamente no caso do INCA, entende-se que será importante desenvolver uma estratégia abrangente de resposta aos riscos identificados ao longo deste estudo, englobando a implementação de medidas mitigatórias eficazes, para minimizar os impactos potenciais dos riscos mapeados e aumentar a probabilidade de que a implementação do *big data* apresente os resultados desejados pela instituição. Esta subseção, teve por objetivo apresentar sugestões de respostas proativas para que o INCA seja capaz de superar estes obstáculos, equilibrando inovação tecnológica com a melhoria das práticas de saúde da instituição.

Quadro 18 - Estratégias de Resposta aos Riscos Identificados para a Implementação do *Big Data* no INCA.

Risco Identificado	Estratégia de Resposta
Governança dos dados	Para mitigar esses riscos, os entrevistados sugeriram a criação de um comitê de governança de dados e a realização de auditorias

Risco Identificado	Estratégia de Resposta
	regulares para assegurar a conformidade e segurança das informações institucionais.
Qualidade dos dados	Implementar processos de validação e tratamento dos dados gerados pelo Instituto, padronizar os processos de coleta, processamento e armazenamento dos dados, bem como investir na capacitação contínua dos profissionais que irão manipular o <i>big data</i> , são ações sugeridas para mitigar esse risco no contexto do INCA.
Privacidade e segurança dos dados	Para mitigar estes riscos, os entrevistados sugerem a implementação de autenticação de dois fatores para garantir que apenas usuários autorizados tenham acesso ao ambiente tecnológico da instituição. Além disso, promover auditorias de segurança regulares e fomentar o conhecimento sobre as melhores práticas de governança dos dados, para que os profissionais de saúde se sintam seguros em manipular os registros dos pacientes, tendo como base os princípios de segurança da informação e em conformidade com as normas de proteção de dados vigentes na instituição.
Integração dos sistemas ativos no INCA	Investir em soluções que facilitem a interoperabilidade e planejar a transição dos dados dos sistemas legados de forma gradual, são estratégias recomendadas pelos profissionais do INCA para que o instituto seja capaz de superar este desafio.
Infraestrutura tecnológica	Para mitigar esse risco, será importante que o Instituto implemente estratégias para promover a modernização do ambiente tecnológico, incluindo a aquisição de tecnologias desenvolvidas especificamente para suportar a captura, tratamento, processamento, armazenamento e análise de grandes volumes de dados complexos.

Risco Identificado	Estratégia de Resposta
Capacitação dos profissionais para manipulação do <i>big data</i>	Para que o Instituto possa superar este desafio, será necessário investir em programas de capacitação contínua e montar uma equipe multidisciplinar, capaz de combinar conhecimentos técnicos e clínicos para manipulação dos dados gerenciados pela instituição.

Fonte: Elaboração própria, com base na análise dos dados coletados no INCA (2024).

Acredita-se que, ao considerar os riscos identificados e definir estratégias de resposta de maneira proativa e direcionada, o Instituto Nacional de Câncer (INCA) estará bem posicionado para enfrentar os desafios associados à implementação do *big data*. As estratégias sugeridas buscam não apenas proteger a integridade e a segurança dos dados, mas também garantir que a infraestrutura do INCA e os profissionais envolvidos estejam preparados para maximizar os benefícios dessa tecnologia nas práticas do Instituto.

No entanto, para que a implementação do *big data* se concretize e gere os resultados esperados, será importante que o INCA adote ações práticas que guiem a execução dessas iniciativas de forma estratégica, assim como apresentado no plano estratégico produto desta dissertação.

8. CONCLUSÃO

Este capítulo apresenta as etapas conclusivas do trabalho, sintetizando as principais contribuições e propondo direções para estudos futuros.

8.1 Síntese dos Resultados e Contribuições

Este estudo explorou o campo das tecnologias em saúde utilizando as bases de dados *Web of Science* e *Scopus*, e, através de técnicas de mapeamento sistemático da literatura, apresentou a diversidade dos conceitos de *big data*, destacando suas possíveis aplicações no contexto da saúde pública. Com base nos resultados obtidos, foi desenvolvido um plano estratégico para a implementação do *big data* no Instituto Nacional de Câncer (INCA), o qual pode servir de referência para outras instituições de saúde pública.

A pesquisa também ressaltou a complexidade e a natureza multidisciplinar envolvidas na implementação do *big data* em instituições de saúde pública, destacando a necessidade de integrar diversos sistemas e fontes de dados para assegurar uma base sólida de informações no setor de saúde. Essa integração é essencial para que o *big data* seja utilizado de maneira eficaz, permitindo análises mais precisas e informadas por parte das instituições, o que impulsiona melhorias na saúde pública e garante a equidade no acesso aos serviços de saúde. As principais contribuições deste trabalho incluem:

- a) **Fundamentação Teórica:** Um mapeamento sistemático que identifica os conceitos-chave e as aplicações potenciais do *big data* na saúde pública, oferecendo uma base sólida para futuras pesquisas e aplicações práticas.
- b) **Elaboração de um Plano Estratégico para o INCA:** Desenvolvimento de um plano estratégico adaptado às necessidades específicas do INCA, abordando uma visão holística para a integração do *big data* na instituição, com foco em apoiar a tomada de decisões estratégicas da instituição com base em evidências.
- c) **Contribuição Metodológica:** A metodologia utilizada combina mapeamento sistemático da literatura e análise de campo, proporcionando uma visão teórica e prática para estudos futuros que abordem o tema de implementação do *big data* em instituições de saúde pública no contexto brasileiro.
- d) **Exploração de Tecnologias Emergentes:** O impacto potencial do uso do *big data* combinado com tecnologias de ponta, como *Machine Learning*, *Deep Learning* e

Inteligência Artificial, foi explorado, destacando como essa combinação pode aprimorar a qualidade dos serviços de saúde e a eficiência dos processos clínicos e administrativos, além de otimizar a gestão de recursos, melhorar a tomada de decisões estratégicas e fortalecer o desenvolvimento de políticas públicas de saúde.

8.2 Reflexão sobre os Desafios e Oportunidades

Com base na análise dos dados coletados em campo, observou-se que implementação do *big data* no INCA poderá enfrentar desafios, como a necessidade de investimentos para a modernização da infraestrutura tecnológica, fundamental para suportar o volume e o processamento dos *big data*. Além disso, será necessário capacitar os profissionais responsáveis por gerenciar e manipular o *big data* e suas tecnologias complementares de forma contínua, garantindo que tenham as habilidades necessárias para lidar com essas tecnologias e transformar os dados em informações úteis para a instituição.

No entanto, as oportunidades também são significativas. O *big data* apresenta potencial de transformar grandes volumes de dados em *insights* valiosos, orientando políticas públicas e práticas de saúde mais eficazes. Esses benefícios são ampliados quando o *big data* é integrado a tecnologias emergentes, como *Machine Learning* e *Inteligência Artificial*, que podem automatizar processos, identificar padrões de doenças complexas como o câncer e prever tendências de evolução e propagação da doença em larga escala, impactando diretamente a qualidade dos serviços de saúde prestados à população.

Além disso, o uso do *big data* pode apoiar a prática médica, oferecendo suporte a diagnósticos precoces de doenças e fortalecendo as iniciativas de medicina preventiva. Essa tecnologia apresenta potencial para otimizar o processo de identificação de sinais iniciais de doenças, permitindo uma maior agilidade no monitoramento contínuo do estado de saúde dos pacientes, resultando em intervenções mais eficazes e oportunas.

8.3 Trabalhos Futuros

Com base nos resultados obtidos, algumas direções para estudos futuros podem incluir a expansão deste estudo para outras instituições de saúde no Brasil, permitindo a comparação entre diferentes contextos de implementação do *big data* e a análise das adaptações necessárias em cada caso. Além disso, seria relevante explorar a aplicação do *big data* em outras áreas da saúde, além da oncologia, para avaliar seu potencial em diferentes especialidades de saúde.

Adicionalmente, é necessário aprofundar a pesquisa sobre a integração do *big data* com tecnologias como *Machine Learning* e *Inteligência Artificial*, com o objetivo de fornecer exemplos

práticos de personalização de tratamentos, automação de processos diagnósticos e análise preditiva do câncer. Esses estudos poderiam ampliar a compreensão do impacto dessas tecnologias no setor de saúde, especialmente em instituições de saúde públicas brasileiras.

8.4 Considerações Finais

Com base na análise dos dados das entrevistas, foi possível identificar que, especificamente no contexto do INCA, a implementação do *big data* é percebida como um processo contínuo que requer avaliação constante e adaptação às mudanças tecnológicas e de processos organizacionais, assim como às necessidades específicas dos pacientes e dos profissionais de saúde da instituição. Como uma referência nacional em oncologia, o INCA poderá enfrentar desafios singulares, que incluem a integração de grandes volumes de dados oncológicos com outras fontes de informação em saúde, essencial para promover uma gestão integrada entre as instituições de saúde da rede pública no Brasil, garantindo assim a continuidade dos serviços prestados ao paciente.

A utilização do *big data* na rotina da instituição é percebida pelos profissionais entrevistados como uma oportunidade de oferecer aos gestores do INCA uma visão abrangente das informações institucionais, o que pode auxiliar na melhoria da tomada de decisões estratégicas. As análises preditivas, por exemplo, têm o potencial de prever demandas, otimizar a alocação de recursos e identificar oportunidades para melhorias contínuas nos serviços prestados pelo Instituto. Essas práticas podem resultar na redução de custos operacionais e na otimização da logística hospitalar, como a gestão de estoques e o planejamento de recursos, aumentando a eficácia dos processos administrativos da instituição.

Além disso, o processamento de grandes volumes de dados pode facilitar a identificação de tendências epidemiológicas e padrões na evolução do câncer, permitindo ao INCA desenvolver políticas de saúde pública mais eficazes e alinhadas com as necessidades da população. Para maximizar o potencial do *big data*, é fundamental que a instituição adote estratégias bem planejadas que mitiguem os riscos associados à implementação de novas tecnologias, aumentando a probabilidade de sucesso na utilização dos dados em operações clínicas e administrativas. Isso inclui o fortalecimento da infraestrutura tecnológica e a capacitação contínua dos profissionais envolvidos, garantindo que o uso de dados seja a base para a tomada das decisões estratégicas do Instituto.

No âmbito das práticas médicas, o *big data* é percebido pelos profissionais entrevistados como uma ferramenta com grande potencial para aprimorar o diagnóstico e otimizar o tratamento do câncer no INCA. As ferramentas avançadas de *big data* possibilitam uma análise mais aprofundada de dados de saúde, cruzando históricos clínicos, informações genéticas e exames de imagem, para fornecer

informações valiosas capazes de apoiar as decisões médicas dos profissionais de saúde. Nesse sentido, o *big data* não apenas permite uma avaliação mais precisa do estado de saúde dos pacientes, mas também, ao integrar dados de biópsias, registros de tratamento e respostas terapêuticas, coloca o INCA em uma posição vantajosa para desenvolver planos de tratamento altamente personalizados.

A implementação do *big data* também é percebida pelos profissionais de TI do INCA como uma oportunidade para fortalecer o conhecimento técnico e atualizar a infraestrutura tecnológica da instituição. Para que essa tecnologia seja efetivamente adotada, o INCA precisará investir em sistemas robustos e na capacitação de seus profissionais em áreas como análise preditiva, aprendizagem de máquina ou *Machine Learning* e inteligência artificial, uma vez que esses investimentos são fundamentais para que a instituição esteja apta para atender às novas demandas complexas que o *big data* impõe, assegurando a sua capacidade de utilizar essa tecnologia de maneira estratégica.

No entanto, para que o INCA possa aproveitar plenamente as oportunidades oferecidas pelo uso do *big data*, é essencial estar atento aos riscos associados à sua implementação e abordá-los com rigor. A avaliação e o gerenciamento cuidadosos desses riscos são fundamentais para permitir a adoção de medidas eficazes que aumentem a probabilidade de uma implementação bem-sucedida dessa tecnologia na instituição.

Por fim, o sucesso da implementação do *big data* no INCA não dependerá apenas de investimentos em tecnologia, mas também do compromisso da instituição em se adaptar continuamente aos avanços do setor e às mudanças nas demandas dos seus usuários. Com uma estratégia bem delineada e uma execução eficaz, o INCA tem o potencial de se tornar um modelo de referência na adoção de *big data* na saúde pública, contribuindo significativamente para a melhoria dos resultados em oncologia e para o avanço das políticas de saúde no Brasil.

REFERÊNCIAS

- ABRAHAM, R., SCHNEIDER, J., & VOM BROCKE, J. (2019). **Data governance: A conceptual framework, structured review, and research agenda.** International Journal of Information Management, 49, 424-438.
- ADJEKUM, A.; BLASIMME, A.; VAYENA, E. Elements of Trust in Digital Health Systems: Scoping Review. **Journal of Medical Internet Research**, v. 20, n. 12, p. e11254, 13 dez. 2018.
- ADLER-MILSTEIN, J., & JHA, A. K. (2017). HITECH Act drove large gains in hospital electronic health record adoption. **Health Affairs**, 36(8), 1416-1422.
- AGRAWAL, R., PRABAKARAN, S. *Big data* in digital healthcare: lessons learnt and recommendations for general practice. **Heredity**, v. 124, n. 4, p. 525-534, 2020.
- ALMEIDA, J. R. C. de .; PEDROSA, N. de L. .; LEITE, J. B. .; FLEMING, T. R. do P. .; CARVALHO, V. H. de .; CARDOSO, A. de A. A. . Marcadores Tumorais: Revisão de Literatura. Revista Brasileira de Cancerologia, [S. l.], v. 53, n. 3, p. 305–316, 2007. DOI: 10.32635/2176-9745.RBC.2007v53n3.1798. Disponível em: <https://rbc.inca.gov.br/index.php/revista/article/view/1798>. Acesso em: 16 jul. 2024.
- BAINBRIDGE, Michael. *Big data* challenges for clinical and precision medicine. In: *Big data, big challenges: a healthcare perspective*. Springer, Cham, 2019. p. 17-31.
- BALTAXE, E. et al. Digital Health Transformation of Integrated Care in Europe: Overarching Analysis of 17 Integrated Care Programs. **Journal of Medical Internet Research**, v. 21, n. 9, p. e14956, 26 set. 2019.
- BARDIN, Laurence. Análise de conteúdo. São Paulo: **Edições 70**, 2011.
- BATKO, K., ŚLĘZAK, A. The use of *Big data* Analytics in healthcare. **Journal of big data**, v. 9, n. 1, p. 1-24, 2022.
- BERATUNGSLEHRE, Landwirtschaftliche Kommunikations-und; HOFFMANN, Volker. The diffusion of innovations—the Hohenheim concept. **Knowledge and Innovation Management**, p. 87, 2007.
- BERTUCCI, F. et al. Santé numérique et “cancer hors les murs”, *Big data* et intelligence artificielle. **Bulletin du Cancer**, set. 2019.
- BEYER, M.A.; LANEY, D. The Importance of “*Big data*”: A Definition. **Gartner**. <https://www.gartner.com/doc/2057415>. 2012.
- BORISSOVA, Vladia. Cultural heritage digitization and related intellectual property issues. **Journal of Cultural Heritage**, v. 34, p. 145-150, 2018.
- BORGES DO NASCIMENTO, I. J.; ABDULAZEEM, H.; VASANTHAN, L. T.; MARTINEZ, E. Z.; ZUCOLOTO, M. L.; ØSTENGAARD, L.; AZZOPARDI-MUSCAT, N.; ZAPATA, T.; NOVILLO-ORTIZ, D. Barriers and facilitators to utilizing digital health technologies by healthcare professionals. **NPJ Digit Med**, v. 6, n. 1, p. 161, 18 set. 2023. DOI: 10.1038/s41746-023-00899-4. Disponível em: <https://doi.org/10.1038/s41746-023-00899-4>.
- BRASIL, Ministério da Saúde. Banco de dados do Sistema Único de Saúde-DATASUS. Disponível em <http://www.datasus.gov.br>. Acesso em 23 de maio.2022.
- BRASIL, Ministério da Saúde. Conecte SUS Cidadão. Disponível em <https://www.gov.br/saude/pt-br/assuntos/conecte-sus/cidadao>. Acesso em 20 de maio.2022.

BRASIL, Ministério da Saúde. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Gestão e Incorporação de Tecnologias em Saúde. Entendendo a Incorporação de Tecnologias em Saúde no SUS. Disponível em https://bvsms.saude.gov.br/bvs/publicacoes/entendendo_incorporacao_tecnologias_sus_envolver.pdf. 2016. Acesso em 20 de julho.2022.

BRASIL. Ministério da Saúde. Estratégia de Saúde Digital para o Brasil 2020-2028. Brasília, DF: UNASUS, 2020. Disponível em: https://bvsms.saude.gov.br/bvs/publicacoes/estrategia_saude_digital_Brasil.pdf. Acesso em: 23 de maio. 2022.

BRASIL. Ministério da Saúde. HU-UFMA participa do projeto *big data* do PROADI-SUS. 2021.

CHAPMAN, C.; WARD, S. How to manage project opportunity and risk: Why uncertainty management can be a much better approach than risk management. 3. ed. Chichester: Wiley, 2019.

CHRISTEN, P. **Data Matching: Concepts and Techniques for Record Linkage, Entity Resolution, and Duplicate Detection**. Springer, 2018.

Disponível em: <https://www.gov.br/ebserh/pt-br/hospitais-universitarios/regiao-nordeste/hu-ufma/comunicacao/noticias/hu-ufma-participa-do-projeto-big-data-do-proadi-sus>. Acesso em: 04 jan. 2023.

BRASIL. Ministério da Saúde. Política Nacional de Atenção Básica. 2012. Disponível em: <http://189.28.128.100/dab/docs/publicacoes/geral/pnab.pdf>. Acesso em: 6 abr. 2022.

BUCKERIDGE, D. L., BURKOM, H., Campbell, M., Hogan, W. R., Moore, A. W., & Pavlin, J. A. (2018). Algorithms for rapid outbreak detection: A research synthesis. **Journal of Biomedical Informatics**, 77, 120-137.

CARDONA, L. A., FRANCO OCAMPO, D. F., ESTRADA ESPONDA, R. D. Aplicaciones de la Datificación y *Big data* en América Latina entre el 2015 y 2019. **Revista Logos Ciencia & Tecnología**, v. 14, n. 2, p. 125–143, 30 jun. 2022.

CAPES. **Produção Técnica:** Grupos de Trabalho. Brasília, 2019. Disponível em: <https://www.gov.br/capes/pt-br/centrais-de-conteudo/10062019-producao-tecnica-pdf>. Acesso em: 06 de abr. de 2024.

CHEN, H.; CHIANG, R. H. L.; STOREY, V. C. Business Intelligence and Analytics: From *Big data* to Big Impact. **MIS Quarterly**, v. 36, n. 4, p. 1165, 2012.

CHEUNG, Sarah. Disambiguating the benefits and risks from public health data in the digital economy. **Big data & Society**, v. 7, n. 1, p. 2053951720933924, 2020.

CHIARELLO, Filippo, et al. Data Science for Engineering Design: State of the Art and Future Directions. **Computers in Industry**, vol. 129, 2021, p. 103447. DOI.org (Crossref), <https://doi.org/10.1016/j.compind.2021.103447>.

CHRISTENSEN, J. H, ANDREASEN, M. M., RUDEBUSCH, G. D. Term structure analysis with *big data*: one-step estimation using bond prices. **Journal of Econometrics**, v. 212, n. 1, p. 26-46, 2019.

CRESWELL, J. W. **Investigaçāo qualitativa e projeto de pesquisa: escolhendo entre cinco abordagens**. 3. ed. Porto Alegre: Penso, 2014.

D'AGOSTINO, M., MARTI, M., MEJIA, F. M., MALEK, V., e SAISO, S. G. Salud pública y la interdependencia digital: evolución tecnológica, sostenibilidad tecnológica y la revolución del usuario. **Revista Panamericana de Salud Pública**, v. 45, 2021.

- DAUGHTON, A. R., & PAUL, M. J. (2020). How data analytics can inform evidence-based policy and improve health outcomes. **Health Affairs**.
- DAVENPORT, Thomas H.; DYCHÉ, Jill. *Big data* in big companies. **International Institute for Analytics**, v. 3, n. 1-31, 2013.
- DAVENPORT, Thomas; KALAKOTA, Ravi. The potential for artificial intelligence in healthcare. **Future Healthcare Journal**. 2019. Disponível em: doi:10.7861/futurehosp.6-2-94. Acesso em: 27 mai. 2022.
- DREDZE, M., BRONIATOWSKI, D. A., & SMITH, M. C. (2019). Computational public health: Automated influenza-like illness detection and forecasting. In **Public Health Intelligence and the Internet** (pp. 197-214). CRC Press.
- FARIAS, J; ALMEIDA, J. Intervening factors in the process of disseminating information technologies in health and hospital services. **Science Review**, v. 39, p. 100318, 2021.
- FLICK, Uwe. **Introdução à Pesquisa Qualitativa** *Introdução à pesquisa qualitativa*. 3. ed. São Paulo: ARTMED, 2009.
- FLYVBJERG, B. Survival of the unfittest: Why the worst infrastructure gets built—and what we can do about it. *Oxford Review of Economic Policy*, v. 25, n. 3, p. 344-367, 2020.
- FORNAZIN, M.; RACHID, R. R.; COELHO NETO, G. C. A saúde digital nos últimos quatro anos e os desafios para o novo governo. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, [S. l.], v. 16, n. 4, p. 753–758, 2022. DOI: 10.29397/reciis.v16i4.3515. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/recciis/article/view/3515>. Acesso em: 25 jan. 2023.
- GANDOMI, A.H. Deep learning in medical imaging: fmri *big data* analysis via convolutional neural networks. In: Proceedings of the Practice and Experience on **Advanced Research Computing**. 2018. p. 1-4.
- GHASSEMI, Marzyeh e MOHAMED, Shakir. Machine learning and health need better values. **npj Digital Medicine**, v. 5, n. 1, p. 1-4, 2022.
- GIL, Antonio Carlos. **Métodos e técnicas de pesquisa social**. 6. ed. São Paulo: Atlas, 2008.
- GUPTA, D., RANI, R. A study of *big data* evolution and research challenges. **Journal of information science**, v. 45, n. 3, p. 322-340, 2019.
- H. DAVENPORT, T. How strategists use “*big data*” to support internal business decisions, discovery and production. **Strategy & Leadership**, v. 42, n. 4, p. 45–50, 15 jul. 2014.
- HILLSON, D. The risk management handbook: A practical guide to managing the multiple dimensions of risk. London: Kogan Page, 2016.
- HOSSAIN, M. M. et al. Digital interventions for people living with non-communicable diseases in India: A systematic review of intervention studies and recommendations for future research and development. **Digital Health**, v. 5, p. 205520761989615, jan. 2019.
- INCA. **Coordenação de Prevenção e Vigilância Estimativa 2014: incidência de Câncer no Brasil**. Rio de Janeiro: INCA, 2014. 124 p. Disponível em: https://www.inca.gov.br/bvscontrolecancer/publicacoes/Estimativa_2014.pdf. Acesso em: janeiro de 2023.
- INCA. **Plano Estratégico do INCA 2020-2023: Revisão Anual 2020**. Rio de Janeiro, 2020. Disponível em: <https://www.inca.gov.br/acesso-a-informacao/relatorios-gestao> Acesso em: 27 mai. 2022.

INCA. **Registros hospitalares de câncer: planejamento e gestão.** 2. ed. rev. Rio de Janeiro: INCA, 2010. Disponível em: [registros-hospitalares-de-cancer-2010.pdf \(inca.gov.br\)](https://www.inca.gov.br/acesso-a-informacao/relatorios-gestao.pdf). Acesso em janeiro de 2023.

INCA. **Relatório de Gestão 2019.** Rio de Janeiro, 2020. Disponível em: <https://www.inca.gov.br/acesso-a-informacao/relatorios-gestao> Acesso em: 27 mai. 2022.

INCA. **Relatório de Gestão 2020.** Rio de Janeiro, 2021. Disponível em: <https://www.inca.gov.br/acesso-a-informacao/relatorios-gestao> Acesso em: 25 mar. 2022.

INCA. **Relatório de Gestão 2021.** Rio de Janeiro, 2022. Disponível em: <https://www.inca.gov.br/acesso-a-informacao/relatorios-gestao> Acesso em: 30 mar. 2022.

ISO 27799:2016. **Health informatics Information security management in health using ISO/IEC 27002.** 2nd ed. Geneva: ISO, 2016.

JANETT, R. S., YERACARIS, P.P. Electronic Medical Records in the American Health System: challenges and lessons learned. **Ciencia & saude coletiva**, v. 25, p. 1293-1304, 2020.

JINDAL, A., DUA, A., KUMAR, N., DAS, A. K., VASILAKOS, A. V., e RODRIGUES, J. J. Providing healthcare-as-a-service using fuzzy rule based *big data* analytics in cloud computing. **IEEE Journal of Biomedical and Health informatics**, v. 22, n. 5, p. 1605-1618, 2018.

KAHN, L. H. (2020). The future of public health surveillance: Predictive, precision, and response analytics. In **Futureproofing Healthcare** (pp. 173-194). Springer.

KARPATHAKIS, K, LIBOW, G., POTTS, H. W., DIXON, S., GREAVES, F., & MURRAY, E. An evaluation service for digital public health interventions: user-centered design approach. **Journal of medical Internet research**, v. 23, n. 9, p. e28356, 2021.

KITCHENHAM, B. A., CHARTERS, S. **Guidelines for performing Systematic Literature Reviews in Software Engineering.** [s.l.]: Keele University and Durham University Joint Report, 2007. Disponível em: https://www.elsevier.com/_data/promis_misc/525444systematicreviewsguide.pdf

LANEY, D (2001) 3D data management: controlling data volume, velocity and variety. META Group, **Gartner**. Disponível em: <http://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>. Acesso em 18 de maio de 2022.

LEI GERAL DE PROTEÇÃO DE DADOS (2018). Disponível em: https://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/l13709.htm. Acesso em 25 de maio de 2023.

LOTFI, Chaimaa, SRINIVASAN, S., ERTZ, M., e LATROUS, I. A tool for study on impact of *Big data* technologies on firm performance. In: Intelligent Communication Technologies and Virtual Mobile Networks. **Springer**, Singapore, 2023. p. 501-515.

LUPTON, Deborah e MICHAEL, Mike. Toward a manifesto for the ‘public understanding of *big data*’. **Public Understanding of Science**, v. 25, n. 1, p. 104-116, 2016.

LUPTON, Deborah. The Commodification of Patient Opinion: The Digital Patient Experience Economy in the Age of *Big data*. **Sociology of Health & Illness**, vol. 36, no 6, julho de 2014, p. 856–69. DOI.org (Crossref), <https://doi.org/10.1111/1467-9566.12109>.

MAYER-SCHÖNBERGER; CUKIER, K. *Big data: A Revolution That Will Transform How, and Think.* Boston, **Elsevier**, 2013

MORRIS, P. W. G. Reconstructing project management. 2. ed. Oxford: Wiley-Blackwell, 2019.

NATIONAL INSTITUTES OF HEALTH OF THE UNITED STATES. The *Big data* to Knowledge (BD2K). U.S. Department of Health and Human Services. **The National Institutes of Health**, 2020. Disponível em: <https://commonfund.nih.gov/bd2k>

National Program of Cancer Registries (NPCR). **Centers for Disease Control and Prevention Division of Cancer Prevention na Control** (CDC), 2022. Disponível em: <http://www.cdc.gov/cancer/npcr/>. Acesso em: janeiro/2023.

OHLHORST, Frank J. *Big data* analytics: turning *big data* into big money. **John Wiley & Sons**, 2012.

Organização Mundial da Saúde (OMS); INTERNATIONAL TELECOMMUNICATION UNION (ITU). **National e-health strategy toolkit**. Genebra: Organização Mundial da Saúde, 2012. Disponível em: <https://apps.who.int/iris/handle/10665/75211>. Acesso em: 23 de maio. 2022

PAIGE, E., DOYLE, K., JORM, L., BANKS, E., HSU, M. P., NEDKOFF, L. e FIGTREE, G. A . A versatile *big data* health system for Australia: driving improvements in cardiovascular health. **Heart, Lung and Circulation**, v. 30, n. 10, p. 1467-1476, 2021.

PENTEADO, B. E.; FORNAZIN, M.; CASTRO, L.; FREIRE, S. L. Mapeando a dinâmica da informática médica: uma análise bibliométrica do campo científico. **Revista Eletrônica de Comunicação, Informação & Inovação em Saúde**, [S. 1.], v. 15, n. 4, 2021. DOI: 10.29397/reciis.v15i4.2395. Disponível em: <https://www.reciis.icict.fiocruz.br/index.php/reciis/article/view/2395>. Acesso em: 25 jan. 2023.

PETERSEN, Kai; VAKKALANKA, Sairam; KUZNIARZ, Ludwik. Guidelines for conducting systematic mapping studies in software engineering: An update. **Information and software technology**, v. 64, p. 1-18, 2015.

RESNYANSKY, L. Conceptual frameworks for social and cultural *Big data* analytics: Answering the epistemological challenge. **Big data & Society**, v. 6, n. 1, p. 2053951718823815, 2019.

RICARTE, Élmano. A expansão do processo de digitalização durante a pandemia de covid-19. **Research Centre for Communication and Culture (CECC)**. Faculty of Human Sciences. Finisterra, v. 55, n. 115, p. 53-60, 2020 Disponível em: DOI: <https://doi.org/10.18055/Finis20350>.

SAUNDERS, G. H., CHRISTENSEN, J. H., GUTENBERG, J., PONTOPPIDAN, N. H., SMITH, A., SPANOUDAKIS, G., e BAMIOU, D. E. Application of *Big data* to Support Evidence-Based Public Health Policy Decision-Making for Hearing. **Ear & Hearing**, v. 41, n. 5, p. 1057–1063, 23 jan. 2020.

SAYANTAN, K., DHIR, A., ISLAM, A. N., e MÄNTYMÄKI, M . *Big data* analytics in healthcare: a systematic literature review. **Enterprise Information Systems**, v. 14, n. 7, p. 878-912, 2020.

SCHLICK, Cary Jo R.; CASTLE, Joshua P.; BENTREM, David J. Utilizing *big data* in cancer care. **Surgical Oncology Clinics**, v. 27, n. 4, p. 641-652, 2018.

SEGATA, J; RIFIOTIS, T. Digitalização e dataficação da vida. **Civitas - Revista de Ciências Sociais** [online]. 2021, v. 21, n. 2 [Acessado 14 Novembro 2022] , pp. 186-192. Disponível em: <<https://doi.org/10.15448/1984-7289.2021.2.40987>>. Epub 18 Out 2021. ISSN 1984-7289.

<https://doi.org/10.15448/1984-7289.2021.2.40987>

SENTHILKUMAR, S. A., RAI, B. K., MESHRAM, A. A., GUNASEKARAN, A. e CHANDRAKUMARMANGALAM, S. *Big data* in healthcare management: a review of literature. **American Journal of Theoretical and Applied Business**, v. 4, n. 2, p. 57-69, 2018.

SILVA, Sandro Luís Freire de Castro. **A Coexistência de Lógicas Institucionais e os Sistemas**

Emergentes no Contexto de Sistemas de Informação em Saúde. 2021. 210 f. Tese (Doutorado em Informática) - Universidade Federal do Estado do Rio de Janeiro (UNIRIO), Rio de Janeiro, 2021.

SILVENNOINEN, P. Professional Master's Degree Students' Perceptions on the Changes Digitalisation Imposes on Counselling in the Social and Health Care Sector. **International Journal of Environmental Research and Public Health**, v. 17, n. 17, p. 6243, 27 ago. 2020.

SINGH, R.P., HALEEM, A., JAVAID, M., KATARIA, R., e SINGHAL, S. Cloud computing in solving problems of COVID-19 pandemic. **Journal of Industrial Integration and Management**, v. 6, n. 02, p. 209-219, 2021

SUPRIYA, M. S. e SASIDARAN, K. Machine Learning for *Big data*. In: Applications of *Big data* in Large-and Small-Scale Systems. **IGI Global**, 2021. p. 56-76.

TEIXEIRA, B. C., TOPORCOV, T. N., CHIARAVALLOTTI-NETO, F., e FILHO, A. D. P. **Spatial Clusters of Cancer Mortality in Brazil: A Machine Learning Modelling Approach**. Available at SSRN 3955672.

TEMPORÃO, José Gomes. **O Sistema de Saúde Brasileiro: Desafios e Perspectivas**. Rio de Janeiro: Editora Fiocruz, 2021.

THOMAS, David Robert. A General Inductive Approach for Analyzing Qualitative Evaluation Data. **American Journal of Evaluation**, [S. l.], v. 27, n. 2, p. 237–246, 2006. DOI: doi.org/10.1177/1098214005283748.

TOPOL, Eric J. High-Performance Medicine: The Convergence of Human and Artificial Intelligence. **Nature medicine**, vol. 25, no 1, janeiro de 2019, p. 44–56. DOI.org (Crossref), <https://doi.org/10.1038/s41591-018-0300-7>

VERGARA, Sylvia Constant. **Métodos de Pesquisa em Administração**. 6a edição ed. São Paulo: Atlas, 2015.

WANG, S., PERSHING, LEE A., ARON Y. *Big data* requirements for artificial intelligence. **Current opinion in ophthalmology**, v. 31, n. 5, p. 318, 2020.

WERUTSKY, G., BARRIOS, C. H., CARDONA, A. F., ALBERGARIA, A., VALENCIA, A., FERREIRA, C. G., ... e CAZAP, E. Perspectives on emerging technologies, personalised medicine, and clinical research for cancer control in Latin America and the Caribbean. **The Lancet Oncology**, v. 22, n. 11, p. e488-e500, 2021.

WESSON, P. et al. Risks and Opportunities to Ensure Equity in the Application of *Big data* Research in Public Health. **Annual Review of Public Health**, v. 43, n. 1, 6 dez. 2021.

WILLIAMS, T. L. Strategic risk management: New tools for competitive advantage in an uncertain age. New York: Routledge, 2017.

WONG, B. L. H. et al. The dawn of digital public health in Europe: Implications for public health policy and practice. **The Lancet Regional Health - Europe**, v. 14, p. 100316, mar. 2022.

WU, X.; LIANG, L.; CHEN, S. How *big data* alters value creation: through the lens of *big data* competency. **Management Decision**, v. ahead-of-print, n. ahead-of-print, 7 jan. 2022.

YIN, Robert K. **Estudo de caso: planejamento e métodos**. 5a edição ed. Porto Alegre: Bookman, 2015.

YUNITA, A.; SANTOSO, H. B.; HASIBUAN, Z. A. "Everything is data": towards one *big data* ecosystem using multiple sources of data on higher education in Indonesia. **Journal of Big data**, v. 9, n. 1, 14 jul. 2022. 158

XIAO, B.; RAHMANI, R.; LI, Yuhong; GILLBLAD, D.; KANTER, T. **Intelligent data-intensive IoT: A survey**. In: 2nd IEEE International Conference on Computer and Communications (ICCC), 2016, Chengdu, China. Proceedings... Chengdu: IEEE, 2016. p. 2362-2368. DOI: 10.1109/CompComm.2016.7925122.

ZHANG, Y., HUANG, T., BOMPARD, E. F. *Big data* analytics in smart grids: a review. **Energy informatics**, v. 1, n. 1, p. 1-24, 2018.

APÊNDICE A: MAPEAMENTO SISTEMÁTICO DA LITERATURA (MSL)

O MSL foi conduzido com base nas orientações de Kitchenham e Charters (2007) e Peterson et al. (2015). As diretrizes propõem a execução das seguintes etapas: 1) definição da questão de pesquisa; 2) definição da *string* de busca; 3) critérios de seleção, inclusão e exclusão; 4) seleção dos estudos. O objetivo do estudo foi definido com base na abordagem *Goal-Question-Metric* (BASILI, 1992) que consistiu em analisar estudos que abordam a evolução tecnológica no setor de saúde para observar a implementação do *big data* em instituições públicas de saúde, com o propósito de compreender o papel do *big data* no processo de tomada de decisões e na prática em saúde do ponto de vista de gestores e profissionais no contexto da saúde pública. A estratégia utilizada envolveu buscas automáticas nas bases digitais Web of Science⁴ e Scopus⁵.

Para a elaboração deste MSL foram utilizadas duas ferramentas de apoio. A primeira delas foi o Parsifal⁶, que é uma ferramenta online para apoiar pesquisadores e que incorpora os processos sugeridos por Kitchenham e Charters, 2007 e Petersen et al., 2015. A segunda foi o software Microsoft Excel, do pacote de ferramentas de escritório Microsoft Office⁷ na versão 365, que auxiliou na catalogação e classificação dos estudos.

Definições da Questão de Pesquisa

Inicialmente, foram executadas consultas *ad hoc* para uma melhor compreensão do estado da arte em diferentes campos de estudo. Essas consultas permitiram a identificação de trabalhos seminais da área e que foram utilizados nas discussões apresentadas no artigo. Em seguida, os critérios do mapeamento foram formulados seguindo os critérios do método PICOC (Population, Intervention, Comparison, Outcomes, Context), amplamente utilizado para a formulação de perguntas de pesquisa em saúde (SACKETT et al, 2000). Pelo fato do presente estudo se tratar do MSL, neste artigo foram utilizados apenas os critérios *Population* (População) e *Outcomes* (Resultados), conforme descritos no Quadro19.

⁴<https://www.webofscience.com/wos/woscc/basic-search>

⁵ <https://www.scopus.com/search/form.uri?display=basic#basic>

⁶ <https://parsif.al/>

⁷ <https://www.microsoft.com/office>

Quadro 19 - Critérios do PICOC adotados

População	<i>Big data</i> no setor de saúde pública
Resultados	Foco nos estudos que abordam a aplicabilidade e os desafios para implementação do <i>big data</i> no setor de saúde pública.

Fonte: Elaboração própria (2023).

Visando dar o devido direcionamento e alinhamento ao objetivo deste estudo, optou-se por definir a seguinte questão de pesquisa (QP: “*Quais os benefícios e desafios da aplicabilidade do big data no setor de saúde pública*”).

Definição da *String* de Busca

Inicialmente, para a elaboração da *string* de busca (SS), optou-se por executar a calibragem inicial nas bases *Web of Science* e *Scopus* em inglês, onde foi utilizada a SS (“*big data AND health*”). Porém, está iteração retornou um elevado número de resultados (17.568 artigos na *Web of Science* e 8.088 artigos na *Scopus*), em sua maioria abordando o uso do *big data* no setor de saúde privada, o que corrobora a premissa de que a perspectiva central do *big data* na saúde é advinda da iniciativa privada. Assim, optou-se por escolher uma cadeia que observasse o *big data* por meio de uma ótica específica para que assim fosse possível atender ao objetivo deste artigo.

Desta forma, optou-se por escolher uma cadeia que buscasse recuperar as tendências do uso do *big data* no setor de saúde pública, com ênfase nos estudos em tecnologia, organizações e saúde. Para isso, foi refinado um termo que melhor representasse o propósito de pesquisa, resultando na seguinte definição da SS e os operadores lógicos: (“*big data*”) e (“*public*”) e (“*health*”). Em seguida, a SS foi executada nas bibliotecas digitais. O Quadro 20 apresenta a SS de busca final.

Quadro 20 - String de busca final

String de Busca Final
(“ <i>big data</i> ”) AND (“ <i>public</i> ”) AND (“ <i>health</i> ”)

Fonte: Elaboração própria (2023).

Critérios de Seleção, Inclusão e Exclusão

Seguindo as recomendações de Kitchenham e Charters (2007) e Petersen et al. (2015), para o processo de seleção de artigos foram definidos os critérios de inclusão (CI) e exclusão

(CE) descritos no Quadro 21, com o objetivo de responder à questão de pesquisa com base em estudos relevantes e dar confiabilidade ao estudo.

Quadro 21 - Critérios de inclusão e exclusão

ID	Critério de Inclusão
CI1	O estudo aborda a aplicabilidade e os desafios para a implementação do <i>big data</i> no setor de saúde pública.
ID	Critérios de Exclusão
CE1	O artigo é uma revisão sistemática.
CE2	O artigo trata o problema de forma lateral.
CE3	O artigo não trata o uso do <i>big data</i> no setor de saúde.

Fonte: Elaboração própria (2023).

Estudos Selecionados

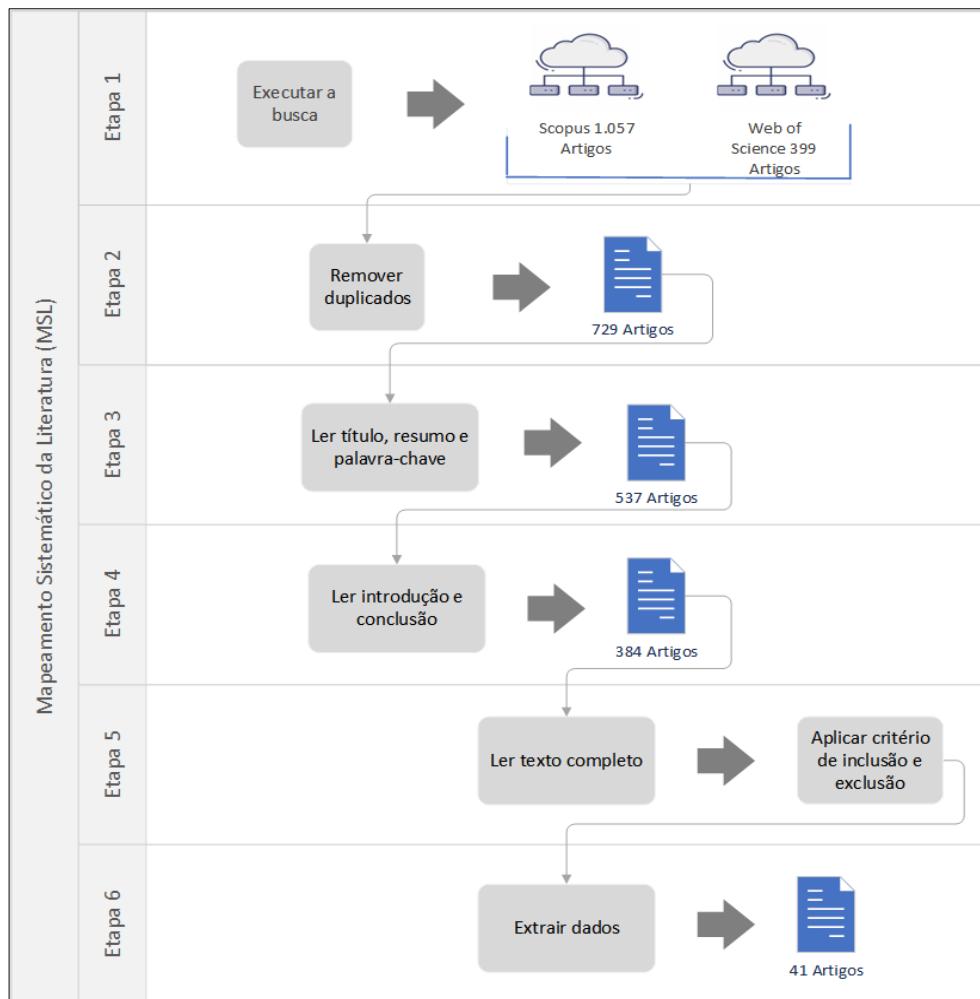
O processo de seleção dos estudos foi dividido em seis etapas: (1) execução da busca; (2) remoção dos documentos duplicados; (3) leitura de títulos, resumos e palavras-chave; (4) leitura da introdução e conclusão; (5) leitura completa; e (6) extração de dados (KITCHENHAM; CHARTERS, 2007).

Inicialmente a busca retornou 2.394 estudos, no entanto foram aplicados cinco filtros por refinamento de escopo. O primeiro no que diz respeito aos tipos de documento, onde optou-se pela seleção de artigo completos, supondo que fossem trabalhos consolidados, resultando em 2.083 artigos. A segunda filtragem se refere ao período de publicação dos artigos, onde optou-se por delimitar as buscas no recorte dos últimos 5 anos (2018 a 2022), visando recuperar o estado da arte sobre o tema de pesquisa, resultando em 1.456 artigos.

Na terceira filtragem foi realizada uma verificação manual, com o objetivo de identificar os trabalhos duplicados, resultando em 729 artigos distintos entre si. A partir da remoção dos arquivos duplicados, foi aplicada uma nova filtragem por meio da leitura do título, resumo e palavras-chave, resultando 537 em artigos.

Na quarta filtragem, uma nova seleção foi realizada por meio da leitura das introduções e conclusões, resultando em 384 artigos. Em seguida, foi realizada a leitura completa dos estudos, resultando em 185 artigos. Após a quarta filtragem, foram aplicados os critérios de inclusão e exclusão, resultando em 41 artigos finais. Estes estudos serão utilizados como base para composição das reflexões propostas nos capítulos a seguir. A Figura 15 apresenta a visão geral do processo de seleção dos artigos.

Figura 15- Visão geral do processo de seleção dos artigos.



Fonte: Elaboração própria (2023), adaptado de Silva (2021).

Os dados extraídos foram organizados em uma planilha contendo as seguintes informações: (1) Título; (2) Autor(es); (3) Fonte; e (4) Ano de publicação. O Quadro 22 apresenta os artigos selecionados que foram utilizados como base para a construção das discussões deste estudo.

Quadro 22 -Lista de artigos selecionados para a composição da Fundamentação Teórica.

ID	Título	Autor	Fonte	Ano
A1	Elements of Trust in Digital Health Systems: Scoping Review.	ADJEKUM, A.; BLASIMME, A.; VAYENA, E.	Journal of Medical Internet Research	2018
A2	<i>Big data</i> in digital healthcare: lessons learnt and recommendations for general practice.	AGRAWAL, R., PRABAKARAN, S.	Heredity	2020
A3	<i>Big data</i> challenges for	BAINBRIDGE,	Springer	2019

	clinical and precision medicine. In: Big data, big challenges: a healthcare perspective.	Michael.		
A4	Digital Health Transformation of Integrated Care in Europe: Overarching Analysis of 17 Integrated Care Programs.	BALTAXE, E.	Journal of Medical Internet Research	2019
A5	The use of <i>Big data</i> Analytics in healthcare.	BATKO, K., ŠLEZAK, A.	Journal of big Data	2022
A6	Santé numérique et “cancer hors les murs”, <i>Big data</i> et intelligence artificielle.	BERTUCCI, F.	Bulletin du Cancer	2019
A7	Cultural heritage digitization and related intellectual property issues.	BORISSOVA, Vladia.	Journal of Cultural Heritage	2018
A8	Aplicaciones de la Datificación y <i>Big data</i> en América Latina entre el 2015 y 2019.	CARDONA, L. A., FRANCO OCAMPO, D. F., ESTRADA ESPONDA, R. D.	Revista Logos Ciencia & Tecnología	2022
A9	Disambiguating the benefits and risks from public health data in the digital economy.	CHEUNG, Sarah.	Big data & Society	2020
A10	Data Science for Engineering Design: State of the Art and Future Directions.	CHIARELLO, Filippo.	Computers in Industry	2021
A11	Term structure analysis with big data: one-step estimation using bond prices.	CHRISTENSEN, J. H, ANDREASEN, M. M., RUDEBUSCH, G. D.	Journal of Econometrics	2019
A12	Salud pública y la interdependencia digital: evolución tecnológica y la sostenibilidad tecnológica.	D'AGOSTINO, M., MARTI, M., MEJIA, F. M., MALEK, V., e SAISO, S. G.	Revista Panamericana de Salud Pública	2021
A13	The potential for artificial intelligence in healthcare.	DAVENPORT, Thomas; KALAKOTA, Ravi.	Future Healthcare Journal.	2019
A14	Intervening factors in the process of disseminating information technologies in health and hospital services.	FARIAS, J; ALMEIDA, J.	Science Review	2021
A15	Deep learning in medical	GANDOMI, A.H.	Advanced Research	2018

	imaging: fmri <i>big data</i> analysis via convolutional neural networks.		Computing.	
A16	Machine learning and health need better values.	GHASSEMI, Marzyeh e MOHAMED, Shakir.	npj Digital Medicine	2022
A17	A saúde digital nos últimos quatro anos e os desafios para o novo governo.	FORNAZIN, M.; RACHID, R. R.; COELHO NETO, G. C.	Revista Eletrônica de Comunicação, Informação & Inovação em Saúde	2022
A18	A study of <i>big data</i> evolution and research challenges.	GUPTA, D., RANI, R.	Journal of information science	2019
A19	Digital interventions for people living with non-communicable diseases in India: A systematic review of intervention studies and recommendations for future research and development.	HOSSAIN, M. M.	Digital Health	2019
A21	Electronic Medical Records in the American Health System: challenges and lessons learned.	JANETT, R. S., YERACARIS, P.P.	Ciencia & saude coletiva	2020
A22	Providing healthcare-as-a-service using fuzzy rule based <i>big data</i> analytics in cloud computing.	JINDAL, A., DUA, A., KUMAR, N., DAS, A. K., VASILAKOS, A. V., e RODRIGUES, J. J.	IEEE Journal of Biomedical and Health informatics	2018
A23	An evaluation service for digital public health interventions: user-centered design approach.	KARPATHAKIS, K., LIBOW, G., POTTS, H. W., DIXON, S., GREAVES, F., & MURRAY, E.	Journal of medical Internet research	2021
A24	<i>Big data</i> analytics in healthcare: a systematic literature review.	SAYANTAN, K., DHIR, A., ISLAM, A. N., e MÄNTYMÄKI, M.	Enterprise Information Systems	2020
A25	A tool for study on impact of <i>Big data</i> technologies on firm performance.	LOTFI, Chaimaa, SRINIVASAN, S., ERTZ, M., e LATROUS, I.	Springer	2023
A26	A versatile <i>big data</i> health system for Australia: driving improvements in	PAIGE, E., DOYLE, K., JORM, L.,	Heart, Lung and Circulation,	2021

	cardiovascular health.	BANKS, E., HSU, M. P., NEDKOFF, L. e FIGTREE, G. A .		
A27	Mapeando a dinâmica da informática médica: uma análise bibliométrica do campo científico.	PENTEADO, B. E.; FORNAZIN, M.; CASTRO, L.; FREIRE, S. L.	Revista Eletrônica de Comunicação, Informação & Inovação em Saúde	2021
A28	Conceptual frameworks for social and cultural <i>Big data</i> analytics: Answering the epistemological challenge.	RESNYANSKY, L.	Big data & Society	2019
A29	A expansão do processo de digitalização durante a pandemia de covid-19.	RICARTE, Élmano.	Research Centre for Communication and Culture (CECC).	2020
A30	Application of <i>Big data</i> to Support Evidence-Based Public Health Policy Decision-Making for Hearing.	SAUNDERS, G. H., CHRISTENSEN, J. H., GUTENBERG, J., PONTOPPIDAN, N. H., SMITH, A., SPANOUDAKIS, G., e BAMIOU, D. E.	Ear & Hearing	2020
A31	Utilizing <i>big data</i> in cancer care.	SCHLICK, Cary Jo R.; CASTLE, Joshua P.; BENTREM, David J.	Surgical Oncology Clinics	2018
A32	Digitalização e dataficação da vida.	SEGATA, J; RIFIOTIS, T.	Civitas - Revista de Ciências Sociais	2021
A33	<i>Big data</i> in healthcare management: a review of literature.	SENTHILKUMAR , S. A., RAI, B. K., MESHRAM, A. A., GUNASEKARAN, A. e CHANDRAKUMA RMANGALAM, S.	American Journal of Theoretical and Applied Business	2018
A34	Professional Master's Degree Students' Perceptions on the Changes Digitalisation Imposes on Counselling in the Social and Health Care Sector.	SILVENNOINEN, P.	International Journal of Environmental Research and Public Health,	2020
A35	Cloud computing in solving problems of COVID-19 pandemic.	SINGH, R.P., HALEEM, A., JAVAID, M.,	Journal of Industrial Integration and	2021

		KATARIA, R., e SINGHAL, S.	Management,	
A36	Machine Learning for Big Data. In: Applications of <i>Big data</i> in Large-and Small-Scale Systems.	SUPRIYA, M. S. e SASIDARAN, K.	IGI Global	2021
A37	A General Inductive Approach for Analyzing Qualitative Evaluation Data.	THOMAS, David Robert.	American Journal of Evaluation	2021
A38	High-Performance Medicine: The Convergence of Human and Artificial Intelligence.	TOPOL, Eric J.	Nature medicine	2019
A39	Perspectives on emerging technologies, personalised medicine, and clinical research for cancer control in Latin America and the Caribbean.	WERUTSKY, G., BARRIOS, C. H., CARDONA, A. F., ALBERGARIA, A., VALENCIA, A., FERREIRA, C. G., CAZAP, E.	The Lancet Oncology	2021
A40	Risks and Opportunities to Ensure Equity in the Application of <i>Big data</i> Research in Public Health.	WESSON, P.	Annual Review of Public Health	2021
A41	The <i>big data</i> in public health in Europe: Implications for public health policy and practice.	YUNITA, A.; SANTOSO, H. B.; HASIBUAN, Z. A.	Journal of Big Data	2022

Fonte: Elaboração própria (2023).

APÊNDICE B: ROTEIRO DE ENTREVISTA SEMIESTRUTURADO - PROFISSIONAIS DE SAÚDE DO INCA



**UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO
INSTITUTO DE CIÊNCIAS SOCIAIS APLICADAS
PROGRAMA DE PÓS-GRADUAÇÃO EM GESTÃO E ESTRATÉGIA**

Documento Diretivo - Projeto de Pesquisa

Orientador: Prof. Dr. Saulo Barbará

Coorientador: Prof. Dr. Sandro Freire

Aluna: Luana da Silva Alves

Título da pesquisa: Plano de recomendações estratégicas para a adoção do *big data* na saúde pública: um estudo aplicado no Instituto Nacional de Câncer.

1. PROBLEMA DA PESQUISA:

Tema: Implementação do *big data* no setor de saúde pública brasileiro.

Problema: Quais instrumentos de gestão podem ser adotados pelo Instituto Nacional de Câncer (INCA) para implementação do *big data* no Instituto?

2. OBJETIVOS:

Objetivo Geral: Elaboração de um relatório técnico com recomendações estratégicas para implementação do *big data* no Instituto Nacional de Câncer (INCA).

Objetivos Específicos:

- Descrever, com base na literatura, a relação entre tecnologia e o processo de evolução tecnológico na área da saúde;
- Identificar os impactos no setor de saúde ocasionados pelo aumento exponencial do volume dos dados dos pacientes e como o uso estratégico do *big data* tem sido aplicado no setor;

- Apresentar um relatório técnico com estratégias que viabilizem a implementação do *big data* no INCA, por meio de um estudo aplicado; e
- Validar as recomendações propostas no relatório técnico.

ROTEIRO PARA ENTREVISTAS SEMIESTRUTURADA - PROFISSIONAIS DE SAÚDE DO INCA

Cenário: No ano de 2020 o Setor de Tecnologia da Informação (SETI), em conjunto com os gestores das áreas de planejamento estratégico do INCA, apresentaram uma iniciativa para a implementação da área de Ciência de Dados e Gestão do Conhecimento no instituto (INCA,2021).

As áreas envolvidas no projeto, justificaram a relevância do projeto ao afirmar que o INCA é produtor de grandes volumes de dados de saúde, estruturados e não estruturados, originados de diferentes fontes, e que a criação de uma área para análise descritiva e preditiva desta base de dados de saúde, poderá propiciar a geração de conhecimentos relevantes para a área da oncologia.

Estas informações, poderão ser utilizadas para apoiar à tomada de decisões estratégicas da instituição, tendo em vista a necessidade existente de mapear, armazenar e compartilhar as experiências e conhecimento produzidos pelo INCA com outras instituições de saúde do país (INCA, 2021).

Paralelamente, tomando como base os estudos analisados ao longo desta pesquisa, observou-se que as instituições de saúde orientadas por dados aplicaram terapias mais assertivas, controlaram seus ativos com maior eficiência e beneficiaram os pacientes com tratamentos personalizados, fatores que melhoraram a qualidade do tratamento e aumentaram as chances de sobrevivência dos pacientes com câncer.

Desta forma, pressupõem-se que, por se tratar de uma instituição pública de saúde, o INCA não está alheio a esta realidade. Neste sentido, tecnologias como o *big data* podem se apresentar como uma ferramenta estratégica de valor para as instituições de saúde que buscam alcançar seus objetivos estratégicos a médio e longo prazo.

Fato-chave: Avaliar quais instrumentos de gestão podem ser adotados pelo Instituto Nacional de Câncer (INCA) para implementação do *big data* no Instituto?

PROPOSTA DE ROTEIRO:

1. Qual atividade você desempenha no INCA e qual a sua experiência na área da saúde?
2. De maneira geral, como o *big data* pode ser usado para melhorar a qualidade dos cuidados de saúde no INCA?
3. Na sua percepção, quais são os principais desafios e oportunidades para a implementação do *big data* no INCA?
4. De maneira geral, como você avalia os impactos que a adoção do *big data* no INCA teria para o desenvolvimento do seu trabalho?
5. Na sua percepção, existe uma padronização para registrar eletronicamente os dados dos pacientes no INCA?
6. Poderia citar as fontes de dados dos pacientes disponíveis no INCA?
7. Na sua percepção, o INCA possuí os requisitos tecnológicos necessários para implementar com sucesso o uso do *big data* no instituto? Em caso de resposta negativa, poderia por favor citar os motivos que te fizeram chegar a essa conclusão?
8. De maneira geral, você considera que a implementação do *big data* no instituto traria desafios éticos e legais quanto a coleta de dados dos pacientes? Caso a resposta seja positiva, poderia apontar os principais desafios e como eles poderiam ser superados?
9. Tendo em vista que o INCA é uma instituição que atua no desenvolvimento e coordenação de ações integradas para a prevenção e controle do câncer no Brasil, na sua percepção, quais seriam os impactos que a implementação do *big data* traria para o Instituto?
10. De maneira geral, você considera que a adoção do *big data* no INCA poderia auxiliar na tomada de decisões estratégicas da instituição? Caso a resposta seja positiva, poderia por favor, explicar os motivos que te fizeram chegar a essa conclusão?
11. De maneira geral, você considera que a adoção do *big data* no INCA poderia auxiliar na tomada de decisões clínicas dos profissionais de saúde? Caso a resposta seja positiva, poderia por favor, explicar os motivos que te fizeram chegar a essa conclusão?

ANEXO A: TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

<p>MINISTÉRIO DA SAÚDE</p>  <p>INSTITUTO NACIONAL DE CÂNCER</p>	<p>www.inca.gov.br</p>
<p>TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO</p>	
<p>RECOMENDAÇÕES ESTRATÉGICAS PARA A ADOÇÃO DO BIG DATA NA SAÚDE PÚBLICA: UM ESTUDO DE CASO NO INSTITUTO NACIONAL DE CÂNCER</p>	
<p>Prezado (a) Senhor (a),</p> <p>Meu nome é Luana da Silva Alves e estou realizando a pesquisa acadêmica aplicada sobre o tema Implementação do <i>Big Data</i> em uma Intituição de Saúde Pública. Esta pesquisa compõe a minha dissertação de mestrado realizada no MPGE/UFRRJ, sob orientação do Prof. Dr. Saulo Barbará de Oliveira e sobre a coorientação do Prof. Dr. Sandro Luis Freire de Castro Silva. As informações a seguir destinam-se a convidá-lo (a) a participar voluntariamente deste projeto na condição de fonte, ou seja, o sujeito que fornece as informações primárias para a pesquisa em curso.</p> <p>A presente proposta de pesquisa tem como objetivo entender como você percebe a implementação de tecnologias como o Big Data no INCA. Para tal, serão feitas perguntas sobre o seu dia a dia, sobre a sua opinião quanto ao uso de tecnologias como o <i>Big Data</i> no setor de saúde e se o seu trabalho seria impactado pelo uso do <i>Big Data</i> no Instituto.</p> <p>Com as informações obtidas, acredita-se que será possível realizar uma generalização que permitirá entender melhor os impactos do uso do <i>Big Data</i> no setor de saúde, em especial na oncologia. O estudo pretende apresentar como o <i>Big Data</i> pode ser utilizado como uma ferramenta estratégica importante no setor de saúde, uma vez que permite coletar, armazenar e analisar grandes volumes de dados médicos em tempo real.</p> <p>A sua participação envolve os seguintes riscos previsíveis: Os riscos desta investigação, embora sejam mínimos, podem envolver situações de constrangimento ou desconforto ao apontar erros ou equívocos cometidos por pessoas conhecidas, bem como questões relacionadas à privacidade e confidencialidade. Com a intenção de mitigar esses riscos, serão adotadas algumas medidas de prevenção, conforme descrito a seguir:</p> <p>Para garantir a privacidade e confidencialidade dos participantes, será assegurado o sigilo dos</p>	
<p>Rubrica do participante ou representante legal</p>	<p>Rubrica do investigador responsável</p>
<p>Página 1 de 3</p>	



dados pessoais coletados. Todas as informações obtidas serão tratadas de forma estritamente confidencial e utilizadas exclusivamente para fins acadêmicos e científicos. Nomes ou quaisquer informações que possam identificar individualmente os participantes serão mantidos em sigilo e apenas resultados agregados e anonimizados serão divulgados. Ademais, o participante tem a liberdade e o direito de interromper a entrevista a qualquer momento e a mesma poderá ser remarcada para momento mais oportuno, caso assim o entrevistado deseje.

Sua participação possui os seguintes benefícios: Esta pesquisa contribuirá para identificar quais os caminhos que podem ser adotados pelo INCA, uma instituição pública de saúde especializada no tratamento de doenças complexas, para adoção e implementação do *Big Data* na instituição. Ademais, contribuirá para o conhecimento acadêmico acerca de mapear as potencialidades do uso de tecnologias como o *Big Data* no setor de saúde e os caminhos que podem ser adotados por uma intituição de saúde pública brasileira para implementar esta tecnologia no setor.

Você está sendo consultado sobre seu interesse e disponibilidade de participar desta pesquisa. Você é livre para recusar-se a participar, retirar seu consentimento ou interromper sua participação a qualquer momento. A recusa em participar não acarretará penalidade alguma.

Caso você desista de participar da pesquisa, você poderá solicitar a exclusão dos dados coletados a qualquer momento. Nesse caso, você deverá enviar e-mail para tec.luanaalves@gmail.com, solicitando a exclusão dos seus dados coletados. Caso você desista, eu, na condição de pesquisador responsável, assumo o compromisso de retirar seu consentimento, assim como excluir seus dados da pesquisa e do banco de dados.

Você não será remunerado por ser participante da pesquisa. Se houver gastos com transporte ou alimentação, eles serão resarcidos pelo pesquisador responsável. Todas as informações obtidas por meio de sua participação serão de uso exclusivo para esta pesquisa e ficarão sob a guarda do/da pesquisador/a responsável. Caso a pesquisa resulte em dano pessoal, o resarcimento e indenizações previstos em lei poderão ser requeridos pelo participante. Você poderá ter acesso aos resultados da pesquisa quando ela terminar, caso tenha interesse, solicitando diretamente à pesquisadora através do e-mail informado abaixo.

Se você tiver qualquer dúvida em relação à pesquisa, você pode entrar em contato com o pesquisador através do(s) telefone(s) 21 969536179 e pelo e-mail tec.luanaalves@gmail.com.

Rubrica do participante ou
representante legal

Rubrica do investigador
responsável

Página 2 de 3



Este estudo foi analisado e aprovado por um Comitê de Ética em Pesquisa (CEP) sob o registro CAAE _____. O CEP é responsável pela avaliação e acompanhamento dos aspectos éticos de pesquisas envolvendo seres humanos, visando garantir o bem-estar, a dignidade, os direitos e a segurança de participantes de pesquisa; bem como assegurando a participação do(a) pesquisador(a) sob os mesmos aspectos éticos.

Caso você tenha dúvidas e/ou perguntas sobre seus direitos como participante deste estudo, entre em contato com o Comitê de Ética em Pesquisa (CEP) da Universidade Federal Rural do Rio de Janeiro, situada na BR 465, km 7, Seropédica, Rio de Janeiro, pelo telefone (21) 2681-4749 de segunda a sexta, das 09:00 às 16:00h, pelo e-mail: eticacep@ufrj.br ou pessoalmente às terças e quintas das 09:00 às 16:00h.

No caso de aceitar participar da pesquisa, você e o pesquisador devem rubricar todas as páginas e também assinar as duas vias deste documento. Uma via é sua e a outra via ficará com o(a) pesquisador(a).

Para mais informações sobre os direitos dos participantes de pesquisa, leia a **Cartilha dos Direitos dos Participantes de Pesquisa** elaborada pela Comissão Nacional de Ética em Pesquisa (Conep), disponível no site:

http://conselho.saude.gov.br/images/comissoes/conep/img/boletins/Cartilha_Direitos_Participantes_de_Pesquisa_2020.pdf

Rubrica do participante ou
representante legal

Rubrica do investigador
responsável

Página 3 de 3

ANEXO B: TERMO DE ANUÊNCIA PARA AUTORIZAÇÃO DA PESQUISA NO INCA



Ministério da Saúde
 Secretaria de Atenção Especializada à Saúde
 Instituto Nacional de Câncer
 Gabinete
 Serviço de Tecnologia da Informação

TERMO

TERMO DE ANUÊNCIA

O Serviço de Tecnologia da Informação do Instituto Nacional de Câncer - INCA, CNPJ 00.394.544/0171-50, com sede em Praça Cruz Vermelha, 23 – Centro, 20230-130 da cidade do Rio de Janeiro - RJ, representado pelo Sr. Roberto Luiz Silva dos Santos abaixo assinado, na qualidade de Chefe do Serviço de Tecnologia da Informação, vem por meio desta confirmar, para os devidos fins junto a Universidade Federal Rural do Rio de Janeiro - UFRRJ, a firme intenção de participar do projeto de pesquisa Recomendações estratégicas para a adoção do *big data* na saúde pública: um estudo de caso no Instituto Nacional de Câncer, coordenado pelo Dr. Sandro Luís Freire de Castro Silva e que terá como pesquisador responsável nessa instituição o Dr. Saulo Barbará de Oliveira.

Declaramos conhecer e cumprir as Resoluções Éticas Brasileiras, em especial a Resolução 466/2012 do CNS. Estamos cientes das responsabilidades como instituição participante deste projeto de pesquisa, e do compromisso no resguardo da segurança e o bem-estar dos participantes de pesquisa nela recrutados, dispondo de infraestrutura necessária para a garantia de tal segurança e bem-estar.

Rio de Janeiro, 19 de julho de 2023.

Roberto Luiz Silva dos Santos

Chefe do Serviço de Tecnologia da Informação

Portaria nº 553 de 17 de Março de 2022, publicada no D.O.U. nº 53 de 18/03/2022.

Matrícula SIAPE nº 2262649

Instituto Nacional de Câncer

Documento assinado eletronicamente por **Roberto Luiz Silva dos Santos, Chefe do Serviço de Tecnologia da Informação**, em 19/07/2023, às 15:11, conforme horário oficial de Brasília, com fundamento no § 3º, do art. 4º, do [Decreto nº 10.543, de 13 de novembro de 2020](#); e art. 8º, da [Portaria nº 900 de 31 de Março de 2017](#).

 A autenticidade deste documento pode ser conferida no site http://sei.saude.gov.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0034861227** e o código CRC **B189D8F7**.

Referência: Processo nº 25410.010627/2023-72 **SEI nº** 0034861227

Serviço de Tecnologia da Informação - SETI/INCA
Rua do Rezende, nº 195 - Bairro Centro, Rio de Janeiro/RJ, CEP 20230-091
[Site](#)

ANEXO C: PARECER CONSUBSTANIADO DO COMITÊ DE ÉTICA EM PESQUISA



UNIVERSIDADE FEDERAL
RURAL DO RIO DE JANEIRO
(UFRRJ)

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: RECOMENDAÇÕES ESTRATÉGICAS PARA A ADOÇÃO DO BIG DATA NA SAÚDE PÚBLICA: UM ESTUDO DE CASO NO INSTITUTO NACIONAL DE CÂNCER

Pesquisador: LUANA DA SILVA ALVES

Área Temática:

Versão: 3

CAAE: 71597923.3.0000.0311

Instituição Proponente: PPGE - Programa de Pós-graduação em Gestão e Estratégia

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 6.528.135

Apresentação do Projeto:

Trata-se de um projeto de pesquisa que busca a implementação do "Big Data" como uma ferramenta para análise dos dados obtidos de pacientes com câncer atendidos no INCA. "A pesquisa mostra-se oportuna dada a potencialidade do Big Data quando aplicado a saúde. A iniciativa tem como base o Planejamento Estratégico elaborado pela gestão do INCA, para o período de 2020-2023, em conformidade com o Plano Nacional de Saúde (PNS), tomado como ponto de partida para a elaboração estratégica do Instituto".

Esta pesquisa será realizada em duas fases: uma de exploração, que consiste na construção da parte teórica deste estudo e outra de imersão no campo. A metodologia adotada na fase de pesquisa de campo será a do estudo de casos, no qual, para a seleção dos sujeitos serão utilizados os critérios de acessibilidade e a técnica de "snowball", que permite a seleção de novos respondentes por meio das redes de referência dos participantes. A coleta de dados será feita por meio do levantamento bibliográfico, para uma melhor compreensão do corpus em estudo, bem como da análise documental da Instituição. Como instrumento de coleta de dados foi elaborado, pela própria pesquisadora, um roteiro de entrevista semiestruturado. O roteiro da entrevista foi criado com base na Fundamentação Teórica, com o objetivo de compreender a percepção dos entrevistados sobre os seguintes temas: (1) a relação do processo de desenvolvimento tecnológico e a evolução dos conceitos de saúde; (2) a utilização dos dados como

Endereço: BR 465, KM 7, Zona Rural, Biblioteca Central, 2º andar
Bairro: ZONA RURAL **CEP:** 23.897-000
UF: RJ **Município:** SEROPEDICA
Telefone: (21)2681-4749 **E-mail:** e

E-mail: eticacep@ufrri.br

**UNIVERSIDADE FEDERAL
RURAL DO RIO DE JANEIRO
(UFRRJ)**



Continuação do Parecer: 6.528.135

ferramentas de apoio para as decisões estratégicas das instituições de saúde; (3) a relevância dos dados para a formulação de políticas públicas na área da saúde; e (4) a relação do desenvolvimento tecnológico com a qualidade dos serviços de saúde prestados à população. A análise de conteúdo se desenvolverá em 3 etapas: (1) pré-análise, (2) exploração do material e codificação e, por fim, (3) categorização, tratamento dos resultados, inferência e interpretação.

Desfecho primário:

A partir dos dados encontrados, espera-se construir um produto técnico que permita identificar quais instrumentos de gestão podem ser adotados pelo INCA para a implementação do Big Data no instituto.

Critérios de inclusão e exclusão:

Para a primeira fase da pesquisa o critério de inclusão inclui estudos que abordam o uso do Big Data no setor de saúde pública. Serão excluídas investigações de utilizam método de revisão da literatura e que abordam o problema de forma lateral.

Para a segunda fase da pesquisa, serão incluídos os profissionais entrevistados devem atender os seguintes requisitos: (1) ser gestor de áreas voltadas para o planejamento estratégico do INCA; (2) ser profissional atuante da área da saúde; (3) ser profissional da área de tecnologia da informação. Serão excluídas entrevistas omissas, e limitações para os participantes, como a influência das relações trabalhistas em suas respostas e a inteligibilidade das perguntas realizadas e das respostas dadas. As análises indutivas e interpretativas têm como limitação a interpretação inadequada das percepções dos participantes obtidas pelo pesquisador, o que pode ser mitigados com a categorização dos dados empíricos.

Objetivo da Pesquisa:

A proponente descreve como objetivos:

Objetivo geral/principal:

Propor recomendações de melhorias com base nos recursos do Big Data aplicáveis à tomada de decisão estratégica no Instituto Nacional de Câncer (INCA).

Objetivos específicos/secundários:

Endereço: BR 465, KM 7, Zona Rural, Biblioteca Central, 2ºandar	CEP: 23.897-000
Bairro: ZONA RURAL	
UF: RJ	Município: SEROPEDICA
Telefone: (21)2681-4749	E-mail: elicacep@ufrj.br

Página 02 de 05

**UNIVERSIDADE FEDERAL
RURAL DO RIO DE JANEIRO
(UFRRJ)**



Continuação do Parecer: 6.528.135

1. Descrever, com base na literatura, a relação entre tecnologia e o processo de evolução tecnológico na área da saúde;
2. Identificar os impactos no setor de saúde ocasionados pelo aumento do volume exponencial de dados e a contribuição do Big Data para melhoria da gestão;
3. Apresentar um plano de ação com estratégias que viabilizem a implementação do Big Data no INCA, por meio de um estudo de caso; e
4. Validar as recomendações propostas no plano.

Avaliação dos Riscos e Benefícios:

Segundo a pesquisadora, os benefícios envolvem a construção de um produto técnico que permita identificar quais instrumentos de gestão podem ser adotados pelo INCA para a implementação do Big Data no Instituto. A implementação do Big Data possibilita que os profissionais do setor tomem decisões estratégicas eficazes para o enfrentamento e monitoramento da incidência de câncer, assim como gerar conhecimentos significativos para os profissionais e pesquisadores da área.

Riscos para os participantes da pesquisa e propostas para mitigá-los ainda não foram apresentados pela pesquisadora, embora a pesquisadora reconheça que os sujeitos para participarem da pesquisa devam reconhecê-los.

Comentários e Considerações sobre a Pesquisa:

Trata-se de uma terceira avaliação de um protocolo original de pesquisa, que tem como equipe de pesquisa: SAULO BARBARA DE OLIVEIRA e SANDRO LUIS FREIRE DE CASTRO SILVA.

Considerações sobre os Termos de apresentação obrigatória:

Todos os documentos apresentados no protocolo de pesquisa pelo(a) proponente não possuem pendência, segundo as normas vigentes.

Recomendações:

Recomenda-se que o pesquisador acompanhe a tramitação do projeto de pesquisa na Plataforma Brasil com regularidade, atentando-se às diferentes fases do processo e seus prazos:

- a) O pesquisador deverá submeter relatórios parciais a cada semestre;

Endereço: BR 465, KM 7, Zona Rural, Biblioteca Central, 2º andar	CEP: 23.897-000
Bairro: ZONA RURAL	
UF: RJ	Município: SEROPEDICA
Telefone: (21)2681-4749	E-mail: eticacep@ufrj.br

Página 03 de 05

**UNIVERSIDADE FEDERAL
RURAL DO RIO DE JANEIRO
(UFRRJ)**



Continuação do Parecer: 6.528.135

- b) quando da necessidade de emendas ou notificações no projeto, consultar a Norma Operacional 001/2013 - Procedimentos para Submissão e Tramitação de Projetos.
- c) quando da finalização do projeto, submeter relatório final.

Conclusões ou Pendências e Lista de Inadequações:

A adequação às RESOLUÇÕES Nº 466 de 12 de dezembro de 2012 e Nº 510 de 24 de maio de 2016, , foi plenamente atendida pelo(a) pesquisador(a).

Considerações Finais a critério do CEP:

O pesquisador apresentou dois Termos de Consentimento Livre e Esclarecido, cumprindo com as exigências do capítulo IV da Resolução CNS nº 466/2012. Portanto, o TCLE que deverá ser aplicado aos participantes da pesquisa será o TCLE INCA.

Orientamos e exclusão do TCLE/UFRRJ pela inadequação as normativas do sistema CEP/CONEP.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJECTO_2085947.pdf	02/11/2023 09:02:32		Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_UFRRJ.pdf	02/11/2023 09:00:04	LUANA DA SILVA ALVES	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLE_INCA.pdf	02/11/2023 08:59:54	LUANA DA SILVA ALVES	Aceito
Declaração de concordância	AssinaturaSEITermodeAnuenciaLuanaAlves.pdf	20/07/2023 12:28:50	LUANA DA SILVA ALVES	Aceito
Cronograma	Cronograma.pdf	20/07/2023 12:27:21	LUANA DA SILVA ALVES	Aceito

Endereço: BR 465, KM 7, Zona Rural, Biblioteca Central, 2º andar

Bairro: ZONA RURAL

CEP: 23.897-000

UF: RJ

Município: SEROPEDICA

Telefone: (21)2681-4749

E-mail: eticacep@ufrj.br

**UNIVERSIDADE FEDERAL
RURAL DO RIO DE JANEIRO
(UFRRJ)**



Continuação do Parecer: 6.528.135

Projeto Detalhado / Brochura Investigador	Projeto_Mestrado_LuanaAlves.docx	22/05/2023 10:00:37	LUANA DA SILVA ALVES	Aceito
Folha de Rosto	FolhadeRosto_PlataformaBrasil_LuanaA lves.pdf	22/05/2023 09:59:01	LUANA DA SILVA ALVES	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

SEROPEDICA, 23 de Novembro de 2023

Assinado por:
Valeria Nascimento Lebeis Pires
(Coordenador(a))

Endereço: BR 465, KM 7, Zona Rural, Biblioteca Central, 2ºandar	CEP: 23.897-000
Bairro: ZONA RURAL	Município: SEROPEDICA
UF: RJ	E-mail: eticacep@ufrj.br
Telefone: (21)2681-4749	

Página 05 de 05