

# EPIDEMIA DA ZIKA E SEUS DESDOBRAMENTOS

*Uma abordagem multidisciplinar e  
integrativa dos seus efeitos no Brasil*



MAURICIO LIMA BARRETO  
WAGNER DE JESUS MARTINS  
BETHÂNIA DE ARAÚJO ALMEIDA  
MARIA GLÓRIA TEIXEIRA  
MARIA YURY TRAVASSOS ICHIHARA  
RICARDO BARROS SAMPAIO

**Organizadores**

**CENTRO DE INTEGRAÇÃO DE DADOS E CONHECIMENTOS PARA SAÚDE  
CIDACS/FIOCRUZ BAHIA**

**Coordenação**

Mauricio Lima Barreto  
Maria Yury Travassos Ichihara

**Diretora do Instituto Gonçalo Moniz - Fiocruz Bahia**

Marilda de Souza Gonçalves

**Diretora da Gerência Regional de Brasília - Gereb/Fiocruz**

Maria Fabiana Damasio Passos

**Presidente da Fiocruz**

Mario Santos Moreira

**Produção Técnica e Revisão**

Alayde Lopes Sarno Carvalho  
Maria Jenny Silva Araújo  
Mariana Rodrigues Sebastião de Almeida

**Projeto Gráfico**

Gilson Rabelo de Almeida Neto

**Dados Internacionais de Catalogação na Publicação (CIP) (Câmara Brasileira do Livro, SP, Brasil)**

Epidemia da zika e seus desdobramentos [livro eletrônico] : uma abordagem multidisciplinar e integrativa dos seus efeitos no Brasil / organização Mauricio Lima Barreto...[et al.].

-- 1. ed. -- Salvador, BA : Ed. dos Autores, 2022.

PDF.

Outros organizadores: Wagner de Jesus Martins, Bethânia de Araújo Almeida, Maria da Glória Lima Cruz Teixeira, Maria Yury Travassos Ichihara, Ricardo Barros Sampaio.

Bibliografia.

ISBN 978-65-00-57442-5

1. Saúde pública 2. Síndrome Congênita do Zika vírus (SCZv) 3. Zika - Prevenção I. Barreto, Mauricio Lima. II. Martins, Wagner de Jesus. III. Almeida, Bethânia de Araújo. IV. Teixeira, Maria da Glória Lima Cruz. V. Ichihara, Maria Yury Travassos. VI. Sampaio, Ricardo Barros.

22-136867

CDD-362.1

**Índices para catálogo sistemático:**

1. Crianças com Síndrome Congênita do Zika vírus: Saúde pública

362.1

Aline Grazielle Benitez - Bibliotecária - CRB-1/3129



# EPIDEMIA DA ZIKA E SEUS DESDOBRAMENTOS

*Uma abordagem multidisciplinar e  
integrativa dos seus efeitos no Brasil*

MAURICIO LIMA BARRETO

WAGNER DE JESUS MARTINS

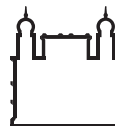
BETHÂNIA DE ARAÚJO ALMEIDA

MARIA GLÓRIA TEIXEIRA

MARIA YURY TRAVASSOS ICHIHARA

RICARDO BARROS SAMPAIO

**Organizadores**

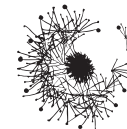


Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz

Instituto Gonçalo Moniz



**cidacs**

Centro de Integração de Dados  
e Conhecimentos para Saúde

# **PLATAFORMA DE VIGILÂNCIA DE LONGO PRAZO PARA A ZIKA E MICROCEFALIA NO AMBITO DO SISTEMA ÚNICO DE SAÚDE (PLATAFORMA ZIKA)**

## **ORGANIZADORES DA PUBLICAÇÃO**

Mauricio Lima Barreto  
Wagner de Jesus Martins  
Bethânia de Araújo Almeida  
Maria Glória Teixeira  
Maria Yury Travassos Ichihara  
Ricardo Barros Sampaio

## **INSTITUIÇÃO EXECUTORA DA PLATAFORMA ZIKA**

FUNDAÇÃO OSWALDO CRUZ – FIOCRUZ  
Centro de Integração de Dados e Conhecimentos para Saúde (Cidacs/Fiocruz Bahia)  
Gerência Regional de Brasília (Gereb/Fiocruz)

## **INSTITUIÇÃO FINANCIADORA DA PLATAFORMA ZIKA**

Ministério da Saúde, Brasil

## **INSTITUIÇÕES PARCEIRAS DO PROJETO**

Universidade Federal da Bahia (UFBA)  
Secretaria de Vigilância em Saúde e Ambiente do Ministério da Saúde (SVSA)  
Universidade de Brasília (UNB)  
London School of Hygiene & Tropical Medicine (LSHTM)

## **AUTORES DA PUBLICAÇÃO**

Alexandro Rodrigues Pinto  
Bethânia de Araújo Almeida  
Bruno Justino Garcia Praciano  
Bruno Santos Ferreira  
Carlos Antônio de Souza Teles Santos  
Daniela Almeida  
Eduardo Hage Carmo  
Eliseu Silva Torres

Enny Cruz Paixão  
Eslaine Santos e Santos  
Fábio Castro Gouveia  
George Caique Barbosa  
Giovanny Vinícius Araújo de França  
Júlia Moreira Pescarini  
Juliane Fonseca Oliveira  
Larissa Catarina Costa

Laura Cunha Rodrigues  
Lavínia Schuler-Faccini  
Liliana Cabral  
Luciana Lobato Cardim  
Luiz Conrado de Farias Neto  
Maíra Lima de Souza  
Márcia Luz da Motta  
Márcio Aldrin França Cavalcante  
Marcos Aires  
Maria da Conceição Nascimento Costa  
Maria Glória Teixeira  
Maria Yury Travassos Ichihara  
Mauricio Lima Barreto  
Micael Filipe Ribeiro de Lima  
Mônica Angélica Carreira Fragoso  
Moreno Magalhães de Souza Rodrigues

### **EQUIPE TÉCNICA**

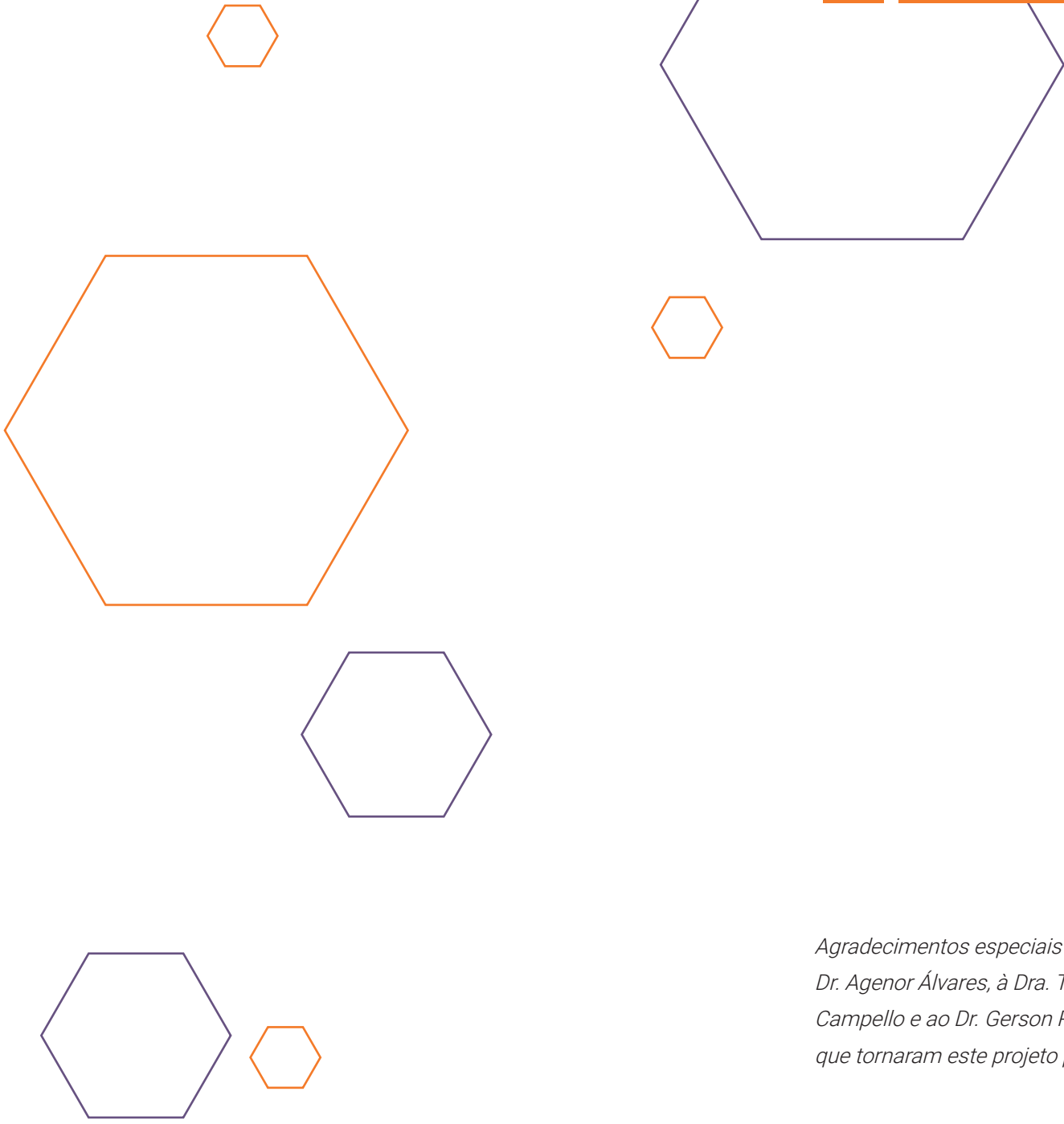
Alayde Lopes Sarno Carvalho  
Cristiana Paula Matos de Almeida  
Danilo Doneda (*In memoriam*)  
Denise Moraes Pimenta  
Érika Carvalho de Aquino  
Fernanda Eustáquio  
Marcelo Souza de Jesus  
Maria Jenny Silva Araújo  
Mariella Silva de Oliveira Costa  
Paulo Roberto Pereira Rocha  
Pedro Henrique Aguiar  
Pedro Ornelas  
Sávia Silva Soares Pueyo  
Wagner Robson Manso de Vasconcelos

Nívea Bispo da Silva  
Paula Xavier dos Santos  
Priscilla Normando  
Qeren Hapuk Rodrigues Ferreira Fernandes  
Rafael Valente Veiga  
Rebeca da Silva Barros  
Ricardo Barros Sampaio  
Rita de Cassia Carvalho-Sauer  
Roberto Fernandes Silva Andrade  
Roberto Perez Carreiro  
Robespierre Dantas da Rocha Pita  
Rosemeire Leovigildo Fiaccone  
Samila Oliveira Lima Sena  
Vanessa Arruda Jorge  
Wagner de Jesus Martins  
Wanderson Kleber de Oliveira

### **COLABORADORES**

Cinthy Vivianne de Souza Rocha Correia  
Cláudia de Souza Ferreira Martins  
Claudio Maierovitch Pessanha Henriques  
Danielle Soares Cavalcante  
Fernanda Machiner  
Flávia Elias  
Gabriela de Oliveira Silva  
Gerson Oliveira Penna  
Gustavo Buss  
José Agenor Álvares da Silva  
Joyce Mendes Scrham  
Kellen Santos Resende  
Patrícia Henning  
Paulo Cezar Vieira Guanaes  
Paulo Ernani Gadelha Vieira  
Raiza Tourinho dos Reis Silva Lima





*Agradecimentos especiais ao Dr. Agenor Álvares, à Dra. Tereza Campello e ao Dr. Gerson Penna, que tornaram este projeto possível.*

## Prefácio

---

Esta publicação vem a público em um momento estratégico e oportuno da vida social e política brasileira, marcado pelo fim da Declaração de Emergência de Saúde Pública de Interesse Internacional (ESPII) relacionada à Pandemia de Covid-19 pela Organização Mundial da Saúde (OMS) e pelo início de um novo ciclo político no Brasil que aponta para a retomada de uma agenda voltada à defesa dos direitos humanos, justiça social e ambiental.

É nesse contexto de retomada que se torna fundamental retomar a epidemia de Zika Vírus e suas repercussões como um analisador das iniquidades sociais, científica, sanitária e ambiental em contextos globais e locais.

Dividida entre as imagens dos mosquitos e dos bebês microencefálicos, a epidemia de Zika tomou de assalto a comunidade científica, os noticiários e a sociedade entre os anos de 2015 e 2016. Esse período foi insuficiente para tornar politicamente evidente, muito além do controle de vetores e da espetacularização das imagens das crianças, os problemas e estigmas mobilizados pela epidemia e suas consequências. Foram expostas a invisibilidade e precariedade da atenção a crianças com múltiplas deficiências e o desfinanciamento da Rede de Atenção à Pessoa com Deficiência, as questões de gênero, de direitos sexuais e reprodutivos, a iniquidade do financiamento e apoio à pesquisa nas instituições brasileiras, o desafio da expansão das cooperações internacionais, a força da atenção primária e da vigilância em saúde no Sistema Único de Saúde (SUS), a desigualdade da resposta de estados e municípios, a inteligência científica, resiliência e capacidade das pesquisadoras brasileiras, os desafios da inclusão escolar, entre outros.

Após o término da ESPII relacionada ao Zika Vírus (ZIKV) em novembro de 2016, ainda com muitas perguntas não respondidas, os recursos e importância política da epidemia foram perdendo espaço, associados a uma política de austeridade

econômica e ameaça democrática, impactando de forma deletéria a pesquisa, o desenvolvimento de imunobiológicos relacionados à vacina e diagnóstico sorológico do ZIKV, bem como, e principalmente, a assistência e proteção social das famílias e crianças afetadas.

O Cidacs/Fiocruz Bahia teve sua origem em 2016 e juntamente com a Gerência Regional de Brasília (Gereb) e a Secretária de Vigilância Sanitária do Ministério da Saúde (SVS/MS)<sup>1</sup>, rapidamente organizaram um termo de cooperação para investigação e análises de dados sobre a epidemia de Zika, como também, subsidiar políticas e iniciativas sociotécnicas, resultando numa intensa e profícua atividade de pesquisa e inovação social e tecnológica.

Este livro é resultado desse enorme esforço e dedicação de um coletivo que reconhece e objetiva intervir sobre os determinantes e políticas sociais que impactam na qualidade de vida e saúde da população. As diferentes iniciativas de pesquisa e ação social apresentadas nos capítulos a seguir, representam uma retomada de experiências e análises que não ficarão relegadas apenas ao registro e memória de uma epidemia. Outrossim, constituem uma base técnica, científica e social para o enfrentamento do Zika Vírus e suas consequências, bem como, modelos para análise e resposta a crises sanitárias emergentes e reemergentes.

Desde a integração de grandes bases de dados sobre as coortes epidemiológicas e de vigilância longitudinal da Síndrome Congênita do Zika (SCZ), passando pela identificação e disseminação de inovações tecnológicas para o enfrentamento das arboviroses, até os desafios e regulação da Ciência Aberta e compartilhamento de dados, a publicação apresenta fortes evidências de um problema inacabado, bem como, oferta uma

---

1. A Secretaria de Vigilância Sanitária do Ministério da Saúde (SVS/MS), passou a ser denominada Secretaria de Vigilância Sanitária e Ambiente (SVSA/MS) por meio do Decreto nº 11.358 de 1º de janeiro de 2023.



agenda de continuidade e inovação nos estudos e ações sobre o Zika, e também para o desenvolvimento da pesquisa engajada no Brasil.

Neste sentido, gostaria de destacar para o debate, três elementos presentes nesta publicação que se impõem de forma urgente como uma agenda para o Brasil e países do chamado Sul Global:

1. O uso e integração de distintas fontes de dados constituem um poderoso instrumento para análise e subsídio para formulação e avaliação de políticas públicas que interfiram na redução das desigualdades sociais e ambientais;
2. A necessidade do fortalecimento de uma ciência cidadã voltada para a defesa dos direitos humanos, da saúde como direito, e da participação da sociedade para a produção e desenvolvimento de conhecimentos e tecnologias que atendam as necessidades das populações, especialmente as mais vulnerabilizadas;
3. A construção de redes sociotécnicas com a participação e ações sinérgicas entre pesquisadores, sociedade, poder público e iniciativa privada, para busca de soluções concretas para problemas de saúde e sociais de base territorial.

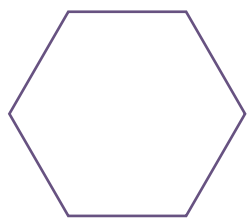
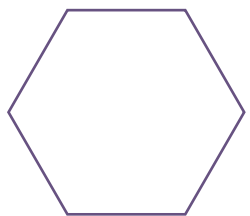
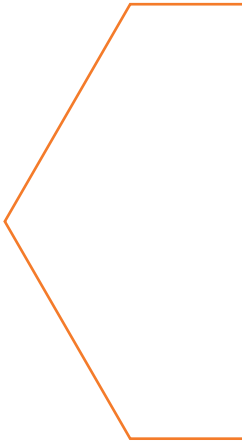
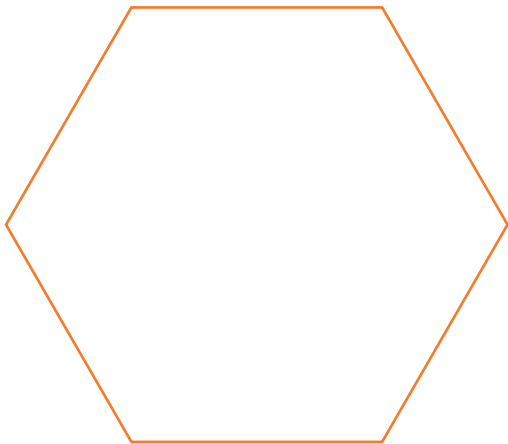
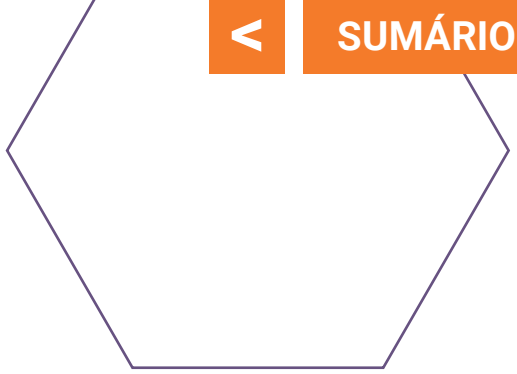
Convido-os a ler atentamente as pesquisas, análises e experiências aqui reunidas para juntos, construirmos conhecimentos, modelos de práticas e políticas que defendam a saúde e a vida.

### **Gustavo Corrêa Matta**

Pesquisador Titular em Saúde Pública da Fundação Oswaldo Cruz (Fiocruz)

Pesquisador do Cidacs / Fiocruz Bahia

Coordenador do Núcleo Interdisciplinar sobre Emergências em Saúde Pública (NIESP) e da Rede Zika Ciências Sociais da Fiocruz.



# Sumário

APRESENTAÇÃO	12
RESUMO EXECUTIVO	18
INTRODUÇÃO	24
TRANSFORMANDO DADOS ADMINISTRATIVOS EM PRODUÇÃO DE CONHECIMENTO: APLICANDO NA PESQUISA DA SÍNDROME CONGÊNITA DO ZIKA	28
VIGILÂNCIA LONGITUDINAL DA ZIKA E DA SÍNDROME CONGÊNITA DO ZIKA	68
PROSPECÇÃO DAS INICIATIVAS DE PESQUISA SOBRE CONTROLE E TRATAMENTO DE ARBOVIROSES	126
FEIRAS DE SOLUÇÕES PARA A SAÚDE	172
CIÊNCIA ABERTA E DADOS PARA PESQUISA: CONTEXTO, DESAFIOS E OPORTUNIDADES	239
CONSIDERAÇÕES FINAIS	276



# Apresentação

*Wagner de Jesus Martins*

O Brasil, em 11 de novembro de 2015, declarou Emergência em Saúde Pública de Importância Nacional (Espin) por conta da ocorrência do número de bebês nascidos com microcefalia, que rapidamente teve demonstrada sua associação causal com o Zika Vírus (ZIKV). Nesse contexto o Cidacs/Fiocruz Bahia, juntamente com a Gereb/Fiocruz Brasília e a SVS/MS, firmaram um Termo de Cooperação para a realização do Projeto “Plataforma de vigilância de longo prazo para a Zika e Microcefalia no âmbito do Sistema Único de Saúde (SUS)”. O Projeto teve como objetivo constituir a Plataforma de integração de conhecimentos centrado em uma coorte epidemiológica (coorte Zika/microcefalia) com diferentes bases de dados da saúde (Sistema de Informações sobre Nascidos Vivos - SINASC / Sistema de Vigilância Alimentar e Nutricional - SISVAN / Sistema de Informação sobre Mortalidade - SIM / Sistema de Informações Hospitalares - SIH / Sistema de Informações Ambulatoriais - SIASUS) e do desenvolvimento social (Cadastro Único para Programas Sociais do Governo Federal CadÚnico / Benefício de Prestação Continuada - BPC), para acompanhamento de longo prazo das condições de vida da população acometida pelo Zika Vírus, em especial, aquela com microcefalia.

O projeto é composto por cinco metas, que se articulam de modo a contemplar a epidemia da Zika na sua complexidade e interfaces:

1. Acompanhar a população acometida pelo Zika Vírus/microcefalia, em forma de coorte epidemiológica;
2. Realizar prospecção das iniciativas de pesquisa e tratamento das arboviroses como Zika Vírus/Dengue/Chikungunya/microcefalia;

3. Fortalecer propostas de colaboração/cooperação na rede do complexo produtivo da saúde (Instituições de Ciência e Tecnologia - ICT/empresas);
4. Definir protocolos de acesso a informações de saúde pública e pessoais;
5. Desenvolver bancos de dados de evidências científicas seguindo os preceitos de *open science/open data*.

Com esta publicação buscamos disseminar os conhecimentos gerados, as ferramentas criadas, conceitos incorporados e metodologias desenvolvidas. Ao longo do projeto tivemos inúmeras pessoas envolvidas na implementação das suas metas, foram muito mais que a lista de autores dos capítulos. Agradecemos a cada acadêmico, pesquisador(a), técnico(a), gestor(a), militante, comunicador(a), estudante enfim, a todos que nos ajudaram nessa jornada.

O livro é composto por cinco capítulos que trazem a síntese do que foi produzido por cada uma das metas propostas.

Em seu primeiro capítulo (Meta 4), a professora Maria Yury Travassos Ichihara e sua equipe nos apresentam o uso de dados administrativos através das análise integrada de grande volume desses dados, para a produção de conhecimentos e evidências para uma vigilância a longo prazo da Síndrome Congênita do Zika (SCZ). A inovadora metodologia de transformar os dados administrativos em dados de pesquisa em saúde, exigiu a produção de marcos regulatórios éticos e de segurança e privacidade da informação, para que o processamento e a análise de dados pudessem ser realizadas em conformidade com as leis e demais instrumentos regulatórios, em um ambiente controlado e seguro. *“A iniciativa trazida pelo Cidacs/Fiocruz Bahia apresenta um modelo inovador, para a realidade brasileira, de governança e gestão de dados.”* dizem Ichihara e colaboradores.

Com base nesses parâmetros estabelecidos, os autores explicam como foi construída a Coorte de Nascimentos, com mais de 28 milhões de nascidos vivos, como foram feitas as análises dos efeitos das condições obstétricas pré-natais, infecções congênitas e potenciais determinantes sociais e ambientais, além do impacto das políticas sociais no nascimento, crescimento, morbidade e sobrevida, em geral e em subgrupos de interesse.

No Capítulo 2 (Meta 1), a professora Maria Glória Teixeira e equipe, nos apresentam a Plataforma de Vigilância Longitudinal da Zika e Síndrome Congênita do Zika (Meta 1), como uma ferramenta de integração de diferentes bases de dados sociais (saúde, educação, previdência e desenvolvimento social), que possibilitou o desenvolvimento de estudos e pesquisas longitudinais de longo prazo para gerar conhecimentos sobre o espectro da Síndrome Congênita do Zika (SCZ), seu impacto epidemiológico e nas condições de vida dos indivíduos e populações acometidas por microcefalia/SCZ e pela Zika. Os estudos de longo prazo foram capazes de detectar e acompanhar desfechos adversos nas crianças, diagnosticadas, ou não, com microcefalia desde o nascimento.

No Capítulo 3 (Meta 3), que trata da realização da prospecção das iniciativas de pesquisa, controle e tratamento de arboviroses relacionadas com a tríplice epidemia (Zika, dengue e chikungunya), Ricardo Barros Sampaio e colegas dedicam-se a explicar o que estava acontecendo em relação a Zika no mundo científico. Foram realizadas avaliações sobre a produção científica e demais iniciativas de pesquisa nas áreas de controle e tratamento das arboviroses. Isso foi feito utilizando-se da Cientometria como ferramenta para o mapeamento das pesquisas em Zika e dos pesquisadores envolvidos no estudo da tríplice epidemia. Foram desenvolvidas ferramentas digitais como o e-Lattes, para extração e análise dos currículos, além de utilizar-se do VOSviewer do *Centre for Science and Technology Studies* (CWTS) da Universidade de Leiden, para a visualização dos dados.

Aqui o grupo apresenta como as pesquisas e estudos utilizaram-se de dados administrativos, coletados de forma sistemática e contínua pela rede de serviços de saúde do SUS, os quais revelaram-se potentes para obter evidências longitudinais de longo prazo. Por isso, apresentam a perspectiva de incluir nesta plataforma integrada, bancos de dados de outras áreas e instituições como de Previdência, Educação e Desenvolvimento Social. Demonstra-se também, a importância da utilização da técnica de *linkage* de dados aplicados a banco de dados administrativos e epidemiológicos. O desenvolvimento de algoritmo para classificação automatizada de casos de SCZ, baseado na técnica de aprendizagem de máquina, pôde corrigir a prevalência nacional das Anomalias Congênitas (AC) em 17,9% no período analisado.

No capítulo quatro (Meta 3) estão descritas as Feiras de Soluções para a Saúde ,que formaram um programa com os seguintes objetivos:

1. Identificar e mapear os atores, tecnologias nacionais/internacionais, e horizontes tecnológicos;
2. Criar vínculos sistêmicos de cooperação e redes interorganizacionais; apresentação e difusão de inovações tecnológicas, sociais e de serviços de saúde aplicadas à Zika, Dengue e Chikungunya e à Síndrome Congênita do Zika (SCZ) e a discussões de questões de saúde pública relevantes, como a emergência de saúde pública causada pelo vírus Zika e da pandemia do novo coronavírus, Agenda 2030, *One Health*, Territórios Saudáveis e Sustentáveis, dentre outros temas, é apresentada pela equipe que trabalhou na realização do ciclo de feiras sob a coordenação de Wagner de Jesus Martins.

Foram realizadas quatro feiras tecnológicas, denominadas de Feiras de Soluções para a Saúde, em diferentes regiões do Brasil, de modo a identificar e disseminar inovações tecnológicas em saúde e contribuir, não somente com o enfrentamento das arboviroses, como também, que estivessem em consonância com a Agenda de Prioridades em Pesquisa do Ministério da Saúde, com os Objetivos de Desenvolvimento Sustentável (ODS) e a Agenda 2030 da Organização das Nações Unidas.

As Feiras se constituem como uma estratégia base para a ciência cidadã, pois seu caráter científico e inovador visa a promoção de intercâmbio de experiências, conhecimentos e a formação de parcerias e redes cooperativas entre públicos de três pilares da ciência e tecnologia em saúde (social, tecnológico e de serviços).

As experiências exitosas das quatro Feiras de Soluções evidenciaram a capacidade catalisadora dos eventos para o compartilhamento de experiências bem-sucedidas, divulgação e fortalecimento de ações de saúde, assim como, ativação de redes socio-técnicas nos territórios.

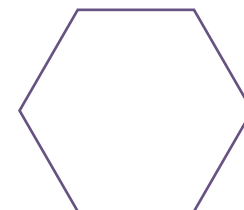
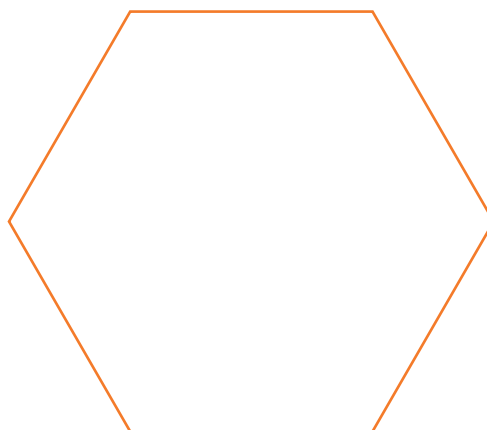
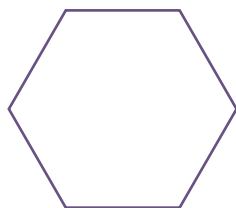
Destaca-se ainda, o uso da Plataforma Ágora Saúde, uma plataforma digital multiuso, que permitiu a geração e consolidação de um banco de dados (hoje com 349 soluções), que poderá servir de base para estudos continuados de inovação em saúde pública, e cuja divulgação ampla promoverá a maior integração entre ciência e o universo social e empresarial. As Feiras tiveram um importante papel na identificação do potencial da sociedade em desenvolver soluções para o enfrentamento de crises sanitárias independente das iniciativas governamentais, constituindo assim, uma vigilância popular de caráter territorial.

O movimento da Ciência Aberta vem ganhando força e visibilidade, por isso, Bethânia Araújo Almeida *et al.*, no Capítulo 5 (Meta 5), trabalharam com o objetivo de nos mostrar como obter uma ciência mais transparente e colaborativa com progressos mais rápidos e alinhada às necessidades e desafios das sociedades contemporâneas. A partir de 2002 com a *Budapest Open Access Initiative*, defende-se no movimento da Ciência Aberta a disponibilização livre e gratuita de publicações científicas. Assim, esse novo paradigma científico deve ser pautado no compartilhamento e reutilização de dados, de métodos e do conhecimento científico.

Por isso, o grupo nos traz o material produzido por ele para fechar, com chave de ouro, esta publicação. São eles: o Livro Verde - Ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e inter-

nacional; o livro intitulado Marcos regulatórios e a abertura de dados para pesquisa em saúde: dados pessoais, sensíveis ou sigilosos e propriedade intelectual; Política de gestão, compartilhamento e abertura de dados para pesquisa: princípios e diretrizes da Fiocruz, instituída por meio da Portaria nº 5.760, de 7 de dezembro de 2020; Cenários sobre Zika - contribuições da pesquisa científica com uso de dados administrativos vinculados; Percepções e experiências sobre compartilhamento e vinculação de dados para pesquisa e avaliação de políticas públicas na área da saúde e o Termo de referência sobre gestão, compartilhamento e abertura de dados para pesquisa.

Os pesquisadores e suas instituições estão imersos em contextos que podem potencializar ou dificultar a abertura e o acesso. Dessa forma, a criação de uma política institucional que defina novos parâmetros e contornos de produção, organização e comunicação da ciência, de forma democrática e inclusiva, torna-se fundamental, advoga o grupo. Os pesquisadores também apontaram a falta de transparência e impessoalidade no processo de cessão de bases de dados administrativos, com base em critérios bem estabelecidos e transparentes, pautados pela isonomia. Na era do digital, esta pode ser uma importante contribuição para que outros grupos se inspirem nessas ações e avancem na ciência digital em saúde, acelerando a criação de conhecimentos e na sua aplicação, esse é o desejo de todos nós.







# Resumo Executivo

---

Esta publicação apresenta resultados e produtos obtidos em um projeto coordenado pelo Centro de Integração de Dados e Conhecimentos para Saúde (Cidacs/Fiocruz Bahia) e Gerência Regional de Brasília (Gereb/Fiocruz Brasília), financiado pelo Ministério da Saúde por meio da Secretaria de Vigilância em Saúde e Ambiente (SVSA/MS).

Estruturada em cinco capítulos, o primeiro aborda a experiência de criação de uma plataforma voltada à integração de dados administrativos para produção de conhecimento e evidências científicas, em particular estudos longitudinais voltados à Síndrome Congênita do Zika (SCZ).

Dados administrativos são informações rotineiramente coletadas por diferentes instâncias governamentais para subsidiar atividades administrativas e prestação de serviços à população, como os sistemas de informação da área de saúde, por exemplo, podendo ser integrados com dados de outras fontes que, entre outras finalidades, possibilitam subsidiar a pesquisa científica e estudos avaliativos diversos.

Nessa perspectiva, o primeiro capítulo destaca a governança e gestão dos dados (aquisição, recebimento, armazenamento, elaboração de metadados, processamento, integração e disponibilização de *datasets* integrados e anonimizados a pesquisadores e gestores), em conformidade com os marcos regulatórios éticos, legais e de segurança da informação.

Métodos de curadoria de dados originais e integrados, vinculação e estimativa de acurácia foram desenvolvidos e implementados em ambiente computacional de alto desempenho, com alta segurança da informação. A estrutura de governança e gestão instituída na referida plataforma permitiu aos pesquisadores acesso a um conjunto de dados aptos a responder questões sobre a SCZ.

Diante da necessidade de melhor entender alguns aspectos do espectro clínico da SCZ e seus impactos sobre as crianças acometidas, percebeu-se que lacunas sobre a disseminação do vírus Zika no Brasil e a evolução do desenvolvimento dos fetos

e bebês acometidos só poderiam ser preenchidas mediante investigações de longo prazo. Assim, a equipe de pesquisadores conduziu estudos e pesquisas transversais e longitudinais apoiados por dados administrativos integrados, particularmente do Sistema Único de Saúde (SUS), os quais encontram-se sintetizados no segundo capítulo.

Entre os principais achados, destacam-se as evidências produzidas de que os casos notificados de Zika no curso dos primeiros anos epidêmicos estavam subregistrados, visto terem sido confundidos com a dengue. A iniciativa do Ministério da Saúde para aprimoramento do instrumento de registro de informações sobre nascidos vivos para captar o maior número possível de casos de microcefalia/SCZ, impactou positivamente na notificação de todas as anomalias congênitas, demonstrando que a prevalência nacional destes agravos vinha sendo subestimada em pelo menos 17%.

Outros estudos conduzidos no âmbito do projeto demonstraram que imagens anormais do cérebro e tamanho reduzido do perímetro cefálico foram os principais fatores para o aumento da mortalidade das crianças com SCZ nos três primeiros anos de vida, sendo o risco de morte 11,3 vezes maior entre os nascidos vivos com esta síndrome do que entre aqueles sem esta condição. Também foi possível, dentre outros resultados, avaliar diferenças espaciais nas estimativas de risco entre municípios brasileiros, desfechos adversos e marcadores sociais associados aos recém-nascidos com a SCZ.

No capítulo seguinte, é apresentado estudo que analisa a produção científica sobre controle e tratamento das arboviroses, tanto em âmbito nacional quanto internacional. O estudo utiliza métodos e ferramentas da Cientometria para analisar a distribuição geográfica da produção científica, os temas de pesquisa mais abordados e as dinâmicas de colaboração entre pesquisadores de diversos países. Adicionalmente, foi elaborado um mapa sobre o conhecimento científico voltado ao vírus Zika, que identifica áreas específicas e agrupamentos temáticos como saúde pública, aspectos clínicos e pesquisa laboratorial. O mapeamento e análise dos pesquisadores envolvidos, suas áreas de pesquisa e colaborações ao longo do tempo foram possíveis através da utilização da plataforma Lattes do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

As colaborações entre pesquisadores foram analisadas por meio de redes de coautoria para demonstrar as dinâmicas de colaboração entre os grupos de pesquisa. Além da utilização de métodos cientométricos, destaca-se a utilização da Plataforma Lattes e de recursos desenvolvidos no âmbito do projeto, como o e-Lattes, os quais possibilitaram a coleta, processamento e análise dos dados relacionados à produção científica dos pesquisadores brasileiros. Nessa área de conhecimento, a análise cientométrica envolveu a utilização de métodos estatísticos e matemáticos específicos para avaliar a produção bibliográfica e científica, além de análises de redes de colaboração visualizadas e analisadas através do software VOSviewer.

O quarto capítulo apresenta as Feiras de Soluções para a Saúde, concebidas como arranjos de interação entre diversos atores que se organizaram para lidar com os efeitos, bem como, a proposição de soluções para epidemia da Zika, com destaque para a atuação da sociedade civil. As Feiras tiveram por objetivo agregar uma diversidade de atores com vistas ao estímulo da formação de redes sociotécnicas.

Foram realizadas quatro Feiras entre os anos de 2017 e 2020. A primeira foi na Bahia, e teve por tema “Zika vírus, microcefalia e síndromes congênitas”. As duas seguintes ocorreram em 2019, uma no Rio Grande do Sul, na qual o tema da “Saúde Única para Territórios Saudáveis e Sustentáveis” foi abordado, e outra no Ceará, que teve por tema “Saúde Digital para Territórios Saudáveis e Sustentáveis”. Em virtude da pandemia de Covid-19 ocasionada pelo vírus SARS-Cov2, em 2020 a Feira de Soluções para a Saúde foi realizada de forma virtual. Inspirada pela convivência com o vírus SARS-Cov2, explorou o tema “Enfrentando as crises sanitárias e epidemias: panoramas e perspectivas”.

As iniciativas inscritas nas Feiras de Soluções para a Saúde foram organizadas em três categorias, Social, Serviço e Industrial, que geraram um banco de dados com 288 soluções, sendo 65% para o serviço, 27% sociais e 8% industrial. Percebe-se que as iniciativas de Soluções Industriais foram as de menor número em todas as edições das Feiras, provavelmente devido à necessidade de investimento para o desenvolvimento

tecnológico e escalonamento de produção. Destaca-se ainda a maturidade tecnológica das Soluções Industriais inscritas, pois mais da metade (52%) estavam em fase final de desenvolvimento, uma vez que os produtos estavam sendo comercializados, ou em processos avançados de viabilidade.

Concebemos que as Feiras poderão auxiliar na identificação e fomento de investimentos em soluções alinhadas à Agenda Nacional de Prioridades de Pesquisa em Saúde do Ministério da Saúde, como também, possíveis arranjos produtivos entre diferentes agentes do Complexo Econômico e Industrial da Saúde (CEIS).

O quinto e último capítulo aborda o movimento da Ciência Aberta com ênfase na gestão, compartilhamento e abertura de dados para pesquisa científica. Inicialmente foi realizado um estudo sobre políticas, infraestruturas e estratégias internacionais, principalmente de pesquisas financiadas com recursos públicos. Dentre os achados, destacam-se que os principais mecanismos de indução para gestão e compartilhamento de dados de projetos de pesquisa ocorrem como requisitos nos editais de financiamento, e que as instituições de pesquisa passaram a investir em infraestrutura (*hardware* e *software*) e pessoal especializado para adequada gestão, armazenamento, preservação e acesso aos dados primários e secundários produzidos e (re)utilizados com a finalidade de pesquisa.

Outro estudo analisou os marcos legais nacionais que incidem sobre a abertura de dados para pesquisa em saúde. Seus achados indicam que as instituições de pesquisa, universidades públicas e agências de fomento brasileiras necessitam criar normas internas na forma de políticas para o adequado acesso, compartilhamento e abertura dos dados custodiados com finalidade de pesquisa na área da saúde. Políticas que devem estar aderentes aos marcos éticos e legais de proteção de dados pessoais e sensíveis, como também, dos direitos autorais sobre os bancos de dados.

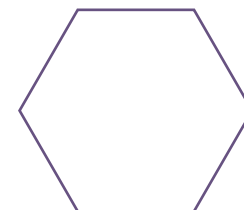
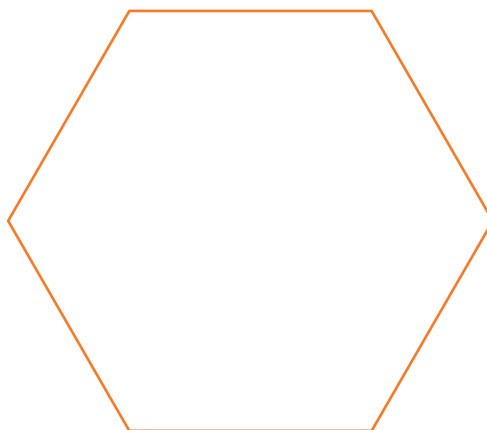
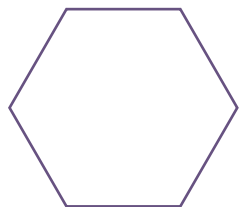
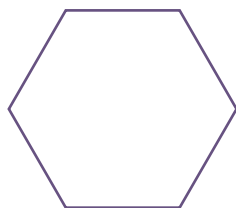
Os dois estudos mencionados serviram de subsídio para elaboração da Política de gestão, compartilhamento e abertura de dados para pesquisa da Fiocruz, cujos princí-

pios e diretrizes poderão ser utilizados como orientações pelo Ministério da Saúde e outras instituições públicas que custodiam dados utilizados para pesquisa científica.

Em seguida, uma pesquisa exploratória teve por objetivo compreender as percepções e atitudes de indivíduos de determinados segmentos da sociedade brasileira sobre o compartilhamento e a vinculação de dados pessoais para pesquisas científicas e políticas públicas. A investigação deu especial atenção aos dados administrativos, entendendo-os como bens públicos de grande valor para a sociedade.

Por fim, destaca-se que o livro é um registro sucinto da nossa experiência para um público diverso e amplo, notadamente gestores, pesquisadores e sociedade civil, com indicações e referências bibliográficas para aqueles que desejarem saber mais sobre os resultados e produtos apresentados.

A despeito da epidemia de Zika no Brasil ter ocorrido entre 2015 e 2016, seus impactos continuam a ser vividos cotidianamente pelas crianças e famílias acometidas pela SCZ. As experiências relatadas nesta publicação suscitam reflexões e apresentam contribuições ao debate, além de apontar caminhos que certamente continuaremos a trilhar para responder perguntas complexas de interesse para o SUS e a sociedade brasileira.







# Introdução

---

*Mauricio Lima Barreto*

Ao final de 2015, no Brasil foi observado o aumento inusitado de crianças nascidas com malformações congênitas importantes, principalmente malformações neurológicas caracterizadas por marcada redução do perímetro cefálico – microcefalia.

A princípio várias hipóteses foram levantadas para este novo problema, levando algum tempo até que se consolidasse a ideia de que a causa era o vírus Zika. Este vírus conhecido por várias décadas, até então não havia sido associado com efeitos patogênicos e em especial com malformações fetais. No Brasil, sua transmissão ocorre pelo mosquito *Aedes aegypti*, também transmissor da dengue e da chikungunya. Com a Zika, a partir de 2015, o grupo das arboviroses prevalentes no país ficava acrescido de mais um componente.

A dengue foi reintroduzida no Brasil em 1981 e, em 2011, uma revisão sobre os avanços no controle das doenças infecciosas no Brasil, mostrava-se que, apesar de esforços, a dengue estava entre aquelas doenças em que nenhum avanço havia sido obtido no tocante ao controle, e, pelo contrário, sucessivas epidemias e a introdução de novos sorotipos mostrava sinal de agravamento. Isto ocorria pela completa impossibilidade de controle do seu vetor que, dadas as condições dos ambientes urbanos, encontrava e continua encontrando amplas possibilidades de proliferação.

No segundo semestre de 2015, obstetras e neonatologistas, principalmente em Pernambuco, começam a relatar aumento na frequência de casos de microcefalia entre recém-nascidos. Em seguida, é comprovada a associação de casos de infecção pelo Zika vírus e microcefalia, que passa a ser denominada Síndrome Congênita do Zika (SCZ). Os casos relatados avolumam-se rapidamente e, em 11 de novembro de 2015, o Ministério da Saúde do Brasil decreta Emergência em Saúde Pública de Importância Nacional (Espin).

Desde a declaração desta nova epidemia, ficou nítido que governos, instituições de saúde pública, agências de financiamento de pesquisa, universidades e institutos de pesquisa, comunidades profissionais e científicas e a sociedade civil deveriam unir esforços para mostrar a capacidade de enfrentar mais um, dentre vários e complexos

problemas de saúde emergentes. Nesse contexto, a comunidade científica brasileira mobilizou-se e em colaboração com pesquisadores de vários outros países, inicia intenso processo de investigação sobre esta nova doença que, a partir da Região Nordeste, rapidamente dissemina-se pelo Brasil, e, em sequência, para outros países da América Latina.

Uma nova epidemia é em si um fenômeno complexo, e a reação da sociedade inicia-se por entender as suas causas e a diversidade e magnitude dos seus efeitos. O passo seguinte é o desenvolvimento de medidas e tecnologias possíveis de mitigar os seus efeitos e que possam contribuir para a prevenção de novas ocorrências. Tudo isto mobiliza diferentes atores e grupos de uma sociedade, que precisam atuar de forma coordenada, para ampliar as chances de sucesso na contenção dos efeitos do processo epidêmico.

Diante das dificuldades e necessidades impostas pelas arboviroses, sobretudo pela emergência de saúde causada pelo vírus Zika, um Termo de Cooperação entre a Fundação Oswaldo Cruz (Fiocruz) e a Secretaria de Vigilância em Saúde e Ambiente do Ministério da Saúde (SVSA/MS) foi constituído para a realização do Projeto “Plataforma de vigilância de longo prazo para a Zika e Microcefalia no âmbito do Sistema Único de Saúde (SUS)”. O Projeto teve como objetivo principal constituir uma plataforma de integração de conhecimentos e bases de dados dos campos da saúde e do desenvolvimento social para acompanhamento de longo prazo das condições de vida da população acometida com microcefalia e pelo Zika Vírus.

O presente projeto nasceu como mais uma entre diversas iniciativas arrojadas daquele momento. O núcleo central seria a construção de uma coorte que pudesse integrar grande volume de dados de distintas fontes sobre casos da Síndrome Congênita do Zika (SCZ) que permitisse trazer informações do que acontece com esta população de crianças que apresentam, ou não, patologia detectável (Meta 1). Associado a isto, elementos de suporte necessários à infraestrutura de dados originais e integrados nos seus aspectos tecnológico, ético e legal (Meta 4), a serem compatibilizados em acordo com os princípios da gestão e abertura de dados na perspectiva da Ciência Aberta

(Meta 5). Acoplaram-se ainda braços do estudo que buscariam entender dois outros aspectos de grande relevância no entendimento da gênese da epidemia e da sua dinâmica científica, tecnológica e social.

A busca de compreender a epidemia, suas causas, possibilidades de mitigação e prevenção gera uma intensa onda mobilizadora de colaboração na comunidade científica nacional e internacional, requerendo identificação dos atores envolvidos e análise das suas dinâmicas de colaboração (Meta 2), e ações voltadas à aproximação e estabelecimento de parcerias entre os desenvolvedores de tecnologias no complexo econômico industrial da saúde, de tecnologias sociais desenvolvidas pelos diversos movimentos de mães e famílias das crianças acometidas e demais movimentos da sociedade civil, no esforço de mitigar os efeitos da epidemia através de redes de solidariedade (Meta 3).

Portanto, os capítulos que compõem esta publicação sumarizam os esforços que mobilizaram mais de 40 pesquisadores de diversas áreas científicas, e que, de forma articulada, buscaram atingir os objetivos de cada uma das metas estabelecidas. A epidemia em sua complexidade é um fenômeno agudo e único, cujos efeitos se espraiam em diversas esferas da vida social. Com este espírito, a equipe de investigadores buscou, usando conhecimentos e métodos diversos, integrar suas investigações para tentar construir um corpo de conhecimentos efetivamente multidisciplinar, que contribua para o avanço científico, porém de forma mais importante, que contribua para aliviar o sofrimento das crianças e familiares que por circunstâncias indesejadas foram vítimas da infecção pelo vírus Zika.

No curso dos trabalhos, fomos surpreendidos com a pandemia de Covid-19, ainda mais global e devastadora que a epidemia de Zika. Porém, foi importante para nos fazer lembrar que epidemias e pandemias são eventos recorrentes na história das sociedades humanas, e que o presente estudo pode ser relevante para que entendamos eventos passados, os quais não tivemos oportunidade de investigar com grandes detalhes e, o mais importante, nos ajude a entender mais rapidamente eventos epidêmicos ou pandêmicos futuros que certamente ocorrerão.



# Transformando dados administrativos em produção de conhecimento: aplicando na pesquisa da Síndrome Congênita do Zika

---

*Maria Yury Travassos Ichihara  
Carlos Antônio de Souza Teles Santos  
Daniela Almeida  
George Caique Gouveia Barbosa  
Eliseu Silva Torres  
Liliana Cabral  
Luiz Conrado de Farias Neto  
Maíra Lima de Souza  
Nívea Bispo da Silva  
Priscilla Normando  
Rebeca da Silva Barros  
Roberto Perez Carreiro  
Robespierre Dantas da Rocha Pita  
Rosemeire Leovigildo Fiaccone  
Samila Oliveira Lima Sena*

No Brasil, os dados administrativos têm sido intensamente utilizados para produzir evidências e responder aos problemas de saúde vivenciados pela sociedade. Extensa literatura demonstra as potencialidades e limitações no uso desses dados na pesquisa em saúde, ressaltando-se o crescimento da qualidade das informações contidas nos sistemas de informação de saúde nas últimas décadas. O crescimento do uso de dados para produzir conhecimento em saúde intensificou-se em anos mais recentes, no contexto da Quarta Era Industrial, na qual ocorre um aumento exponencial da coleta e circulação de dados com seu consequente uso em vários campos do complexo produtivo e inclusive na pesquisa (CHEONG; CHANG, 2007).

O uso de ecossistemas de dados integrados passa a ser então, um dos grandes desafios na produção de conhecimentos de saúde. Isso implica em lidar não só com o grande volume de dados, mas com a heterogeneidade da sua estrutura, com a acessibilidade aos dados identificados geridos por diferentes proprietários, em compreender e seguir os marcos regulatórios de segurança e privacidade da informação, utilizar recursos computacionais e tecnologias cada vez mais complexos, além do desenvolvimento de metodologias de análises estatísticas adequadas para responder questões científicas.

No Brasil, o ecossistema de dados é constituído por sistemas de informação que coletam dados rotineiramente para efeito de registro ou transações, associados à prestação de um serviço ou, no caso de dados de saúde, para orientação de políticas, monitoramento de programas, vigilância, registros oficiais (por exemplo, óbitos e nascimentos) e auditorias. Assim, é possível obter dados socioeconômicos, demográficos e de saúde individualizados, provenientes do cadastro da população pobre e extremamente pobre elegível para programas sociais (Cadastro Único), da base de dados de beneficiários de programas sociais, como o Bolsa Família, Minha Casa Minha Vida, Cisternas, informações sobre nascimentos, óbitos, doenças de notificação compulsória, internações, entre outros.

Este contexto traz uma grande oportunidade única de integração e uso dos dados administrativos nominais individualizados para produzir conhecimento sobre os

problemas de saúde que afligem a população brasileira. Porém, para utilizar os dados administrativos em pesquisa em saúde, torna-se fundamental que eles sejam adequadamente curados e processados e assim transformados em evidências do mundo real (RWE), como são conhecidos os dados para pesquisa oriundos de atividades rotineiras, os quais tendo sido curados e processados podem tornar-se evidências científicas (CHODANKAR, 2021). Para tal, precisam ser tratados, integrados e organizados em estruturas que possibilitem a aplicação de métodos analíticos adequados. Exemplo dessas estruturas são a Coorte de 100 milhões de brasileiros e a Coorte de Nascimentos do Cidacs/Fiocruz Bahia .

Essas coortes são resultantes da vinculação de informações provenientes dos registros do Cadastro Único para programas sociais do governo federal e de benefícios recebidos em programas sociais (Bolsa Família, Minha Casa Minha Vida ou Cisternas) e de eventos de saúde (nascimentos, óbitos, internações, monitoramento da situação nutricional, notificações de doenças) (BARRETO, *et al.*, 2021).

Na Coorte de Nascimentos com mais de 28 milhões de nascidos vivos é possível analisar o efeito das condições obstétricas pré-natais, infecções congênitas e potenciais determinantes sociais e ambientais, além do impacto das políticas sociais no nascimento, crescimento, morbidade e sobrevida, em geral, e em subgrupos de interesse (PAIXÃO, *et al.*, 2021).

O processamento, integração e uso de dados administrativos identificados seguem os marcos regulatórios do país<sup>1</sup>.

Além disso, necessita estar em consonância com os princípios FAIR<sup>2</sup>, que recomendam que dados e metadados relacionados a dados de pesquisa devem ser Localizáveis, Acessíveis, Interoperáveis e Reutilizáveis.

---

1. a) Lei Geral de Proteção de Dados Pessoais (LGPD), Lei 13.709, 14.8.2018 (Brasil, 2018); b) Lei de Acesso à Informação, LAI - Lei nº 12.527, 18.11.2011 (BRASIL, 2011), que regula o acesso a informações; c) Constituição Federal (BRASIL, 1988), inciso XXXIII do art. 5º; inciso II do § 3º do art. 37; § 2º do inciso V, do art. 216 e d) Resolução CNS Nº 466/2012, que define diretrizes éticas referentes à não violação do direito à privacidade e ao anonimato das pessoas, quando do uso de dados na pesquisa envolvendo seres humanos (BRASIL, 2012).

2. Os princípios FAIR são relacionados a boas práticas em Curadoria e são denominados desta forma por conta de seu significado como um acrônimo: F(indable), A(ccessible), I(nteroperable) e R(eusable) (WILKINSON, 2016).

A transformação dos dados administrativos nominais em dados de pesquisa, em concordância com o arcabouço ético e legal, é realizada no Centro de Integração de Dados e Conhecimentos para a Saúde (Cidacs/Fiocruz Bahia) construído e inaugurado em dezembro de 2016, vinculado ao Instituto Gonçalo Moniz/Fiocruz Bahia. Para seu funcionamento, foram elaborados um conjunto de práticas (técnicas e sociais), que planejam e gerenciam o ciclo de vida do dado, visando seu uso em curto, médio e longo prazo, seguindo processos e fluxos de curadoria de dados, tratamento de dados e boas práticas de Segurança e Privacidade da Informação.

Esses processos implicam no recebimento, armazenamento, tratamento, vinculação desses dados, extração e finalmente, na disponibilização do acesso aos *datasets* integrados e anonimizados aos pesquisadores. É de responsabilidade da segurança da informação monitorar ativos, realizar gestão e relatórios de vulnerabilidades de dispositivos e sistemas, gerenciar e monitorar os acessos dos ambientes físicos e lógicos, garantindo o acesso adequado aos ambientes do Cidacs/Fiocruz Bahia. Além disso, deve orientar melhorias contínuas dos processos de segurança da informação e privacidade, gerenciar relatórios e notificação de eventos e incidentes de segurança da informação e promover cursos e treinamentos de maneira a educar e conscientizar os colaboradores do Cidacs/Fiocruz Bahia, sobre as práticas de segurança da informação e privacidade de dados.

Ressalte-se que o Cidacs/Fiocruz Bahia, em virtude da aprovação da LGPD somente em 2018, baseou sua estruturação na Regulação Geral de Proteção de Dados da Comunidade Europeia, a GDPR (*General Data Protection Regulation*) (UNIÃO EUROPEIA, 2016). Assim, no momento da discussão da LGPD em 2018, o Cidacs/Fiocruz Bahia, por seu acúmulo de experiência na implementação de regras e procedimentos para o tratamento de dados pessoais, em conformidade com regras de privacidade e segurança, teve importante participação no debate sobre o uso de dados pessoais na pesquisa em saúde, contribuindo para a elaboração do Artigo 13 dessa lei, o qual prevê que:



“na realização de estudos em saúde pública, os órgãos de pesquisa poderão ter acesso a bases de dados pessoais, que serão tratados exclusivamente dentro do órgão e estritamente para a finalidade de realização de estudos e pesquisas e mantidos em ambiente controlado e seguro, conforme práticas de segurança previstas em regulamento específico. Essas práticas devem incluir, sempre que possível, a anonimização ou pseudonimização dos dados, bem como considerem os devidos padrões éticos relacionados a estudos e pesquisas”.

Os seus parágrafos, ordenam as regras para o uso dos dados em pesquisa, quais sejam:

§ 1º A divulgação dos resultados ou de qualquer excerto do estudo ou da pesquisa de que trata o *caput* deste artigo em nenhuma hipótese poderá revelar dados pessoais.

§ 2º O órgão de pesquisa será o responsável pela segurança da informação prevista no *caput* deste artigo, não permitida, em circunstância alguma, a transferência dos dados a terceiro.

§ 3º O acesso aos dados de que trata este artigo será objeto de regulamentação por parte da autoridade nacional e das autoridades da área de saúde e sanitárias, no âmbito de suas competências.

§ 4º Para os efeitos deste artigo, a pseudonimização é o tratamento por meio do qual um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo, senão pelo uso de informação adicional mantida separadamente pelo controlador em ambiente controlado e seguro.

Considerando a necessidade de revisão e adequação do Cidacs/Ficruz Bahia à LGPD, vem sendo realizado o monitoramento dos riscos relacionados aos procedimentos, processos e atividades que envolvam o tratamento de dados pessoais, em consonância com uma Política de Gestão de Vulnerabilidades (ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS, 2013).

## GOVERNANÇA E GESTÃO DE DADOS

Para garantir a confiabilidade e qualidade do dado para utilização responsável e de forma transparente, é necessária a adoção de um modelo de governança e gestão de dados de forma eficaz (CHEONG; CHANG, 2007; VAN DER AALST; BICHLER; HEINZL, 2017; OECD, 2007). A governança de dados é importante porque define políticas e procedimentos no que diz respeito às regras, aos papéis e responsabilidades em torno do ambiente dos dados. Os processos de governança indicam como os usuários e os dados se relacionam (leis, controle de acesso, acordos de licença de uso, políticas, entre outros) (CHEONG; CHANG, 2007). Gestão de dados refere-se a uma coleção de conceitos, processos e práticas dedicadas ao uso de ativos de dados a fim de obter o máximo proveito para alcançar os objetivos institucionais. Deve cumprir os regulamentos existentes e abranger todo o ciclo de vida de um determinado ativo de dados, desde sua criação original até sua preservação e arquivamento (CHEONG; CHANG, 2007).

A necessidade de uma estratégia de gestão de dados torna-se mais evidente ante as perspectivas de reúso, compartilhamento e interoperabilidade de dados, norteadas pelos princípios FAIR, para a construção de ecossistemas digitais voltados à produção de conhecimento (WILKINSON, 2016).

### Governança e Gestão de Dados no Cidacs/Fiocruz Bahia

Com a finalidade de produzir conhecimento em saúde de forma ágil e menos custosa que as pesquisas com dados primários, e visando apoiar a tomada de decisões em políticas públicas sociais em benefício da sociedade, o Cidacs/Fiocruz Bahia criou uma Plataforma de Dados inovadora, voltada para a transformação de dados administrativos em dados para pesquisa.

Esta plataforma está organizada a partir de um *framework* de governança e gestão de dados para garantir a qualidade e confiabilidade de seus possíveis usos. Inclui como principais processos relativos ao ciclo de vida dos dados: a curadoria, a vinculação e processamento, a extração e o acesso a dados integrados anonimizados, a tecnologia da informação adequada e a segurança e privacidade da informação. A plataforma habilita o recebimento, o armazenamento, o processamento, a integração, a preservação e a disponibilização de grande volume de dados administrativos integrados a pesquisadores e gestores. As diretrizes de gestão de dados no Cidacs/Fiocruz Bahia visam salvaguardar e proteger informações do uso indevido, perda, acesso não autorizado, modificação ou divulgação, objetivando reforçar a cadeia de custódia dos dados

A Plataforma de Dados Integrados segue os regramentos éticos, de segurança e privacidade da informação existentes no Brasil, e suas configurações físicas e lógicas buscam a preservação da confidencialidade das informações pessoais e sensíveis, a fim de restringir os riscos de vazamento de dados e a possibilidade de re-identificação de indivíduos. Dessa maneira, para a produção dos dados, criou-se uma sala segura para recebimento e armazenamento das bases de dados identificadas, criação de metadados (elaboração de dicionário e descritiva das bases de dados), realização do pré-processamento, vinculação com estimativas de acurácia, extração de *datasets* dos dados integrados anonimizados e um ambiente de análise configurado para disponibilização dos *datasets* extraídos aos pesquisadores e gestores.

Para a disponibilização de *datasets* extraídos dos dados integrados para responder às questões de pesquisa sobre a Síndrome Congênita do Zika, foram elaborados protocolos para a obtenção, recebimento, catalogação, tratamento, extração, anonimização e disponibilização dos dados, sempre em aderência aos marcos legais e éticos, que preservam a segurança e a privacidade (Figura 1)<sup>3</sup>.

A Figura 1 apresenta o *framework* da governança e gestão de dados da Plataforma de Dados Integrados operacionalizada no Cidacs/Fiocruz Bahia.

---

3. Maiores detalhes sobre a plataforma do Cidacs/Fiocruz Bahia podem ser encontrados em BARRETO; ICHIHARA; ALMEIDA (2019).

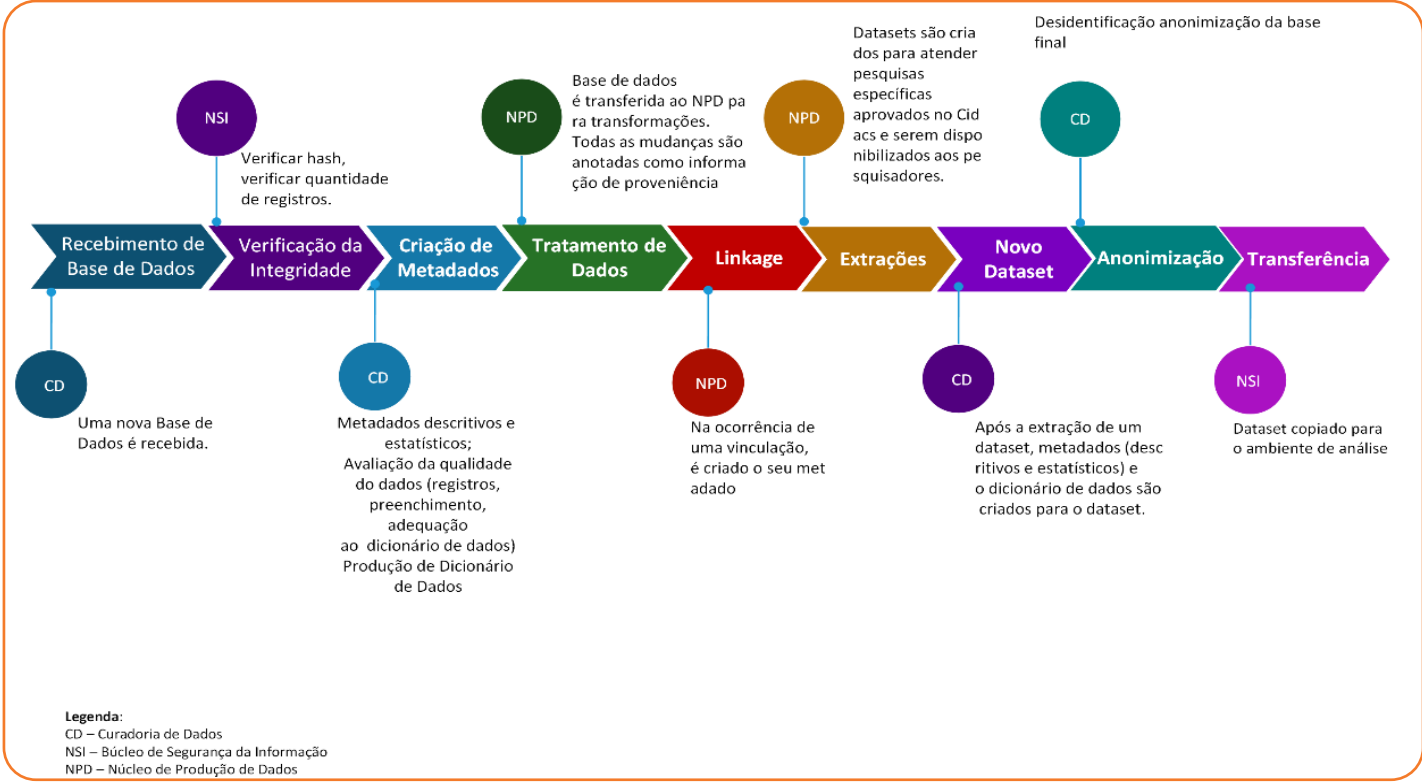
Figura 1: Plataforma de dados integrados do Cidacs/Fiocruz Bahia



Fonte: Elaborada pelo Cidacs/Fiocruz Bahia

A seguir, serão descritos os fluxos e protocolos de gestão de dados em pesquisa em saúde (Figura 2) referentes a: 1) Cessão e recebimento de bases de dados; 2) Inspeção das bases de dados; 3) Catalogação e disseminação dos metadados das bases de dados; 4) Pré-processamento padrão das bases de dados; 5) *Linkage* e acurácia das bases de dados; 6) Anonimização dos *datasets* integrados, 7) Acesso aos *datasets* integrados e anonimizados.

Figura 2: Fluxos e protocolos aplicados para gestão de dados em pesquisa em saúde



Fonte: Elaborada pelo Cidacs/Fiocruz Bahia

## 1) Cessão e recebimento de bases de dados

A cessão das bases de dados nominais do Ministério da Saúde para o Cidacs/Fiocruz Bahia para realização do estudo da Síndrome Congênita do Zika, foi realizada formalmente mediante ofício de solicitação das bases necessárias e assinatura de termos de transferência de responsabilidade pela instituição cedente (no caso a SVSA-MS), e termo de recebimento pelo Cidacs/Fiocruz Bahia, conforme o Quadro 1. O Cidacs/Fiocruz Bahia acessa as bases de dados nominais através da sua transferência por meio de dispositivo seguro, seguindo os protocolos de segurança e garantia da confidencialidade da informação. O responsável pela Segurança da Informação recebe as bases de dados, verifica a sua integridade física e realiza a sua transferência para a Sala Segura, mediante autorização do Curador. A base de dados nominal é copiada no HDFS, em local com permissão de acesso a Curadoria de Dados, que após realizar a inspeção e elaborar os metadados, libera o seu acesso para o tratamento do dado pela equipe de produção de dados. Toda a equipe que acessa os dados nominais assume compromisso de confidencialidade e privacidade através da assinatura do Termo de habilitação para acesso e manuseio de dados na sala segura do Cidacs/Fiocruz Bahia.

As bases de dados e suas atualizações solicitadas e cedidas para serem utilizadas no projeto abrangem o Sistema de Informação de Nascimentos (SINASC), Sistema de Informação de Mortalidade (SIM), Registro de Eventos de Saúde Pública (RESP), Sistema de Informação de Agravos de Notificação (SINAN), com registros de dengue, Zika e chikungunya.

No Quadro 1 encontra-se a relação de termos de transferência e recepção de dados nominais pelo Cidacs/Fiocruz Bahia.

Quadro 1: Relação de termos de transferência e recebimento de base de dados nominais

TERMO	TIPOLOGIA DO DOCUMENTO	USO
<b>Termos de solicitação de bases de dados</b>	Ofício	Solicitação de base de dados nominais com assinatura do Coordenador do Cidacs/Fiocruz Bahia e seu Curador.
	Termo de responsabilidade e confidencialidade para a cessão das bases de dados identificadas nos Sistemas de Informação em Saúde (SIM, SINAN, SINASC) sob gestão do Departamento de Análise em Saúde da Secretaria de Vigilância da Saúde e Ambiente/Ministério da Saúde.	Termo emitido pela SVSA/MS que deve ser assinado pelo Coordenador e Curador do Cidacs/Fiocruz Bahia.
<b>Termos de entrada das bases de dados</b>	Termo emitido pela SVSA/MS que deve ser assinado pelo Coordenador e Curador do Cidacs/Fiocruz Bahia.	Entrada de novas bases no Cidacs/Fiocruz Bahia (conta com assinatura do controlador para cessão da base).
	Termo de recebimento de dados pela Curadoria.	Entrada de novas bases no Cidacs/Fiocruz Bahia (conta com assinatura do curador para recebimento da base).
	Termo de sigilo e confidencialidade.	Entrada de novas bases no Cidacs/Fiocruz Bahia (conta com assinatura do curador para confidencialidade dos dados).

Fonte: Elaborado pelo Cidacs/Fiocruz Bahia

## 2) Inspeção das bases de dados

A inspeção da base de dados é etapa fundamental para a inclusão dos dados na plataforma do Cidacs/Fiocruz Bahia. O processo é realizado pela Curadoria de Dados e consiste em verificar a conformidade da base com seus descritores e possíveis modificações, em relação a versões anteriores já abrigadas na plataforma. Inclui: 1) análise da quantidade de registros e a correspondência entre as variáveis presentes no banco, em relação ao dicionário de dados disponibilizado pelo órgão cedente; 2) verificação da completude das variáveis, da presença de novas variáveis e da compatibilidade das novas versões das bases com as anteriores. Cumpridas essas etapas, são emitidos os termos de recebimento da base e o Termo de transferência de responsabilidade e guarda de base de dados nominais à Curadoria de Dados do Cidacs/Fiocruz Bahia - Lei n. 12527/2011 - (BRASIL, 2011), cujos signatários são representantes oficiais do Cidacs/Fiocruz Bahia e do Ministério da Saúde.

## 3) Catalogação e disseminação dos metadados das bases de dados

O Acesso Aberto refere-se à disponibilidade e acesso gratuito por qualquer pessoa aos resultados de pesquisas científicas. Baseia-se na premissa de que o conhecimento científico é um bem público, portanto, deve estar disponível a todos, e significa a livre disponibilidade de conteúdo digital de caráter científico a qualquer usuário.

Para promover o compartilhamento de dados de pesquisa na perspectiva da Ciência Aberta, o Cidacs/Fiocruz Bahia vem investindo no uso de um repositório de metadados em código aberto – o *Dataverse*. É uma ferramenta gratuita e de acesso aberto, permitindo a criação de um repositório estruturado de metadados, com base nos parâmetros do *Digital Documentation Initiative* e nos princípios FAIR. Fornece funcionalidades que ampliam a replicabilidade do dado de pesquisa em saúde, a partir da disponibilização de informações relevantes, como metadados e algoritmos sob medida, fluxos de trabalho, modelos e *software* (incluindo códigos-fonte).



Os metadados do *Dataverse* oferecem suporte para descrições de arquivos e conjunto de dados e são compatíveis com o esquema *DataCite* para registro DOI (baseando-se nos mais abrangentes padrões para interoperabilidade).

O *Dataverse* do Cidacs/Fiocruz Bahia<sup>4</sup> foi instalado localmente, permitindo a geração dos identificadores persistentes dos *datasets* integrados utilizados nas pesquisas. Possui uma interface *web*, que pode ser acessada pelo navegador por qualquer usuário habilitado, a fim de realizar consultas e recuperar informações sobre as bases de dados integradas disponíveis. Trata-se de uma plataforma de disseminação de dados e metadados, cujo principal benefício é oferecer a possibilidade de atribuição de um identificador digital persistente ao conjunto de dados inseridos na mesma. Estes identificadores servem especialmente para a área acadêmica, uma vez que proporcionam a replicabilidade e reprodutibilidade científica, além de expandirem a capacidade de disseminação.

Para o cadastro dos dados no *Dataverse*, as seguintes etapas foram realizadas:

- Construção do modelo lógico dos dados para cadastro, que envolve a criação de descritores de dados segundo padrão de vocabulário *Mesh*<sup>5</sup>;
- Criação de normas de cadastro e normatização dos parâmetros de citação;
- Criação de termos de acesso e disponibilização de dados.

O modelo conceitual de metadados para os conjuntos de dados do Cidacs/Fiocruz Bahia foi criado para refletir informações acerca dos *datasets* que ficarão disponíveis para os pesquisadores. Representa informações presentes (em nível abrangente), sem detalhes de implementação. Segundo nossos requisitos, devem ser gerados modelos lógicos para os seguintes tipos de metadados: descritivos, estatísticos, de *linkage* e de proveniência.

---

4. O *dataverse* do Cidacs/Fiocruz Bahia pode ser acessado em <http://dataverse.cidacs.org:8080/>

5. <https://www.ncbi.nlm.nih.gov/mesh/>

O Quadro 2 mostra como estes metadados do Cidacs/Fiocruz Bahia foram alinhados ao conjunto de metadados disponível na plataforma do *Dataverse*.

Quadro 2: Criação de termos de acesso e disponibilização de dados.

<b>TÍTULO</b>	Consiste na desabreviação das siglas envolvidas no conjunto de dados. <i>Exemplo: Pop 100_SIM = Extração da coorte de cem milhões de brasileiros vinculada ao sistema de informações sobre mortalidade</i>
<b>AUTORIA</b>	Expõe as instituições responsáveis pela produção das bases envolvidas em um conjunto de dados
<b>CONTATO</b>	E-mail para requerer informações sobre o conjunto de dados. <i>"cidacs.curadoria@fiocruz.br"</i>
<b>DESCRIÇÃO</b>	Breve texto indicando relação entre o conjunto de dados e os objetivos do projeto.
<b>ÁREA DO CONHECIMENTO</b>	Campo múltipla escolha, no caso do projeto Zika preenchido com: <i>"Medicine, Health and Life Sciences"</i> e <i>"Social Sciences"</i> .

Fonte: Elaborado pela Curadoria de Dados do Cidacs/Fiocruz Bahia

## TRANSFORMAÇÃO DE DADOS

Os processos de transformação de dados administrativos em dados para pesquisa encontram-se resumidos na Figura 3. Todo tratamento das bases de dados contendo dados sensíveis é realizado por pessoas autorizadas dentro da Sala Segura. O tratamento dos dados tem como principal objetivo processar e assegurar a qualidade dos dados para permitir estudos complexos que respondam a questões científicas relevantes com maior detalhe e precisão.

Para a etapa de vinculação de grandes bases de dados (*linkage*), o Cidacs/Fiocruz Bahia desenvolveu algoritmos avançados e inovadores, com alto grau de acurácia, ajustados a cada vinculação das bases necessárias aos estudos<sup>6</sup>. Os dados vinculados produzidos são catalogados e preservados. Por fim, são extraídos os *datasets* necessários para subsidiar as pesquisas e disponibilizados aqueles anonimizados, para atender à demanda específica do estudo previamente aprovado por comitê de ética e cujo plano de dados é aprovado pela Curadoria de dados após análise de sua viabilidade. Os *datasets* anonimizados que serão analisados pelo pesquisador são transferidos do ambiente da Sala Segura para o ambiente de análise pelo responsável pela Segurança da Informação, mediante autorização da Curadoria de Dados.

Os processos de transformação de dados administrativos em evidências de mundo real encontram-se na Figura 3.

Figura 3: Processos de transformação de dados administrativos em dados de pesquisa



Fonte: Elaborada pelo Cidacs/Fiocruz Bahia.

### Pré-processamento padrão das bases de dados

Foi desenvolvido um Procedimento Operacional Padrão para pré-processamento das bases de dados, o qual foi aplicado de forma geral, em todas as bases utilizadas para o projeto, preparando-as para o *linkage*. O Quadro 5 mostra a relação das bases de dados que foram pré-processadas para o projeto durante o período.

6. Maiores detalhes sobre o algoritmo de vinculação encontram-se BARBOSA *et al.* (2020).

Quadro 3: Relação das bases de dados pré-processadas para o projeto

BASE	PERÍODO	Nº DE REGISTROS
SINASC	2001-2018	53.209.685
SIM	2001-2018	20.900.881
RESP	2015-2020	18.688
Baseline da coorte (POP100 V2)	2001-2015	114.008.317
Atualização do baseline da coorte	2016-2018	17.689.483
SINAN-NI* Zika	2007-2018**	363.259
SINAN-NI* Dengue	2007-2018**	14.224.251
SINAN-NI* Chikungunya	2007-2018**	711.177

\*NI =Notificação Individual

\*\*2007-2016 refere ao período da versão da base SINAN-NET

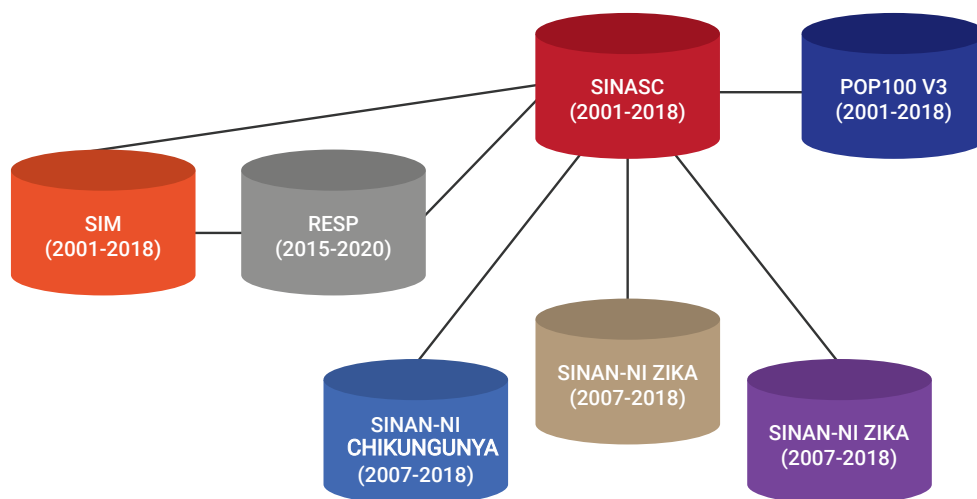
Fonte: Elaborado pela Produção de Dados do Cidacs/Fiocruz Bahia

## Linkage e acurácia das bases de dados

O *linkage* consiste em encontrar informações do mesmo indivíduo em diferentes bases de dados, por meio de integração entre elas. Foi utilizado o Cidacs-RL<sup>7</sup>, ferramenta desenvolvida pelo Cidacs/Fiocruz Bahia, que consiste na busca por similaridade de atributos comuns dos indivíduos (nome, nome da mãe, data de nascimento, sexo e município de residência) entre as bases de dados, utilizando-se técnicas de busca por indexação. Esta ferramenta apresenta alta escalabilidade e acurácia (BARBOSA *et al.*, 2020).

Foram realizados 31 *linkages* para produzir as extrações finais dos *datasets* para os pesquisadores do projeto, envolvendo: 1) SINASC e as bases do SINAN-NI; 2) SINASC e RESP; 3) SINASC e SIM; e 4) *Baseline* da Coorte de 100 milhões (POP100 V3) com SINASC (Figura 4).

Figura 4: Ilustração de alguns *linkages* realizados

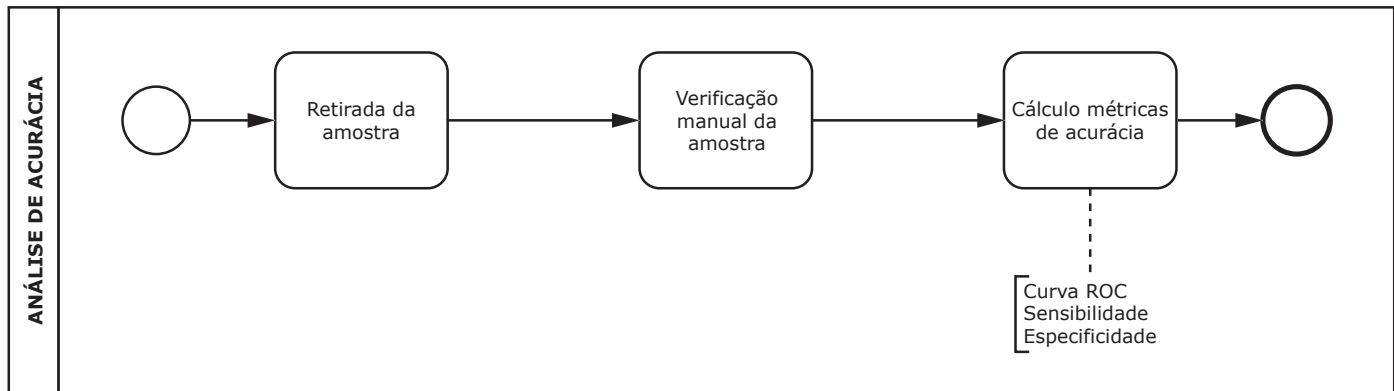


Fonte: Elaborada pelo Cidacs/Fiocruz Bahia

7. Esse algoritmo é um dos produtos deste projeto (BARBOSA *et al.*, 2020)

Foram realizados *linkages* entre as bases do SINASC e os registros de cada agravo (dengue, Zika e chikungunya) no SINAN-NI (Notificação Individual). De acordo com os agravos, as vinculações entre essas bases buscavam identificar as mães dos nascidos vivos do SINASC que foram notificadas no SINAN, e os nascidos do SINASC que foram notificados no SINAN, e realizar a identificação de entradas duplicadas. Assim, para cada agravo (em diferentes períodos) foram realizadas quatro vinculações, totalizando, para todo período do projeto, 33 execuções de *linkage* e cerca de 70 verificações manuais.

Figura 5: Fluxo de análise de acurácia dos *linkages* realizados



Fonte: Elaborada pelo Processamento de Dados do Cidacs/Fiocruz Bahia

Na análise de acurácia da vinculação de dados de duas bases de dados, retira-se uma amostra proporcional de 2.000 pares de registros conforme três faixas de escores (índice de similaridade): escore alto (acima de 0,95), escore intermediário (valores entre 0,90 e 0,95) e escore baixo (abaixo de 0,90). Em seguida, são realizadas as seguintes etapas: cria-se uma variável que informa a qual faixa de escore o par de registros pertence, calcula-se a proporção de registros em cada faixa de escore, retira-se uma amostra

proporcional a cada faixa de escore e divide-se as variáveis de data de nascimento em campos de ano, mês e dia para posterior comparação. Posteriormente, seleciona-se na amostra da base vinculada, sete campos de comparação, na ordem a seguir: a) Nome do indivíduo; b) Nome da mãe; c) Município de residência; d) Sexo; e) Ano de nascimento; f) Mês de nascimento; g) Ano de nascimento, que serão comparados entre os pares para realizar a classificação final do par vinculado e realiza-se a conferência da distribuição do escore em termos de medidas estatísticas (mínimo, máximo, média, mediana, quartis 1 e 3). Por fim, estima-se a Curva ROC para avaliar o melhor ponto de corte da base vinculada, considerando-se os valores de sensibilidade e especificidade. Os valores iguais a 1 são desconsiderados, pois automaticamente já são classificados como pares de registros verdadeiros, por apresentarem similaridade máxima. Para cada base oriunda de uma vinculação, é feita uma estratificação do escore, para que seja retirada uma amostra proporcional de 2.000 registros.

A estratificação da base vinculada e obtenção da amostra são feitas da seguinte forma:

- a) Divide-se o escore em três faixas de categorias: escore alto (acima de 0,95), escore intermediário (valores entre 0,90 e 0,95) e escore baixo (abaixo de 0,90), e cria-se uma variável que informa a qual faixa de escore o par de registro pertence;
- b) Calcula-se a proporção de registros em cada faixa de escore;
- c) Retira-se uma amostra proporcional a cada faixa de escore;
- d) Divide-se as variáveis de data de nascimento em campos de ano, mês e dia, para posterior comparação.

## Coorte de Nascimentos

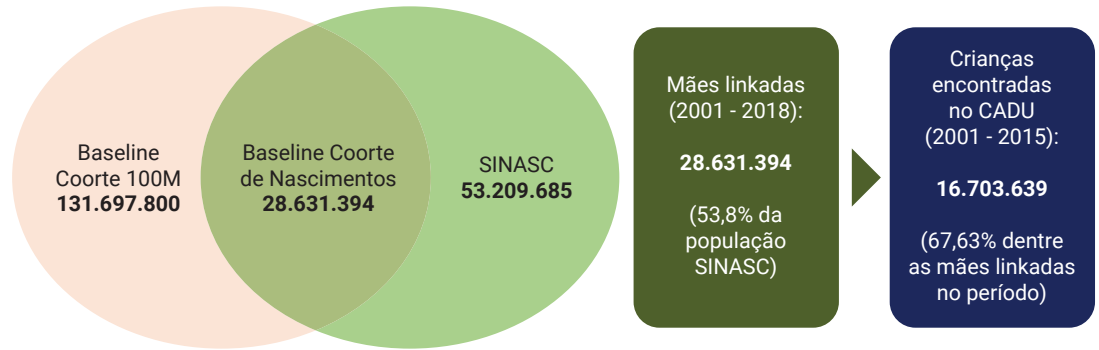
A Coorte de Nascimentos foi construída a partir da vinculação dos dados da população da linha de base da Coorte de 100 Milhões de Brasileiros e dos nascidos vivos (SINASC) para o período de 2001 a 2018 (Figura 6). Em primeiro lugar, foram identificadas as mães dos nascidos vivos no SINASC, que se encontravam no *baseline* da coorte. Para essa vinculação, foram utilizadas as variáveis: nome da mãe, município de residência e data de nascimento materna, ou idade materna. A data de nascimento materna foi introduzida no SINASC a partir do ano de 2011 e por isso, antes desse período, ou quando a data estava em *missing*, a idade materna foi considerada na vinculação.

Como resultado dessas duas estratégias, do total de registros do SINASC (53.209.685), 28.631.394 (53,8%) de nascidos vivos foram vinculados ao *baseline* da coorte pelas informações maternas.

Em segundo lugar, a partir das mães vinculadas, verificamos quais indivíduos do SINASC também estavam no *baseline* da Coorte. Para isso, foram considerados todos os códigos familiares (código de identificação da família nas bases do Cadastro Único, que pode variar ao longo do tempo), que as mães possuíam ao longo dos anos na base do Cadastro Único, e buscou-se registros de indivíduos que possuíam o mesmo sexo e a mesma data de nascimento do SINASC. Os resultados dessa busca mostraram que 16.703.639 nascidos vivos foram encontrados no *baseline* da coorte, representando 67,6% das mães vinculadas no período de 2001 a 2015.



Figura 6: Coorte de Nascimentos



Fonte: Elaborada pelo Cidacs/Fiocruz Bahia

### Extração de *datasets*

Após a vinculação das bases de dados necessárias para a pesquisa, é realizada a extração do *dataset* contendo os dados solicitados nos planos de dados para responder a questões específicas da pesquisa.

O plano de dados é um documento que deve ser preenchido pelos pesquisadores para requisitar o conjunto de dados que serão extraídos da base integrada e comporão o *dataset* adequado para uma estudo específico.

O plano de dados contem o detalhamento dos dados provenientes das bases integradas que serão utilizados pelo solicitante, bem como, informações sobre o estudo requisitante, permitindo o manejo de dados sensíveis e o rastreamento do uso de dados no Cidacs/Fiocruz Bahia. Sinaliza uma demanda de um projeto de pesquisa, ou de um gestor (ambos necessariamente e devidamente cadastrados no Cidacs/Fiocruz Bahia).

Para a elaboração do plano de dados é necessário especificar o período e as variáveis segundo os dicionários disponibilizados pelo Cidacs/Fiocruz Bahia, após o pré-processamento das bases originais.

Um plano de dados é considerado aprovado quando a seleção de variáveis se adequa ao propósito e finalidade do estudo, segundo parecer ético, e aprovação da coordenação. A seleção do conjunto de dados para extração é feita levando-se em conta a adequação dos dados à finalidade da pesquisa e o parecer ético aprovado.

Após a extração do *dataset* integrado será elaborado o seu metadado (dicionário, e descritivas e informações sobre acurácia dos *linkages* realizados).

### **Anonimização dos *datasets* integrados**

Segundo o artigo 5º, item XI da Lei Geral de Proteção de Dados (BRASIL, 2018), a anonimização consiste na utilização de meios técnicos razoáveis e disponíveis no momento do tratamento, por meio dos quais um dado perde a possibilidade de associação, direta ou indireta, a um indivíduo. Segundo Bioni (2020), anonimização é o processo que não pode ser revertido nem mesmo por quem o procedeu. Assim, esse processo visa minimizar o risco de reidentificação dos indivíduos através do uso de informações sensíveis, garantindo a provacidade dos dados.

## POLÍTICA DE ANONIMIZAÇÃO EM DESENVOLVIMENTO NO CIDACS/FIOCRUZ BAHIA

A LGPD conceitua dado anonimizado e anonimização, respectivamente, à “ocasião” e ao “momento” no qual o processo é realizado (BIONI, 2020). Nesse sentido, as decisões em torno do processo de anonimização, segundo as definições presentes no ADF (*The Anonymisation Decision-Making Framework*) estão atreladas fortemente à relação estabelecida entre o contexto e o propósito de uso dos dados, de forma que os dados e o ambiente são um sistema total (situação de dados) (MACKEY; ELLIOT; O’HARA, 2016).

Para desenvolver a política de anonimização mediante adequação à LGPD e aplicação de métodos de anonimização visando manter a privacidade e liberdade do titular dos dados, observa-se as seguintes vertentes:

- 1) Jurídica: relacionada às questões da LGPD;
- 2) Científica: referente às questões ligadas à governança e ao acesso aos dados segundo marcos legais e regulatórios. Tem relação com o desenvolvimento da parte metodológica, que envolve a implementação do processo de anonimização. Busca-se avaliar as possibilidades, oportunidades de melhorias e limitações do método e do cenário de uso dos dados;
- 3) Técnica: dedicada à aplicação do método de anonimização, desenvolvendo rotinas e executando experimentos.

As seguintes regras foram instituídas pelo Cidacs/Fiocruz Bahia para o acesso e compartilhamento de dados anonimizados para fins de pesquisa em saúde pública:

### *Quem acessa*

- Somente usuários autorizados e devidamente registrados podem acessar dados no Cidacs/Fiocruz Bahia;

- O uso do dado deve ser justificado;
- O acesso aos dados ocorre apenas em ambiente seguro e controlado;
- As análises são executadas em ambientes seguros e apenas os resultados de análises (tabelas, gráficos, etc) podem ser extraídos deste ambiente (mediante solicitação e análise pelo responsável pela Segurança da Informação e Curadoria de Dados).

### ***Métodos e estratégias de anonimização***

As seguintes técnicas de pseudonimização e anonimização vêm sendo utilizadas:

- Dados nominais de indivíduos presentes nos *datasets* são desidentificados e designados e códigos de identificação encriptados;
- O uso do *hash*;
- A supressão de variável e a generalização.

Para estas técnicas de anonimização aplica-se a seguinte classificação das variáveis por sensibilidade de divulgação:

- a) Variáveis Identificadoras: dados diretos de um indivíduo, como nome, CPF, RG, NIS (Número de Identificação Social);
- b) Variáveis Semi-identificadoras (variáveis-chave): referentes a valores que, quando combinados com outras informações, podem revelar a identidade do indivíduo. Exemplos: datas de nascimento, ou óbito, localização (CEP, logradouro, bairro), telefones de contato, raça/etnia, sexo, profissão, idade e nome dos pais;
- c) Variáveis Sensíveis: informações particulares como resultados de exames, diagnósticos, valor de salário.

As variáveis consideradas identificadoras não estarão disponíveis em nenhum *dataset* extraído (nem no dicionário de dados), enquanto as variáveis semi-identificadoras são identificadas no dicionário de dados para definição do nível de agregação para extração do *dataset*. Toda requisição de *dataset* tem por base o plano de dados, o qual indica o recorte de dados de acordo com os dicionários disponibilizados pelo Cidacs/Fiocruz Bahia e as restrições que possam ser aplicadas, dada a natureza dos dados.

Para cada extração, é aplicado um mascaramento de identificadores, com códigos de identificação encriptados, que garante:

- 1) A não integração de dados fora do ambiente seguro;
- 2) A reidentificação dos dados dentro do ambiente seguro (pseudonimização).

Com base na LGPD, o método de anonimização terá como objetivo mapear os riscos de reidentificação segundo o tipo de acesso relacionado a cada *dataset* extraído. Estes riscos são calculados por métodos já reconhecidos em anonimização. As próximas ações incluem: definir uma política de anonimização do centro; anonimizar as bases disponibilizadas e desenvolver um índice de risco de reidentificação nativo do Cidacs/Fiocruz Bahia, para inferir o tipo de acesso que será provido a cada base anonimizada pelo Centro. O processo de anonimização segue o seguinte fluxo:

- 1) Curadoria de Dados aplica supressão de variáveis identificadoras e/ou semi-identificadoras;
- 2) O NPD aplica *hash* nos códigos identificadores;
- 3) Curadoria de Dados aplica generalização em variáveis indicadas pelo pesquisador no plano de dados.

O processo de anonimização para os *datasets* integrados utilizados no projeto da Plataforma Zika baseou-se nas seguintes etapas:

- a) Classificação da sensibilidade das variáveis contidas nas bases integradas para o uso na pesquisa da Zika, com consequente definição de quais variáveis sensíveis necessitam ser anonimizadas;
- b) Definição das estratégias de anonimização;
- c) Aplicação de técnicas de anonimização em cada variável, estimando o limiar de risco. O desafio se coloca no objetivo de desvincular os dados pessoais de seu titular (Privacidade), mantendo a qualidade e utilidade dos dados para análises futuras (Utilidade);
- d) Novo cálculo da medida de risco e então computa-se a medida para avaliar a perda de informação após aplicação das técnicas;
- e) Se o risco for aceitável e a utilidade da base não foi comprometida, a mesma pode ser disponibilizada. Esta etapa tem como objetivo impossibilitar a reidentificação dos indivíduos que integram a base de dados, porém possibilitando a realização da análise para responder ao objetivo do estudo;
- f) Realização de testes iniciais levando em consideração uma base integrada entre RESP, SINASC e SIM: 1) classificação inicial das variáveis constantes nas bases, por sensibilidade destas variáveis; 2) levantamento da completude das informações disponíveis, de modo a identificar quais variáveis possuem entre 50% e 100% de informações faltantes, especialmente entre as variáveis classificadas como sensíveis. Verificou-se que, das 165 variáveis constantes na base, quatro foram classificadas como sensíveis. No entanto, informações sobre raça/cor e outros tipos de exame, que segundo a LGPD seriam classificadas como sensíveis, não foram consideradas como sensíveis, por serem importantes para a pesquisa. Estimou-se um risco inicial alto (>0,90) para a base integrada utilizada para o teste, contendo informações de três origens distintas. Apesar de ter sido identificado um risco de pré-anonimização alto, a base é disponibilizada ao pesquisador em um ambiente seguro e controlado.

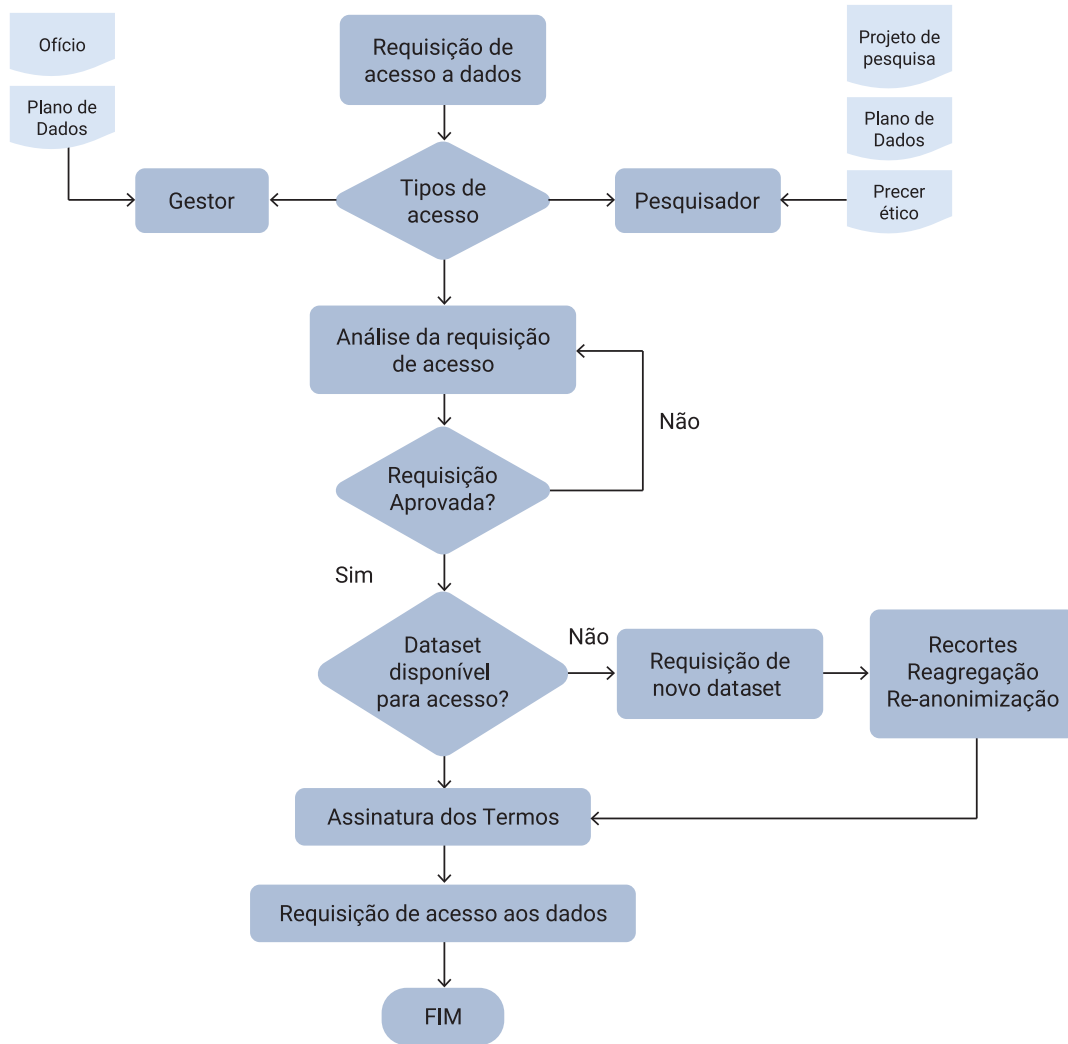
## ACESSO AOS *DATASETS* INTEGRADOS E ANONIMIZADOS

São estes os princípios que orientam o acesso aos *datasets* integrados e anonimizados no Cidacs/Fiocruz Bahia:

- 1) O acesso controlado aos dados é provido aos gestores e pesquisadores autorizados mediante cumprimento dos protocolos de segurança, proteção da privacidade e da confidencialidade do Cidacs/Fiocruz Bahia;
- 2) O acesso controlado aos dados está em conformidade com a legislação pertinente, acordos interinstitucionais e protocolos de segurança da informação;
- 3) O acesso controlado aos dados é garantido mediante a manutenção da responsabilidade legal pelos dados custodiados pela instituição;
- 4) O acesso controlado aos dados será concedido mediante a aprovação do plano de dados e apresentação de parecer ético aprovado por Comites de Ética para responder às questões de pesquisa.

Na Figura 7, encontra-se o detalhamento do fluxo de procedimentos referentes ao tipo de acesso.

Figura 7: Fluxo de acesso aos *datasets* integrados anonimizados para pesquisa em saúde



Fonte: Elaborada pela Curadoria de Dados do Cidacs/Fiocruz Bahia



a) Acesso controlado a *datasets* integrados e anonimizados por pesquisadores

- A pesquisa precisa estar em conformidade com os princípios da finalidade, adequação e necessidade de acordo com a LGPD e dos requisitos éticos do projeto. Assim, para que o(a) pesquisador(a) tenha autorização de acesso controlado aos *datasets* integrados e anonimizados, é impreterível que a pesquisa seja acompanhada de um parecer ético de aprovação no Sistema CEP/ CONEP e da aprovação do plano de dados;
- São gerados códigos distintos para campos-chaves dos *datasets*, com o objetivo de reduzir o risco de reidentificação dos titulares dos dados;
- Os pesquisadores só têm autorização de acesso controlado aos *datasets* integrados e anonimizados, após assinatura do Termo de responsabilidade de acesso e de uso de *datasets* e da Declaração de anuência, para que haja vinculação dos pesquisadores (nacionais e estrangeiros) à pesquisa aprovada no Cidacs/Fiocruz Bahia.

Em termos gerais, o acesso aos *datasets* integrados e anonimizados segue o seguinte fluxo:

- 1) O(a) pesquisador(a) deverá apresentar, para aprovação do Centro de Dados, o projeto, parecer ético favorável e plano de dados para extração de *datasets*;
- 2) Os documentos serão avaliados pela Curadoria de Dados e após validação, o projeto é cadastrado e recebe um código interno de identificação. É por meio deste identificador que ocorre o controle da cadeia de custódia dos dados;
- 3) O pesquisador deve se comprometer ao uso ético e seguro dos dados, que inclui utilizar os dados somente para a finalidade da pesquisa estabelecida, não disponibilizar dados a terceiros e minimizar os riscos de acesso aos dados por pessoas não autorizadas. O pesquisador deve concordar com os termos e

condições de uso designados para o acesso controlado aos *datasets* integrados e anonimizados;

4) Os *datasets* integrados serão extraídos mediante o plano de dados aprovado e os parâmetros de anonimização que envolvem a exclusão de identificadores, semi-identificadores e de potenciais dados sensíveis solicitados;

5) Após a extração do *dataset*, de acordo com o plano de dados, a Curadoria de Dados solicita à Segurança da Informação a transferência do *dataset* para o ambiente de análise, mediante assinatura de Termo de Responsabilidade e uso do *dataset* pelo pesquisador;

6) O acesso controlado aos *datasets* integrados e anonimizados, extraídos para atender ao projeto de pesquisa, será sempre realizado por meio de uma máquina virtual (VM) no ambiente de análise, após a concordância e assinatura dos Termos e Condições de Uso do Data Safe Haven (DSH) - Cidacs/Fiocruz Bahia. Neste ambiente, só é permitido o *download* de resultados de pesquisa (tabelas e gráficos de estatísticas), mediante monitoramento da Curadoria de Dados, em conformidade com as normas de segurança e privacidade de dados do Cidacs/Fiocruz Bahia.

Acesso aos dados originais curados e *datasets* integrados na Plataforma Zika para o gestor:

- A instituição que cede os dados poderá solicitar formalmente as bases de dados originais que tenham sido curadas e tratadas pelo Cidacs/Fiocruz-Bahia para as pesquisas;
- Gestores e profissionais do Ministério da Saúde podem acessar *datasets* integrados e anonimizados utilizados na Plataforma da Zika, por meio de máquina virtual no ambiente de análise, mediante solicitação formal (ofício) com justificativa do estudo, plano de dados e os seguintes itens: identificação (nome

completo); vínculo com a instituição; área técnica/unidade; matrícula Siape (no caso de servidores federais); base(s) de dado(s) solicitada(s), período e recortes; descrição das atividades pretendidas para os dados; finalidade; justificativa e estimativa de tempo para utilização do dado;

- A solicitação é analisada pela Curadoria de Dados, que após sua aprovação, encaminha o plano de dados para que a Produção de Dados identifique o *dataset* integrado extraído. Em seguida, solicita a configuração de máquina virtual no ambiente de análise para a disponibilização dos *datasets*, segundo acordo de acesso;
- Os termos de acesso controlado aos dados integrados e anonimizados (Termo de responsabilidade de acesso e de uso de *datasets*, declaração de Anuência e Termos e Condições de Uso, Termos e declaração de acesso aos dados) devem ser assinados pelos gestores, ou profissionais do Ministério da Saúde autorizados ao acesso.

## **RECURSOS COMPUTACIONAIS PARA O TRATAMENTO DAS BASES DE DADOS E ACESSO DOS PESQUISADORES AOS DATASETS**

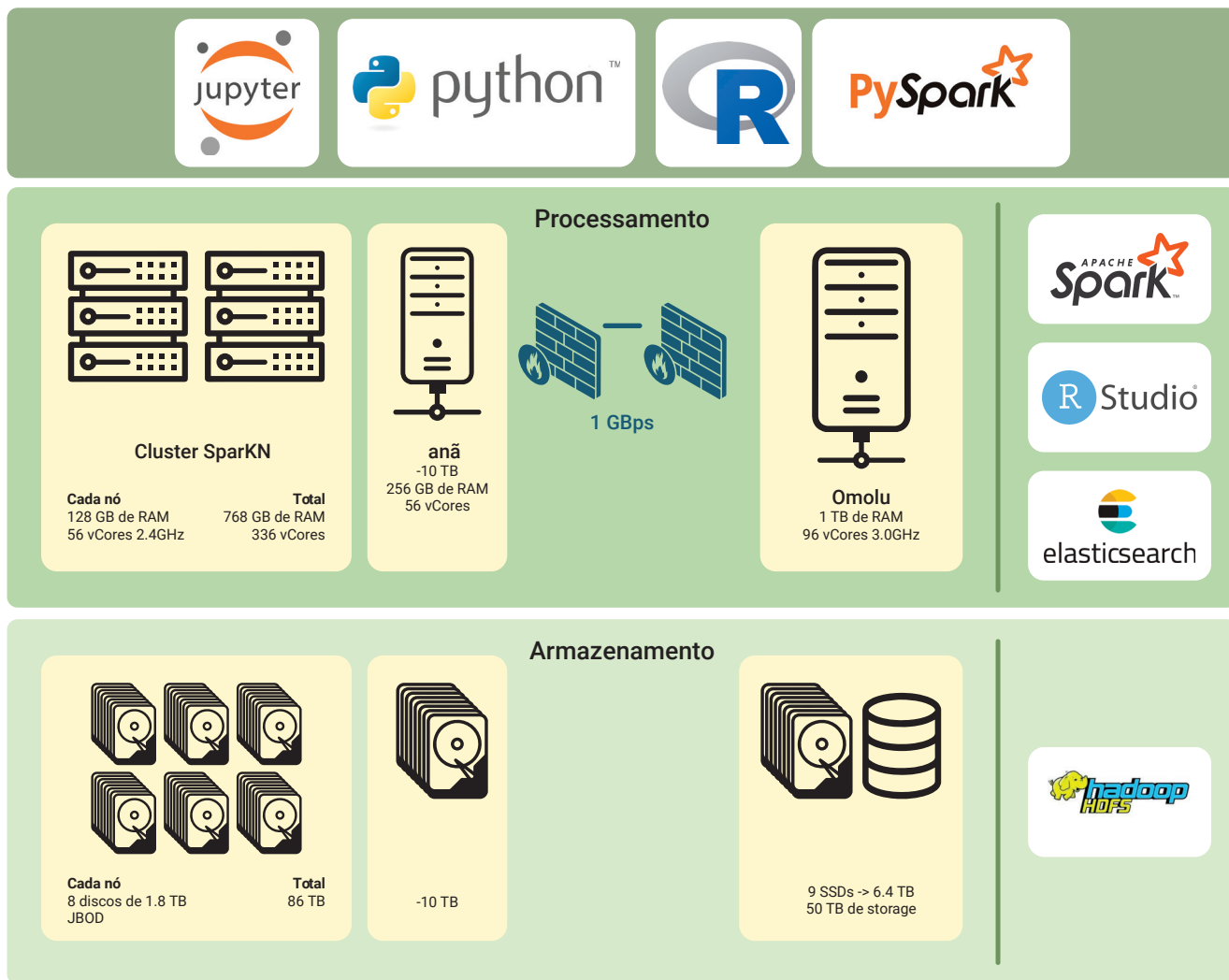
O Cidacs/Ficoruz Bahia opera sua plataforma de dados identificados num ambiente com recursos computacionais de alto desempenho e configurados para atender os marcos regulatórios ético e de segurança e privacidade da informação. Dessa maneira, foram estruturados uma sala segura e um ambiente de análise.

## Sala segura para processamento e integração de dados

É um espaço com acesso restrito e monitoramento permanente realizado pela Segurança da Informação, onde as equipes autorizadas da Produção de Dados e Curadoria de Dados tratam os dados pessoais, após treinamento e assinatura de termos de confidencialidade e de uso e manejo dos dados. O principal objetivo da Sala Segura é a mitigação do vazamento de dados. Por isso, entre outras medidas, todas as áreas internas da Sala Segura possuem controle de acesso duplo (biometria e cartão) e são separadas por portas corta-fogo. Dentre as regras de acesso, está a proibição de entrada de celulares e dispositivos de comunicação, dispositivos de armazenamento como *pen drives* e similares, de todas formas. O monitoramento físico se dá por sistema de Circuito Fechado de Televisão (CFTV), com câmeras cruzadas em todos os ambientes com gravação das imagens.

Para realizar as atividades de tratamento das bases de dados pessoais, foi concebido um *cluster* computacional, contendo soluções (Figura 8) de *software* (e.g. Spark, R Studio, Elasticsearch, Jupyter, Python, PySpark) para atender os processos de ingestão, pré-processamento, *linkage*, acurácia e extração, somente acessíveis localmente pelas equipes autorizadas do Cidacs dentro Sala Segura. Este *cluster* não está conectado na Internet e seu acesso é realizado por rede local, segmentada dos demais ambientes do Cidacs. É conectado por fibra óptica dedicada e ponta-a-ponta a um *cluster* HPC instalado no Centro de Supercomputação do Serviço Nacional de Aprendizagem Industrial - *Campus* Integrado de Manufatura e Tecnologias (Senai-Cimatec), mediante Termo de Cooperação Técnica.

Figura 8: Cluster da Sala Segura



Fonte: Elaborada pelo Cidacs/Fiocruz Bahia

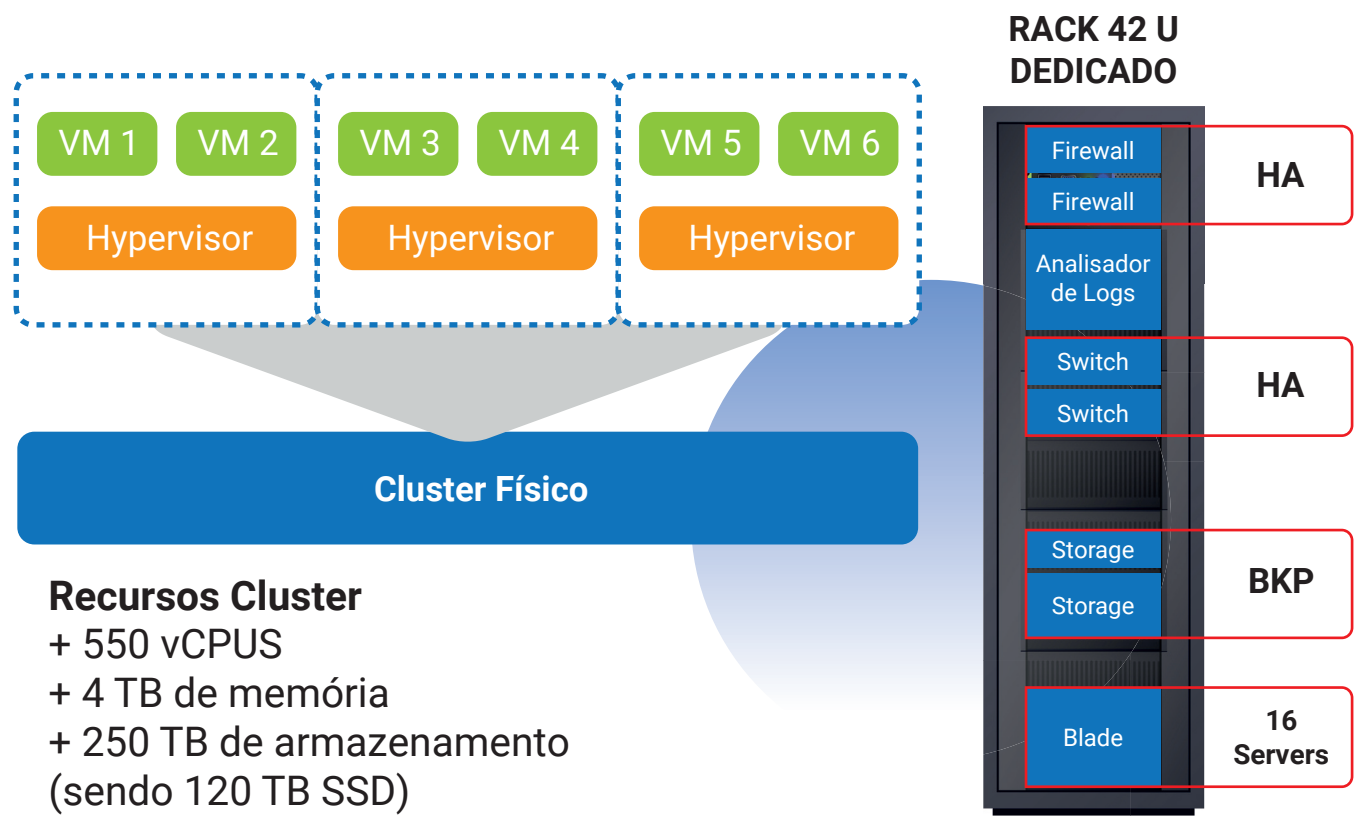
## Ambiente disponível para análise de *datasets* por pesquisadores e gestores

Para o acesso dos pesquisadores a recursos computacionais para analisar os *datasets* extraídos (integrados e anonimizados), foi criado um ambiente controlado, confiável e seguro. Os recursos são configurados para acesso pelos pesquisadores através de VPN (*Virtual Private Network*) com dupla autenticação, com uso de *token* de celular (disponibilizado aos usuários). A transferência de dados (*upload* e *download*) é bloqueada para a área de armazenamento do ambiente. Os recursos para processamento e os *softwares* são previamente instalados e disponibilizados para o usuário em ambiente virtualizado isolado (Figura 9).

O usuário acessa esse ambiente via *desktop* virtual. Os servidores virtuais contam com as ferramentas Python, R/Rstudio e Spark, e bibliotecas adequadas para ciência de dados para *Big Data* são instaladas (ex. tidyverse, dplyr, Pandas, sparklyr e ggplot2). Nesse ambiente, os pesquisadores não têm acesso a Internet, sendo o *input* e *output* dos dados realizado pela equipe de segurança, mediante chamados registrados pela Curadoria de Dados.

A seguir, encontra-se a configuração de máquinas virtuais disponibilizadas para acesso remoto pelos pesquisadores dos *datasets* integrados e anonimizados na Plataforma Zika.

Figura 9: Configuração do ambiente de análise



Fonte: Elaborada pelo Cidacs/Fiocruz Bahia



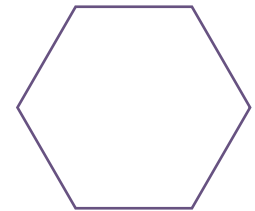
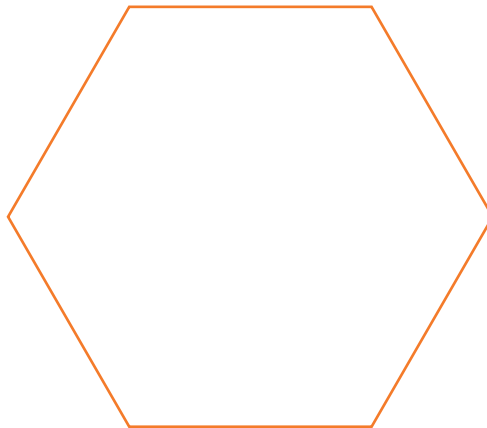
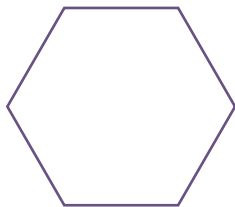
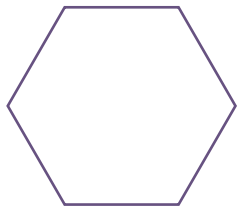
## CONCLUSÕES

Ao ser percebido como um bem da humanidade, o conhecimento científico transpõe o papel de subproduto de uma ação de pesquisa, para ser agente capaz de ampliar conhecimentos e descobertas. Como informam Sayão e Sales (2016), para que isto ocorra, é preciso que eles estejam disponíveis, capazes de serem interpretados, a fim de transmitir conhecimento no tempo e no espaço, e passíveis de reúso, considerando ainda um ambiente interdisciplinar. Isso requer esforço em ações de governança e gestão de dados, cuja execução demanda o estabelecimento de metodologias e instrumentos. São estas ações que serão peças-chaves para evidenciar a transparência dos fluxos do processo de pesquisa, bem como, de sua trajetória de erros e acertos (LEONELLI, 2019). Além disso, acrescidos à noção de bem comum, o usufruto da governança e gestão de dados incorpora os cuidados com a privacidade e a proteção de dados pessoais.

O desafio se torna ampliado quando consideramos que os dados utilizados na pesquisa são dados administrativos, uma vez que apesar das iniciativas para o aprimoramento dos sistemas de informação, resultantes dos esforços e da ampliação do domínio técnico na gestão e produção destes dados, os limites impostos ao manejo dos sistemas de informação impactam em seu uso. Desafios que se postulam quando consideramos a qualidade da informação ante as particularidades distintas presentes entre os diversos sistemas, o preenchimento incompleto de informações e a dificuldade de integração das bases de dados criadas em diferentes estruturas, metodologias e propósitos.

A iniciativa trazida pelo Cidacs/Fiocruz Bahia apresenta um modelo inovador para a realidade brasileira, de governança e gestão de dados. Destaque para o *framework* regulatório relacionado ao acesso, tratamento e uso de dados anonimizados para fins de pesquisa em saúde pública, e a ação de curadoria de dados no estabelecimento e manutenção de gestão de dados por todo o seu ciclo de vida.

Através dele foi possível congregiar as principais bases de dados no domínio temático das infecções causadas por arbovírus em plataforma única; possibilitar aos pesquisadores extração de *datasets* para estudos e pesquisas produzidos no âmbito do projeto Zika; reunir metadados padronizados sobre *datasets* produzidos no âmbito do uso e planejamento da Coorte de Nascimentos para estudos métricos; produzir painéis próprios de indicadores sobre as infecções por arbovírus, bem como, fomentar e integrar redes de acesso aberto a dados e informações.



## ARTIGOS CIENTÍFICOS E PRODUTOS RESULTANTES

BARBOSA, G. C. G. *et. al.* Cidacs-RL: a novel indexing Search and scoring-based record linkage system for huge datasets with high accuracy and scalability. **BMC Medical Informatics and Decision Making**, v. 20, p. 289, 2020. 10.1186/s12911-020-01285-w. Disponível em <https://link.springer.com/article/10.1186/s12911-020-01285-w>. Acesso em novembro de 2022.

BARRETO, M. L; ICHIHARA, M. Y.; ALMEIDA, B. A. *et al.* The Centre for Data and Knowledge Integration for Health (CIDACS): Linking Health and Social Data in Brazil. **International Journal of Population Data Science**. 4:2:04, 2019. Disponível em: <https://orcid.org/0000-0002-0215-4930>. Acesso em novembro de 2022.

CIDACS. **Dataverse**. Disponível em: <<https://dataverse.cidacs.org/>>. Acesso em: novembro de 2022.

## REFERÊNCIAS

ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. **ABNT NBR ISO/IEC 27014: 2013**. Políticas de Segurança da Informação. Rio de Janeiro, 2013.

BIONI, B. R. Compreendendo o conceito de anonimização e dado anonimizado. **Cadernos Jurídicos**, São Paulo, v. 19, p. 191-201, 1 mar. 2020. Disponível em: <https://genjuridico.jusbrasil.com.br/artigos/889500718/compreendendo-o-conceito-de-anonimizacao-e-dado-anonimizado>. Acesso em: 25 nov. 2020.

BRASIL. **Constituição da República Federativa do Brasil de 1988**. Brasília, DF: Presidência da República, 1988. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/constituicao/constituicao.htm](http://www.planalto.gov.br/ccivil_03/constituicao/constituicao.htm). Acesso em: novembro de 2022.

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações previsto no inciso XXXIII do art. 5º, no inciso II do § 3º do art. 37 e no § 2º do art. 216 da Constituição Federal; altera a Lei nº 8.112, de 11 de dezembro de 1990; revoga a Lei nº 11.111, de 5 de maio de 2005, e dispositivos da Lei nº 8.159, de 8 de janeiro de 1991; e

dá outras providências. **Diário Oficial [da] República Federativa do Brasil**. Brasília, DF, 18 de novembro de 2011. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2011-2014/2011/lei/l12527.htm](http://www.planalto.gov.br/ccivil_03/_ato2011-2014/2011/lei/l12527.htm). Acesso em: novembro de 2022.

BRASIL. Lei N° 13.709, de 14 de agosto de 2018 (Lei Geral de Proteção de Dados-LGPD). Dispõe sobre a proteção de dados pessoais e altera a Lei N° 12.965 de 23 de abril de 2014 (Marco Civil da Internet). **Diário Oficial [da] República Federativa do Brasil**. Brasília, DF, 15 ag. 2018. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/\\_ato2015/2018/lei/L13709.htm](https://www.planalto.gov.br/ccivil_03/_ato2015/2018/lei/L13709.htm). Acesso em 20 de outubro de 2019.

BRASIL. Resolução nº 4, de 14 de abril de 2020. Disponibiliza o Guia de Boas Práticas para Implementação da Lei Geral de Proteção de Dados na Administração Pública Federal. **Diário Oficial [da] República Federativa do Brasil**. Brasília, DF, 27 de abril de 2020. Disponível em: <https://www.in.gov.br/en/web/dou/-/resolucao-n-4-de-14-de-abril-de-2020-253999748>. Acesso em: 20 de setembro de 2022.

BRASIL. Resolução CNS nº 466, de 12 de dezembro de 2012. Define diretrizes referentes à não violação do direito à privacidade e ao anonimato das pessoas quando do uso de dados na pesquisa envolvendo seres humanos. **Diário Oficial [da] República Federativa do Brasil**. Brasília, DF, 12 de dezembro de 2012. Disponível em: <https://conselho.saude.gov.br/resolucoes/2012/Reso466.pdf> Acesso em: novembro de 2022.

BARRETO, M.L.; ICHIHARA, M. Y.; PESCARINI, J. *et al.* Cohort Profile: The 100 Million Brazilian Cohort, **International Journal of Epidemiology**, Volume 51, Issue 2, April 2022, Pages e27–e38, <https://doi.org/10.1093/ije/dyab213>

CHEONG, L.K.; CHANG, V. The need for data governance: A case study. In: 18th Australasian Conference on Information Systems. 2007 **ACIS 2007 Proceedings**, Australia.

CHODANKAR, D. Introduction to real-world evidence studies. **Perspectives in Clinical Research**, v. 12, n. 3, p. 171, 2021.

PAIXAO, E.S.; CARDIM L.L.; FALCAO, I.R. *et al.* Cohort Profile: Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS) Birth Cohort, **International Journal of Epidemiology**, Volume 50, Issue 1, February 2021, Pages 37–38, <https://doi.org/10.1093/ije/dyaa255>

LEONELLI, S. Data governance is key to interpretation: Reconceptualizing data in data science. **Harvard Data Science Review**, v. 1, n. 1, p. 10.1162, 2019.

MACKEY, E.; ELLIOT, M.; O'HARA, K. **The anonymisation decision-making framework**. United Kingdom, UKAN Publications, 2016, 119 páginas. Disponível em <https://ukanon.net/framework/>. Acesso em novembro de 2022.

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. **OECD principles and guidelines for access to research data from public funding**. Paris: OECD, 2007. p.24.

PINHO, F. A. S. O. **Anonimização de bases de dados empresariais de acordo com a nova Regulamentação Europeia de Proteção de Dados**. Orientador: Prof. Doutor Manuel Correia, DCC-FCUP. 2017. 94 p. Dissertação (Mestrado em Segurança Informática) - Universidade do Porto, Faculdade de Ciências, Porto, Portugal, 2017. Disponível em: <https://repositorio-aberto.up.pt/handle/10216/105479>. Acesso em: 25 de jul. de 2022.

SAYÃO, L. F.; SALES, L. F. Curadoria digital e dados de pesquisa. **Atoz: novas práticas em informação e conhecimento**, [S.l.], v. 5, n. 2, p. 67-71, dec. 2016. ISSN 2237-826X. Disponível em: <https://revistas.ufpr.br/atoz/article/view/49708>. Acesso em: 07 feb. 2023. doi: <http://dx.doi.org/10.5380/atoz.v5i2.49708>

UNIÃO EUROPEIA. Regulamento (UE) 2016/679 do Parlamento Europeu e do Conselho de 27 de abril de 2016. Relativo à proteção das pessoas singulares no que diz respeito ao tratamento de dados pessoais e à livre circulação desses dados e que revoga a Diretiva 95/46/CE. **Jornal Oficial da União Europeia**. Bruxelas, 27 de abril de 2016. Disponível em: <https://eur-lex.europa.eu/legal-content/PT/TXT/PDF/?uri=CELEX:32016R0679&from=EN>. Acesso em 06 de fevereiro de 2023.

VAN DER AALST, W. M. P.; BICHLER, M.; e HEINZL, A. Responsible data science. **Bus Inf. Syst. Eng.** v. 59, n.5, p.311–313, 2017. Disponível em: <https://doi.org/10.1007/s12599-017-0487-z>. Acesso em 20 de outubro de 2020.

WILKINSON, M. D. *et al.* The FAIR Guiding Principles for scientific data management and stewardship. **Nature Sci. Data** v:3, 2016. Disponível em: <https://www.nature.com/articles/sdata201618>. Acesso em 20 de outubro de 2020.



# Vigilância Longitudinal da Zika e da Síndrome Congênita do Zika

---

*Maria Glória Teixeira  
Eduardo Hage Carmo  
Enny Santos Paixão  
Eslaine Santos e Santos  
Giovanny Vinícius Araújo de França  
Júlia Moreira Pescarini  
Juliane Fonseca Oliveira  
Larissa Catarina Costa  
Laura Cunha Rodrigues  
Lavínia Schuler-Faccini  
Luciana Lobato Cardim  
Maria da Conceição Nascimento Costa  
Maurício Lima Barreto  
Moreno Magalhães de Souza Rodrigues  
Qeren Hapuk Rodrigues Ferreira Fernandes  
Rafael Valente Veiga  
Rita de Cassia Carvalho-Sauer  
Roberto Fernandes Silva Andrade  
Wanderson Kleber Oliveira*

As informações disponíveis sobre a doença causada pelo vírus Zika (ZIKV) até o ano de 2013 eram bastante limitadas e, praticamente, inexistiam estudos epidemiológicos sobre esta arbovirose, com exceção de alguns referentes a uma pequena ilha (Yap) no Oceano Índico, onde ocorreu um surto autolimitado em 2007. Assim, até aquele momento, a maioria das publicações sobre esta virose referia-se a relatos de casos e a um reduzido número de inquéritos sorológicos realizados na Ásia e África (FOY, B.; KOBYLINSKI; FOY, J.; BLITVICH *et al.*, 2011; HEANG; YASUDA; SOVANN; HADDOW *et al.*, 2012). Só em 2013/2014, a partir da epidemia ocorrida na Polinésia Francesa, conhecimentos científicos mais robustos foram gerados quando, além do quadro clínico da fase aguda caracterizada como doença febril autolimitada, acompanhada de exantema, dores e edema articular, foram também observados vários casos de Síndrome de Guillain Barré, após a remissão daquela fase desta virose (PAIXÃO; BARRETO; TEIXEIRA; COSTA; RODRIGUES, 2016).

No Brasil, a emergência do vírus Zika (ZIKV) demorou para ser reconhecida de modo que, inicialmente, houve dificuldades para instituir uma vigilância epidemiológica universal, em função, tanto do desconhecimento de que a doença exantemática que estava ocorrendo em cidades do Nordeste, desde 2014, era produzida pelo ZIKV, como também pela semelhança do seu quadro clínico, na fase aguda, com duas outras arboviroses urbanas (dengue e chikungunya) presentes na grande maioria dos municípios onde o ZIKV passou a circular. Contudo, a partir de novembro de 2015, imediatamente após a detecção do aglomerado de casos de microcefalia em recém-nascidos (RN) relacionada à transmissão vertical do ZIKV, as notificações e investigações de casos de Zika e desta anomalia congênita passaram a ser, sistematicamente, registradas neste país.

Em 11 de novembro de 2015, o Brasil declarou Emergência em Saúde Pública de Importância Nacional (Espin) devido a mudança no padrão de ocorrência de microcefalia ao nascer do aglomerado de bebês nascidos com microcefalia (BRASIL, 2015), ao tempo em que evidências científicas sobre a existência de associação causal entre o ZIKV e aquelas anomalias foram sendo rapidamente produzidas (RASMUSSEN; JAMIESON; HONEIN; PETERSEN, 2016; TEIXEIRA; COSTA; OLIVEIRA; NUNES; RODRIGUES, 2016). Grandes preocupações e desafios foram gerados por este problema para as autorida-



des de saúde e comunidades científicas nacionais e internacionais, ao lado do rápido crescimento de casos e da sua ampla distribuição espacial, inclusive com registros em outros países das Américas. Isso levou a Organização Mundial de Saúde (OMS) a declarar em 1º de fevereiro de 2016, Emergência em Saúde Pública de Importância Internacional (ESPII) pela quarta vez no âmbito do Regulamento Sanitário Internacional - RSI/2005 (BRASIL, 2009), visto que o risco de continuidade da circulação do ZIKV era muito alto, dada a grande disseminação e infestação das cidades brasileiras, de outros países do continente americano e de outras regiões do mundo, por mosquitos do gênero *Aedes*, principalmente o *Aedes aegypti*.

O Ministério da Saúde do Brasil (MS), ainda em 2015, instituiu um Plano de Emergência para este inusitado evento de saúde pública, sob a Coordenação da Secretaria de Vigilância em Saúde (SVS/MS), que incluiu Manuais, Protocolos e ações voltados para mitigação do referido problema (BRASIL, 2017). Estas iniciativas envolviam várias dimensões da atenção à saúde, tais como garantia de rede adequada de assistência às crianças acometidas; fortalecimento das atividades do pré-natal; comunicação e desenvolvimento de ações de educação em saúde para risco de infecção pelo ZIKV, especialmente, em gestantes; estabelecimento de vigilância ativa da Zika e da Microcefalia/Síndrome Congênita associada à infecção pelo vírus Zika, intensificação das ações de prevenção e controle, dentre outras (BRASIL, 2017). Os conhecimentos sobre esta Síndrome Congênita foram sendo produzidos no curso destas duas epidemias, o que possibilitou ampliar a compreensão científica sobre aquela nova síndrome congênita, tanto pela rede de serviços de saúde, como pela mobilização da comunidade científica nacional articulada com algumas instituições internacionais. Nesse ínterim, a SVS disponibilizou recursos para apoiar o desenvolvimento de estudos e pesquisas voltados para este tema.

Apesar de o aglomerado de casos de microcefalia em recém nascidos ter sido o sinal de alerta da existência de uma epidemia de anomalias congênitas, produzida por agente infeccioso, resultando em graves lesões no Sistema Nervoso Central (SNC), imediatamente foi levantada a hipótese de que a ocorrência da infecção pelo ZIKV no conceito poderia não apresentar sinais clínicos ao nascer, ou seja, sem a presença de microcefalia. A hipótese era de que, assim como em outras infecções congênitas, as

manifestações deveriam fazer parte de um largo espectro de lesões, o que motivou a SVS a financiar estudos de longo prazo, capazes de detectar e acompanhar desfechos adversos nas crianças, diagnosticadas, ou não, com microcefalia logo ao nascer, a exemplo de mortalidade precoce, transtornos neurológicos com comprometimento do desenvolvimento físico e cognitivo no decorrer dos primeiros anos de vida, além de outros transtornos que só poderiam ser identificados mais tardiamente, como déficit de aprendizagem escolar.

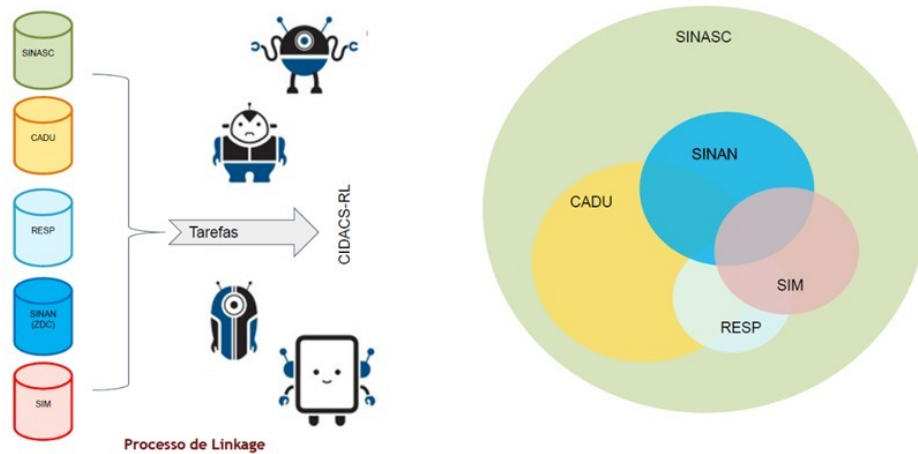
Estes pressupostos levaram à definição dos objetivos e escopo da Meta quatro do projeto denominado “Plataforma de vigilância de longo prazo para a Zika e Microcefalia no âmbito do SUS”, que passou a ser desenvolvida pela equipe de pesquisadores do Centro de Integração de Dados e Conhecimentos para a Saúde (Cidacs/Fiocruz Bahia), em diálogo constante com a equipe técnica da coordenação central do Sistema Único de Saúde (SUS), por meio da SVS/MS. Assim, foi definido como objetivo principal dessa Meta, construir uma plataforma integrada de diferentes bases de dados sociais (saúde, educação, previdência e desenvolvimento social), conforme apresentado no capítulo 1, que possibilitasse desenvolver estudos e pesquisas longitudinais de longo prazo para gerar conhecimentos sobre o espectro da Síndrome Congênita do Zika (SCZ), seu impacto epidemiológico e nas condições de vida dos indivíduos e populações acometidas por microcefalia/SCZ e pela infecção pelo vírus Zika.

## ESCOPO DO PROJETO

O avanço das estratégias que o campo da ciência de dados vem proporcionando, especialmente por meio da vinculação de grandes bases de dados (*bigdata*) individuais (Figura 1), colocou a possibilidade de se desenvolver esta plataforma de estudos sobre alguns aspectos da epidemia de Zika e da SCZ ao longo do tempo, utilizando dados administrativos coletados de forma sistemática e contínua pela rede de serviços de saúde do SUS. Ademais, tem-se em perspectiva incluir nesta plataforma integrada, bancos de dados de outras áreas e instituições como Previdência Social, Educação e Desenvolvimento Social, de modo a possibilitar a condução de estudos e pesquisas longitudinais de longo prazo, para além do campo da saúde.

Figura 1: Bancos de dados da Plataforma Zika

## Plataforma Zika – Bancos de Dados



Fonte: Elaborada pelos autores. SINASC: Sistema de Informações sobre Nascidos Vivos. Cadu: Cadastro Único. RESP: Registro de Eventos em Saúde Pública. SINAN: Sistema de Informações de Agravos de Notificação. SIM: Sistema de Informações de Mortalidade.

Ao lado dos estudos genômicos e daqueles gerados pelos bancos de dados linkados, foram também conduzidos no bojo desta Meta, importantes estudos de agregados espaciais e temporais para responder questões técnico-científicas colocadas pela equipe técnica da SVS. Foi ainda avaliado o desempenho do Sistema de Informação sobre Nascidos Vivos (SINASC) no que se refere ao registro de anomalias congênitas, antes e após as mudanças realizadas no curso da epidemia de SCZ. Dois produtos tecnológicos foram desenvolvidos para contribuir e facilitar a operacionalização de algumas atividades de Vigilância em Saúde, a exemplo da rápida elaboração de análises de dados e automatização da complexa classificação de casos de microcefalia/SCZ registrados no Registro de Eventos em Saúde Pública (RESP).

Neste capítulo apresentaremos os principais resultados dos estudos concluídos e publicados em periódicos científicos, a grande maioria de livre acesso, e destes, dois produtos tecnológicos acima referidos, os quais foram testados e validados e a seguir disponibilizados para uso da rede de serviços de saúde. Por sua vez, a Coorte de Nas-

cimentos construída no Cidacs/ Fiocruz Bahia no bojo desta Plataforma, que envolve, atualmente, 28.631.394 Nascidos Vivos (NV) no Brasil, no período de 1º janeiro de 2001 a 31 de dezembro de 2018, está disponível para uso da comunidade científica em acordo com os preceitos éticos e de proteção de dados vigentes no Brasil.

Importante destacar que, uma das políticas adotadas pela coordenação deste projeto, foi propiciar a formação de uma equipe multidisciplinar envolvendo pesquisadores do Cidacs/Fiocruz Bahia e da Universidade Federal da Bahia (UFBA). Além disso, buscou manter permanente articulação e colaboração com grupos e/ou pesquisadores de outras instituições, bem como, com técnicos e dirigentes da SVS, no propósito de aliar saberes no curso dos processos de investigação, em acordo com as *expertises* necessárias a cada estudo, conforme pode ser observado nas autorias dos artigos publicados.

## ZIKA NO BRASIL

### **Novos *insights* sobre a emergência do vírus Zika nas Américas e reconstrução espaço-temporal da dinâmica epidêmica no Brasil<sup>1</sup>**

Desde 2016, tinha-se conhecimento que o Brasil foi o ponto de introdução do ZIKV no continente americano, que ocorreu entre final de 2013 e início de 2014, conforme indicaram as sequências do genoma deste vírus, cuja origem mais provável foi a Polinésia Francesa (linhagem Asiática) (FARIA; AZEVEDO; KRAEMER; SOUZA *et al.*, 2016; FARIA; QUICK; CLARO; THÉZÉ *et al.* 2017). Em 2015, este agente se disseminou para outros países da América do Sul, Central e Caribe (METSKY; MATRANGA; WOHL; SCHAFFNER *et al.* 2017; THÉZÉ; LI; DU PLESSIS; BOUQUET *et al.* 2018), chegando aos EUA em 2016 (provavelmente vindo do Caribe) (GRUBAUGH; LADNER; KRAEMER; DUDAS *et al.*, 2017). Alguns estudos indicaram os padrões de dispersão pelo continente americano, por meio da história evolutiva do ZIKV, contudo, pouco se sabia sobre a sua difusão no Brasil, epicentro do *cluster* mundial de casos de SCZ (WHO, 2016; TEIXEIRA; COSTA, OLIVEIRA; NUNES; RODRIGUES, 2016).

---

1. COSTA; VEIGA; OLIVEIRA; RODRIGUES *et al.* 2021.

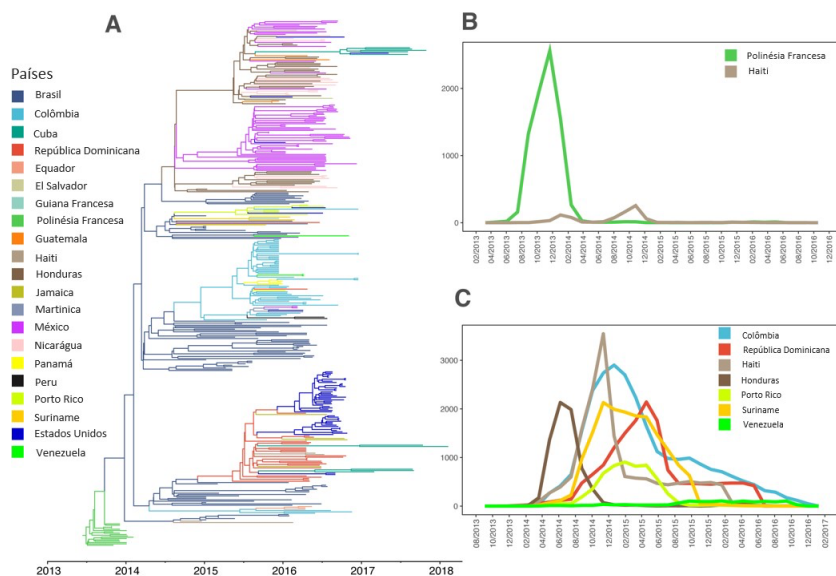
Assim, algumas dúvidas permaneceram sobre a proveniência e momento de emergência, difusão e circulação do ZIKV nas Américas. Então, era imprescindível investigar esta questão, dado que os aportes de conhecimentos genômicos podem trazer novos *insights* sobre a trajetória dos agentes infecciosos, importante informação para entendimento da epidemiologia de novas doenças. Por esta razão, foi desenvolvido este estudo mais abrangente, baseado em um conjunto atualizado de genomas completos e quase completos de ZIKV, oriundos do Caribe e das Américas Central, do Norte e do Sul, com foco especial no Brasil. Para tal, aplicou-se métodos de relógio molecular e filogeografia, para melhor compreender o surgimento do ZIKV nas Américas, com especial ênfase no início da sua disseminação no Brasil, sendo constatado que o Estado de Pernambuco atuou como ponto de partida e polo de disseminação deste vírus pelo país.

O estudo utilizou 431 genomas completos, ou quase completos, de ZIKV do Caribe e das Américas Central, do Norte e do Sul e Polinésia Francesa, disponíveis no GenBank até janeiro de 2019, dos quais 119 eram procedentes de 12 estados das regiões Centro-Oeste, Norte, Sudeste e Nordeste do Brasil. Os resultados encontrados confirmaram que o ZIKV chegou ao Brasil proveniente da Polinésia Francesa, entre novembro de 2013 a março de 2014, ou seja, antes da sua primeira identificação laboratorial em humanos, bem como, dos possíveis casos de Zika detectados no final de 2014 e início de 2015, que estavam sendo notificados pelos serviços de saúde como uma doença exantemática a esclarecer (Figura 2).

No período entre maio e junho de 2013, o Brasil sediou os jogos da Copa das Confederações e, dentre as seleções participantes, a seleção do Taiti, maior Ilha da Polinésia Francesa, jogou em Recife/Pernambuco. Com base nos resultados obtidos no estudo, foi possível aventar a hipótese de que este agente tenha sido introduzido nessa cidade durante o período dessa competição. Além disso, constatou-se que a partir do Brasil o ZIKV se disseminou para diversos países das Américas e do Caribe, como Haiti, em meados de 2014, e no final deste mesmo ano, para Colômbia, República Dominicana, Honduras, Porto Rico, Suriname e Venezuela (Figura 2).

Na Figura 2 observa-se as análises filogenéticas espaciais e temporais do ZIKV nas Américas: (A) Árvore de máxima credibilidade de clado (MCC) utilizando genomas de ZIKV apresentando introdução nas Américas (Central, Norte e Sul) e no Caribe. (B) Transições virais suportadas (BF > 3,0) ao longo do tempo de linhagens de ZIKV com destino ao Brasil de acordo com análises de Markov Jumps; (C) Transições virais suportadas (BF > 3,0) ao longo do tempo de linhagens de ZIKV do Brasil para outros países.

Figura 2: Estimativa de tempo da disseminação do vírus no Brasil no período de 2014 a 2017

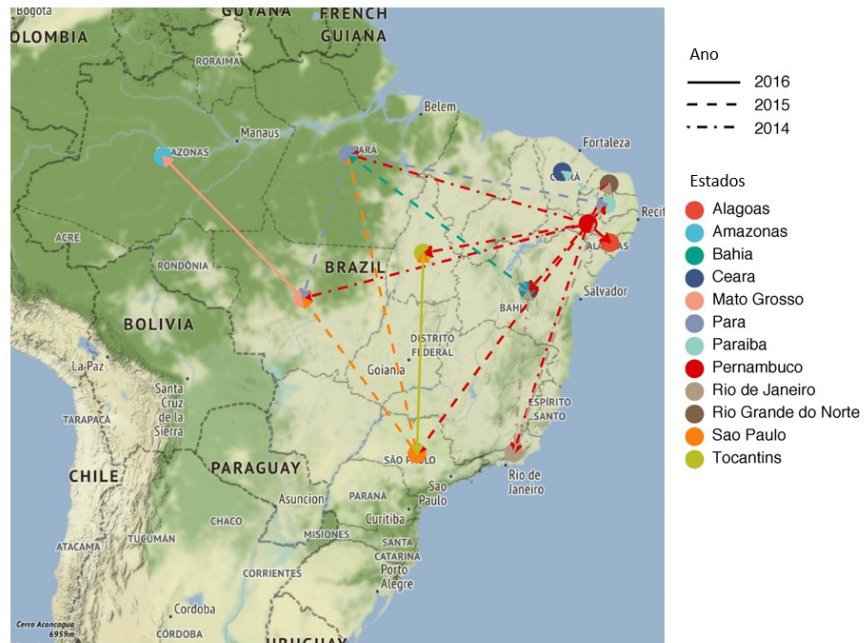


Fonte: Elaborada pelos autores

Como referido, Recife/Pernambuco provavelmente foi o ponto de introdução do ZIKV no Brasil em 2013, sendo ponto de partida para a disseminação para outras cidades de outros estados brasileiros (Alagoas, Bahia, Mato Grosso, Pará, Paraíba, Rio de Janeiro, Rio Grande do Norte, São Paulo e Tocantins), entre meados de 2014 e final de 2016, com maior intensidade de disseminação em meados de 2015 (Figura 3).



Figura 3: Trajetória de disseminação do ZIKV entre 2014-2016 no Brasil



Fonte: Elaborada pelos autores

Embora o ZIKV tenha sido isolado pela primeira vez no Brasil de um paciente oriundo de Camaçari/ Bahia, a Vigilância Epidemiológica (VE) havia observado que os primeiros casos daquela Doença Exantemática a Esclarecer, que posteriormente foi confirmada como Zika, haviam ocorrido em Recife e Natal, evidenciando a capacidade de rápida detecção de novas doenças do sistema de VE no Brasil. É importante salientar que os resultados deste estudo evidenciam que os avanços técnicos e metodológicos na área de estudos genômicos devem ser aplicados à vigilância viral de forma mais sistemática, por trazerem importantes subsídios para a vigilância de doenças. A filogenética e filogeografia têm o potencial de elucidar a origem, o impacto das mutações e a trajetória dos agentes infecciosos, especialmente aqueles emergentes. Essas informações podem ser úteis para apoiar o desenvolvimento de modelos de previsão e a definição de possíveis estratégias de controle.



## **Interdependência entre casos confirmados e descartados de dengue, chikungunya e Zika vírus no Brasil: uma análise multivariada de séries temporais<sup>2</sup>**

A cocirculação de três diferentes arbovírus (DENV, CHIKV, ZIKV) que se estabeleceu em áreas urbanas no Brasil a partir de 2013, representou um importante problema para a saúde pública, que impôs a instituição de uma vigilância ativa para acompanhamento destas arboviroses. Contudo, apesar das iniciativas adotadas pela SVS para implantar a notificação caso a caso, por ser um dos principais requisitos dos sistemas para Vigilância Epidemiológica universal, constatou-se ser muito difícil para os serviços de saúde realizarem o diagnóstico diferencial de cada uma dessas arboviroses, de forma mais precisa, pelo fato de que os sinais e sintomas clínicos apresentados pelos pacientes infectados por estes agentes infecciosos são bastante semelhantes, e a disponibilidade de exames laboratoriais específicos era muito limitada.

A equipe técnica da SVSA percebeu que as séries temporais de notificação de dengue do país como um todo, estavam com excesso de casos descartados quando comparados aos períodos anteriores, gerando a hipótese de que, possivelmente, casos de Zika e/ou chikungunya estavam sendo notificados como suspeitos de dengue, e após investigação epidemiológica eram descartados por serem inconclusivos. Visando elucidar esta questão, foi desenvolvido um estudo com o objetivo de identificar e quantificar a interdependência mútua das séries de casos notificados e descartados das três arboviroses, lançando-se mão de métodos estatísticos e matemáticos mais sofisticados.

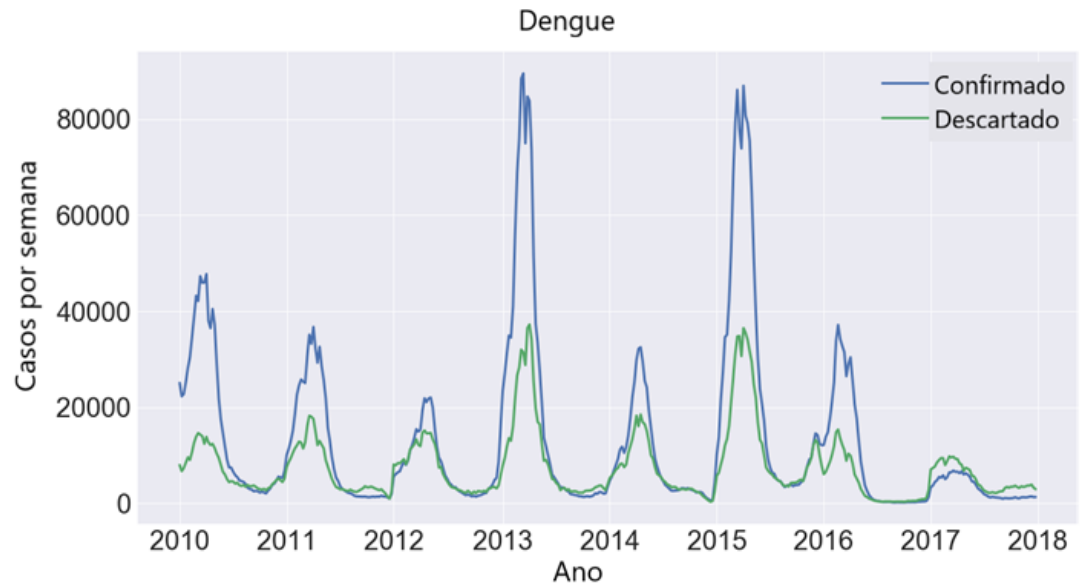
Especificamente, empregou-se um modelo de Vetor Auto-regressivo (VAR), que foi capaz de identificar e caracterizar correlações entre as seis séries de casos confirmados e descartados de dengue, chikungunya e Zika. Em seguida, aplicou-se o Teste de Causalidade de Granger entre cada par de séries, permitindo aferir até que ponto os dados de uma série de casos de uma das arboviroses estaria afetando a série de outra. As bases de dados destas arboviroses registradas no Sistema de Informação de Agravos de Notificação (SINAN), de 2010 a 2018, revelaram que as duas maiores ondas epidê-

---

2. OLIVEIRA; RODRIGUES; SKALINSKI; SANTOS *et al.* 2020.

micas de dengue nesse período ocorreram em 2013 e em 2015, e que a partir de 2013 houve um aumento na proporção de casos descartados dessa virose (Figura 4).

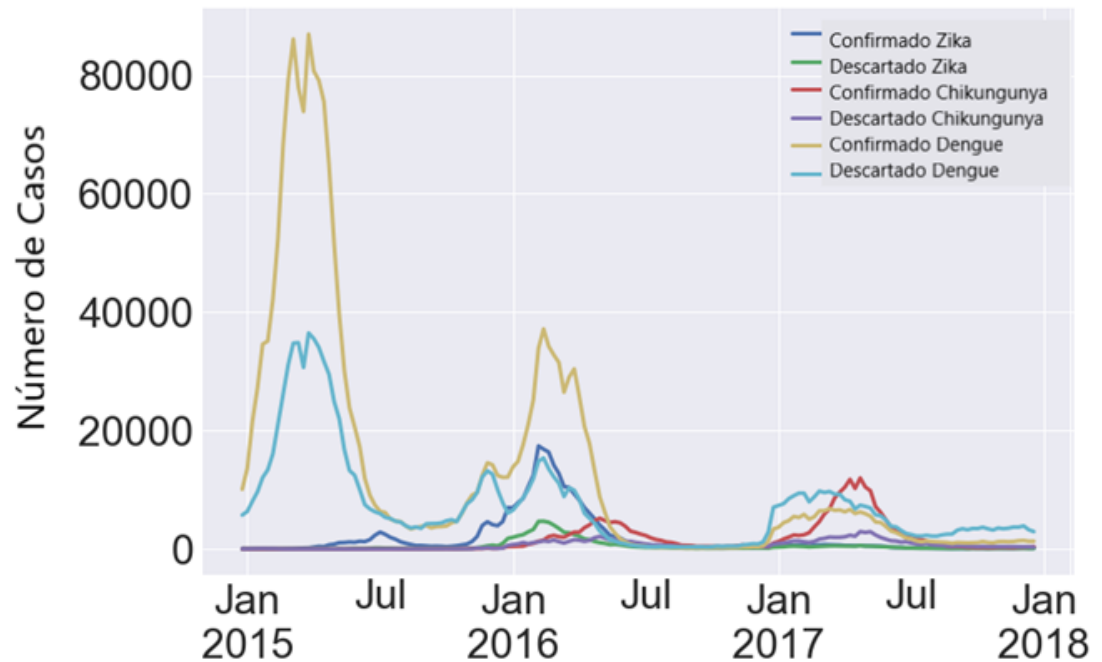
Figura 4: Séries temporais da incidência de casos confirmados e descartados de dengue por semana epidemiológica/ano no Brasil, de 2010 a 2018



Fonte: Sistema de Informação de Agravos de Notificação

A série de casos notificados de chikungunya, que teve início em 2014 com alguns poucos casos, foi seguida por três ondas de 2015 a 2017, sendo a maior magnitude de incidência observada no final de 2016. Já as notificações de Zika tiveram início em 2015, com maior pico nos primeiros meses de 2016 (Figura 5).

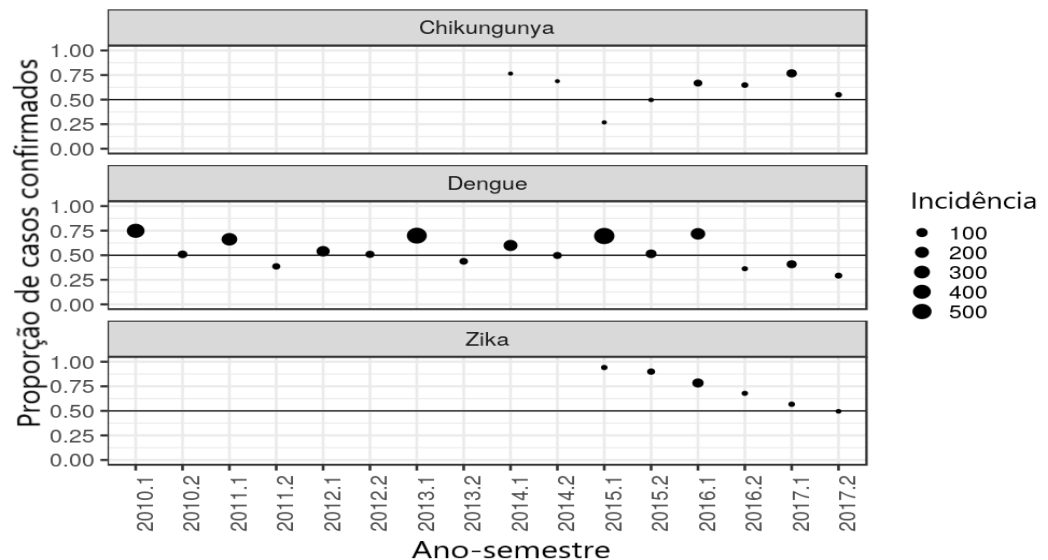
Figura 5: Número de casos confirmados e descartados de Zika, chikungunya e dengue por ano/semestre. Brasil, 2015 a 2018



Fonte: Sistema de Informação de Agravos de Notificação

A Figura 6 mostra a proporção de casos confirmados das três arboviroses, por semestre de 2015 a 2017, dentro o total de casos notificados, onde se observa que os padrões da distribuição da proporção entre casos confirmados e descartados diferem entre essas doenças. Para dengue, a proporção de casos descartados foi maior durante o segundo semestre do ano, quando, em geral, ocorre menor número de casos. Esse padrão, também é observado nos anos com menor incidência desta doença (2012, 2014 e 2017).

Figura 6: Proporção e incidência de casos confirmados de chikungunya, Dengue e Zika, por ano/semestre. Brasil, 2010 a 2017



Fonte: Sistema de Informação de Agravos de Notificação

Os resultados obtidos com a aplicação do Teste de Causalidade de Granger permitiram mostrar que a série de casos de Zika teve um impacto significativo sobre a série de casos de dengue e vice-versa, sugerindo que vários casos de dengue descartados e confirmados podem realmente ter sido casos de Zika. O modelo também sugere que a série de casos confirmados de chikungunya é quase independente das séries de casos de dengue e Zika.

Argumentamos que, ao lado da vigilância de casos baseada nos sinais e sintomas dos pacientes sem confirmação laboratorial, o uso sistemático de métodos similares adotados aqui pelas equipes que analisam essas arboviroses urbanas, poderão auxiliar na redução de erros inerentes às notificações de casos por critério clínico-epidemiológico, e fornecer uma melhor aproximação da magnitude de ocorrência de cada uma dessas viroses. Esses achados evidenciam os desafios relacionados aos vieses de notificação e lançam uma nova luz sobre como lidar com este problema, mesmo que só

se disponha de diagnósticos de casos com base em critérios clínico-epidemiológicos, em situação de circulação simultânea de arbovírus que se expressam clinicamente de forma similar.

## **ANOMALIAS CONGÊNITAS E SÍNDROME CONGÊNITA DO ZIKA**

### **Avaliação do impacto da epidemia de Zika no registro de anomalias congênitas no Brasil: uma análise de séries temporais interrompidas<sup>3</sup>**

Após a identificação do agregado de casos de microcefalia em 2015, ficou evidente que o instrumento de registro de dados do SINASC no que se referia às anomalias congênitas era incompleto e que estas condições eram sub-registradas. Visando aperfeiçoar esse instrumento, o Ministério da Saúde/SVSA realizou algumas alterações com vistas à superação deste problema, razão pela qual a equipe de pesquisadores desta plataforma, ao lado da equipe técnica da SVSA, considerou importante avaliar o impacto da epidemia de Zika no registro de defeitos congênitos no Brasil.

Este estudo foi desenvolvido a partir de uma análise de séries temporais interrompidas, visando identificar mudanças nas tendências do registro das anomalias congênitas relacionadas e não relacionadas à SCZ. Foram utilizados dados mensais do SINASC referentes a: 1) taxa de anomalia congênita do cérebro e do olho, provavelmente ZIKV 2) taxa de anomalia congênita não relacionada ao cérebro, ou olho, e não ZIKV. Como o período entre a infecção materna pelo ZIKV e o diagnóstico da anomalia congênita atribuível à infecção é de cerca de seis meses, considerando que a infecção ocorreu no primeiro trimestre, utilizamos o mês de setembro de 2015 como ponto de interrupção na série temporal, ou seja, seis meses após o início do expressivo aumento de casos de Zika no país (março de 2015).

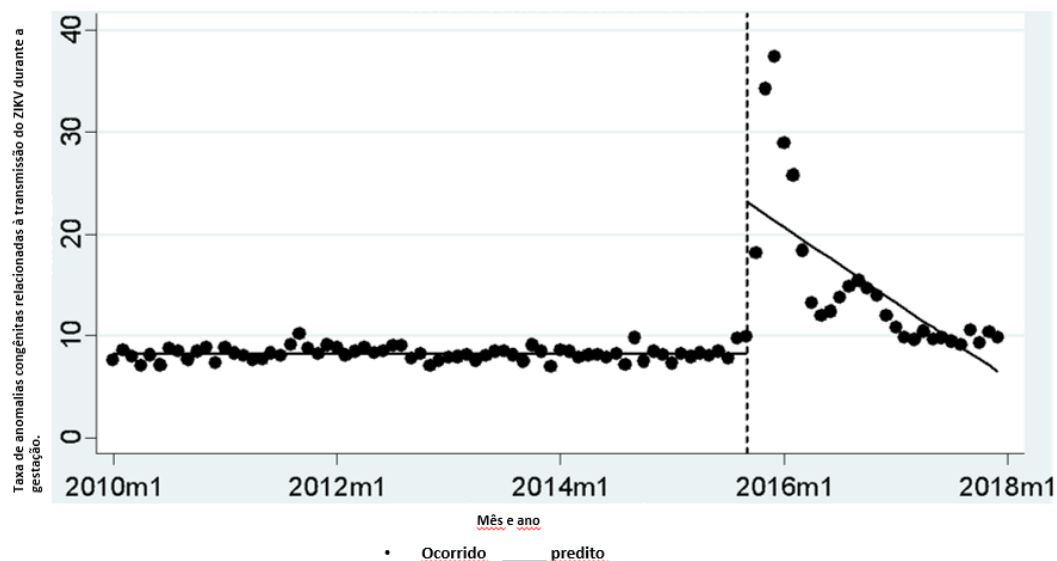
---

3. Para ver o artigo publicado sobre essa análise, consultar PAIXÃO; RODRIGUES; CARDIM; OLIVEIRA *et al.*, 2019.

Para fins desta análise, consideramos o período de janeiro de 2010 a setembro de 2015 como “evento pré-Zika” e o período após setembro de 2015 a dezembro de 2017 como “evento pós-Zika”. As notificações de anomalias congênitas do cérebro e olho estavam apresentando uma tendência de decréscimo no *baseline*, sem significância. Imediatamente após o ponto de interrupção, houve um grande aumento na taxa de notificação de anomalias congênitas de 14,9/10.000 nascidos vivos no grupo cérebro e olho, e de 5,2/10.000 nascidos vivos no grupo não relacionado com anomalias cerebrais ou oculares. Esse aumento na notificação ocorreu em todas as regiões do país (exceto no Sul) e, principalmente, no Nordeste. No período “evento pós-Zika”, diferentemente do grupo cérebro e olhos que apresentou queda mensal, o grupo sem malformações cerebrais ou oculares apresentou um aumento lento, mas significativo (em relação à tendência pré-Zika) de 0,2/10.000 nascidos vivos.

Esta avaliação revelou que houve um avanço no registro de anomalias ao nascimento, incluindo as anomalias não atribuídas ao ZIKV, durante e após a epidemia de Zika, evidenciando que o aperfeiçoamento do instrumento de registro de informações do SINASC, ao lado da sensibilização dos profissionais que atuam na rede de serviços de atenção ao parto, tiveram impacto positivo nos registros das anomalias congênitas.

Figura 7: Série temporal de anomalias congênitas cerebrais e oculares, Brasil- 2010 a 2017



### Tendência temporal da prevalência e mortalidade infantil das anomalias congênitas no Brasil, de 2001 a 2018<sup>4</sup>

As anomalias congênitas (AC) são definidas pela Organização Mundial da Saúde (OMS) como defeitos estruturais, ou funcionais, que ocorrem antes do nascimento. Estima-se que, mundialmente, em média 3 a 6% dos recém-nascidos apresentam algum tipo de anomalia, sendo, portanto, uma das principais causas da carga global de doenças (WHO, [s.d.]; CDC, 2008; SHERIDAN; WRIGHT; SMALL; CORRY *et. al*, 2013) . No Brasil, as AC ocupam a segunda posição dentre os principais grupos de causas de óbito infantil. Ainda assim, não existe uma vigilância ativa das AC no país e as informações disponíveis são provenientes dos registros obtidos nas Declarações de Nascidos Vivos (DNV), fonte de registro de dados do SINASC, cujo campo para registro de AC era

4. Para ver o artigo publicado sobre essa análise, consultar FERNANDES; PAIXÃO; COSTA; TEIXEIRA, 2023.

considerado incompleto até 2015, tendo sido aperfeiçoadas as orientações relativas ao seu preenchimento e captação no curso da epidemia da microcefalia/SCZ detectada no final daquele ano (PAIXAO; RODRIGUES; CARDIM; OLIVEIRA *et. al*, 2019; BRASIL, 2021).

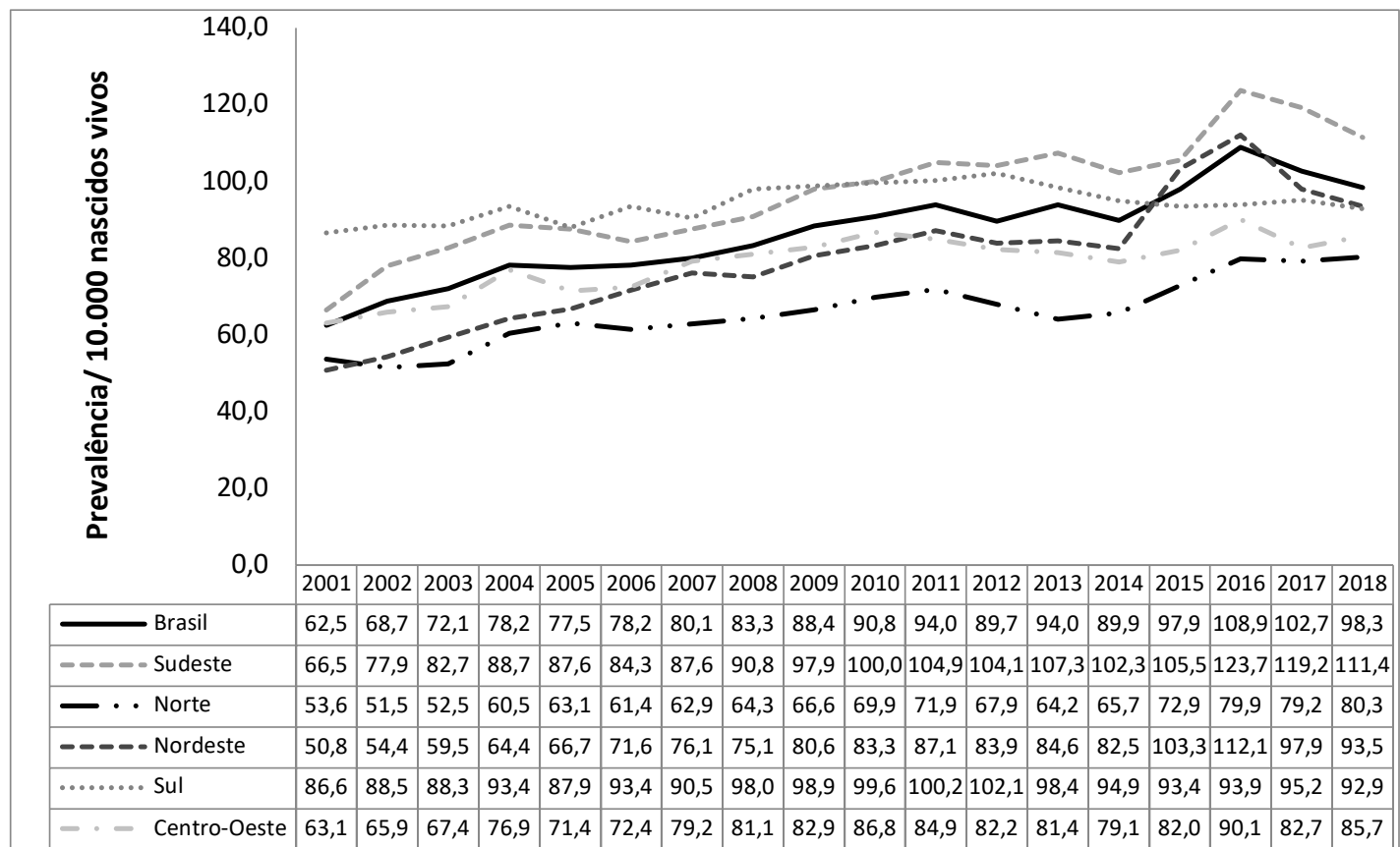
Visando contribuir para ampliar as informações sobre a ocorrência destas anomalias - eventos de saúde de baixa magnitude quando se considera os nascimentos como um todo, este estudo utilizou a estratégia *linkage*, a partir da vinculação das bases de dados de nascidos vivos (SINASC) e óbitos (Sistema de Informações sobre Mortalidade/SIM), e descreveu a tendência temporal da prevalência e mortalidade infantil por AC em nascidos vivos (NV) no Brasil e grandes regiões, de 2001 a 2018. Foram calculadas a prevalência de NV e a taxa de mortalidade infantil (TMI) por AC, para estas variáveis geográficas.

Para verificar a tendência temporal da prevalência de NV e TMI por AC neste período, aplicou-se a regressão linear generalizada de Prais-Winsten, considerando, para cada modelagem, o indicador de interesse como variável dependente e o tempo como variável independente, assumindo-se nível de significância de 5%. Por fim, obteve-se a taxa de variação anual (TVA) em porcentagem das tendências estudadas.

A prevalência e mortalidade infantil por AC mostraram-se crescentes no Brasil (TVA= 5,20% e 4,23%, respectivamente) e em quatro regiões. A prevalência apresentou-se crescente principalmente no Nordeste (TVA = 7,40%) e no Sudeste (TVA = 5,68%), e a região Sul mostrou uma tendência estacionária. Para a TMI, as regiões Norte e Nordeste apresentaram os maiores aumentos anuais (TVA= 10,12% e 9,90%, respectivamente), enquanto na região Sudeste mostrou-se estacionária, conforme pode ser observado na Figura 8.



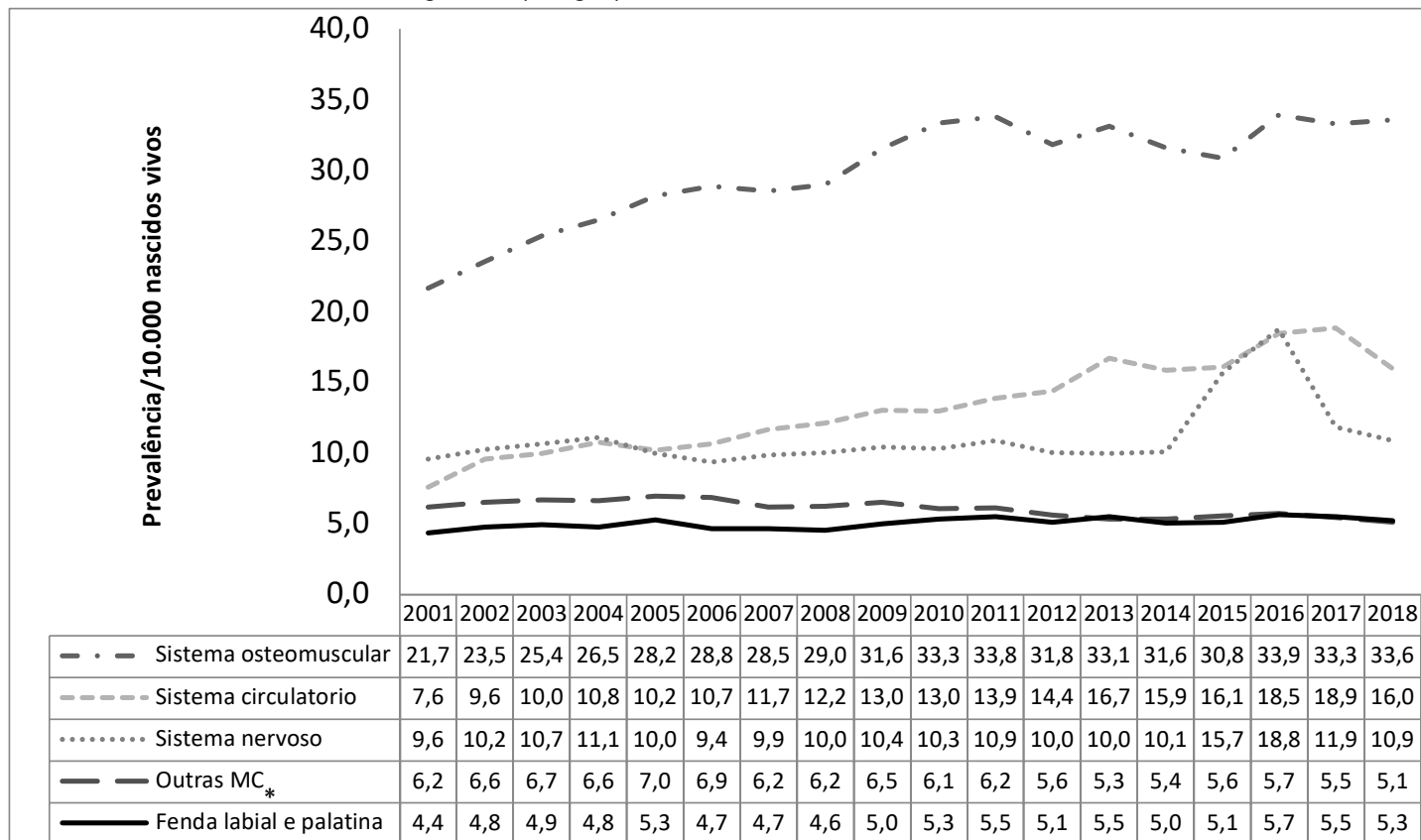
Figura 8: Prevalência (/10.000 nascidos vivos) das Anomalias Congênitas no Brasil e Regiões, 2001 a 2018



Fonte: Sistema de Informações sobre Nascidos Vivos (SINASC), acrescidos daqueles não registrados ao nascer, mas registrados na Declaração de Óbitos/Sistema de Informações sobre Mortalidade-DO/SIM).

As AC do aparelho osteomuscular foram as mais prevalentes ao nascimento (29,8/10.000 NV); as do aparelho circulatório ficaram na segunda posição (12,7/10.000 NV) após o linkage, onde anteriormente ocupavam a 6ª posição (Figura 9).

Figura 9: Prevalência (/10.000 nascidos vivos) das anomalias congênicas por grupo de causa. Brasil, de 2001 a 2018

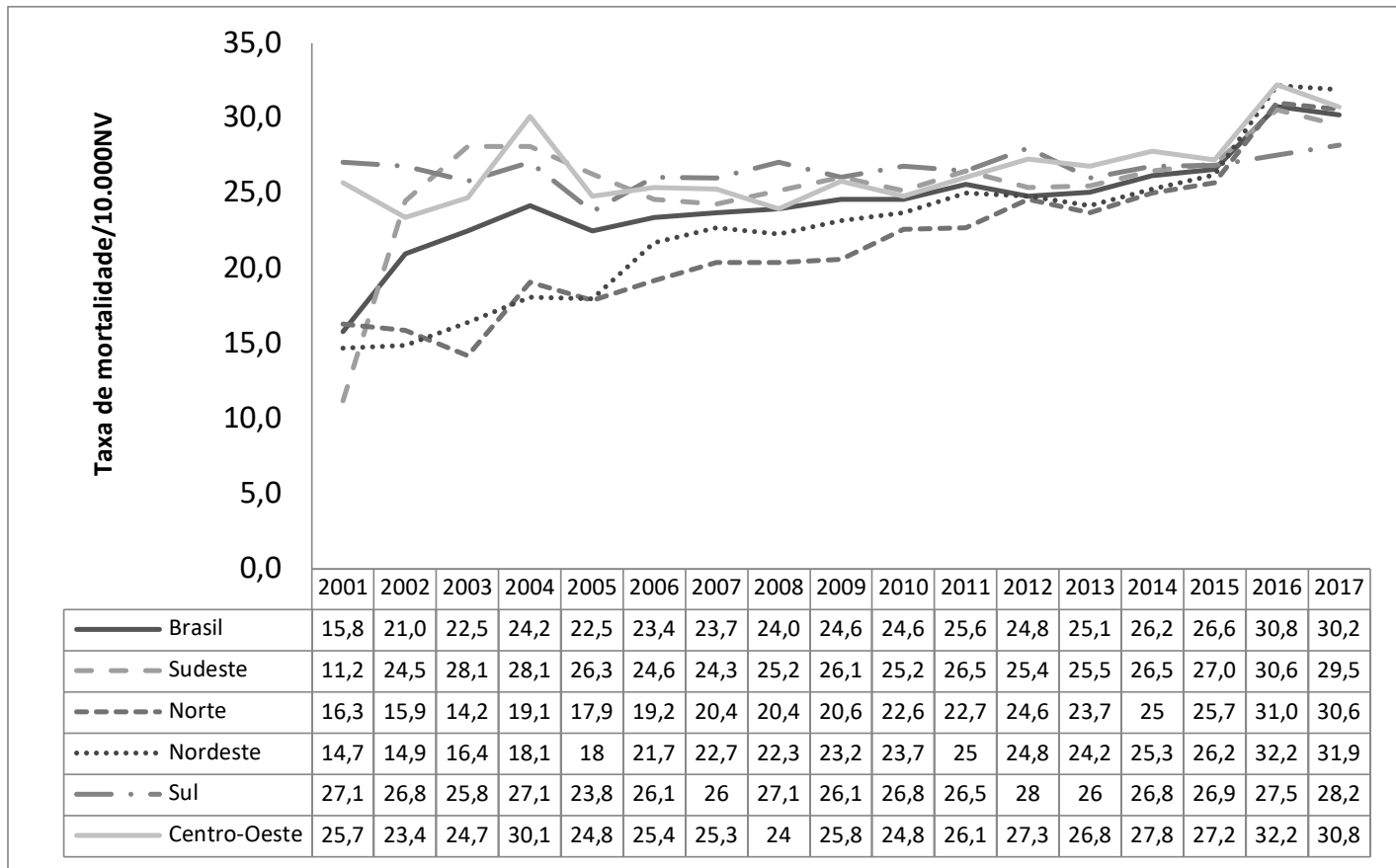


\*Outras MC = outras malformações congênicas

Fonte: Sistema de Informações sobre Nascidos Vivos (SINASC), acrescidos daqueles não registrados ao nascer, mas registrados na Declaração de Óbitos/Sistema de Informações sobre Mortalidade-DO/SIM).

Conforme se observa na Figura 10, as regiões Sul (26,6/10.000 NV) e Centro-Oeste (26,6/10.000 NV) apresentaram as maiores taxas de mortalidade infantil por AC, seguidas pelas regiões Sudeste (25,5/10.000 NV), Nordeste (22,4/10.000 NV) e Norte (21,8/10.000 NV).

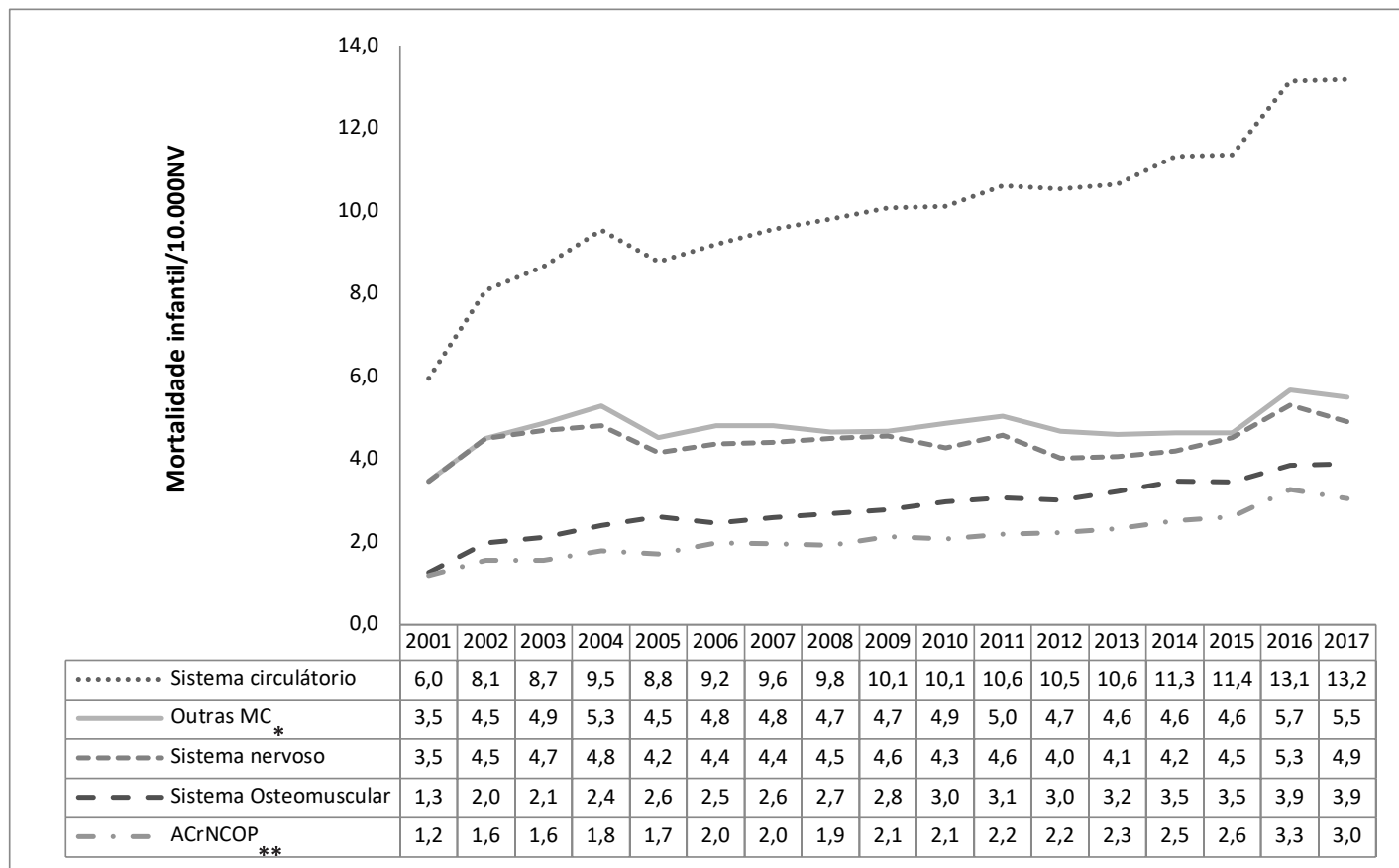
Figura 10: Taxas de mortalidade infantil (/10.000 nascidos vivos) por Anomalias Congênitas no Brasil, segundo região. Brasil, de 2001 a 2017



Fonte: SIM - Sistema de Informações sobre Mortalidade

As AC do sistema circulatório predominaram entre as principais causas de morte dos menores de um ano, apresentando taxa de mortalidade infantil de 10,0/10.000 NV, seguidas pelo grupo de outras MC (4,8/10.000 NV), sistema nervoso (4,4/10.000 NV) e sistema osteomuscular (2,8/10.000 NV). Os casos de anomalias cromossômicas não classificadas em outra parte (ACrNCOP), representaram 2,1/10.000 NV (Figura 11).

Figura 11: Taxas de mortalidade infantil (/10.000 nascidos vivos) por Anomalias Congênicas segundo grupos de causas. Brasil, 2001 a 2017



\*Outras MC = Outras malformações congênicas;

\*\*ACrNCOP = Anomalias cromossômicas não classificadas em outras partes

Fonte: Sistema de Informações sobre Mortalidade (SIM).

Constatou-se que a técnica de *linkage* aplicada corrigiu a prevalência nacional das AC em 17,9% no período analisado, após recuperar as AC notificadas no SIM, mostrando ser uma boa ferramenta para melhorar a quantidade e qualidade das informações sobre as AC no Brasil. Estes achados podem contribuir para entender melhor a ocorrência das AC e seu impacto na saúde infantil neste país.

## **Diferenças espaciais nas estimativas de risco da Síndrome Congênita pelo vírus Zika (SCZ) entre os municípios brasileiros, no período de 2015 a 2017**

O conhecimento sobre o risco dos eventos adversos associados à infecção congênita pelo vírus Zika tem sido produzido, predominantemente, a partir de análises de taxas relativas ao registro desses eventos em relação ao número estimado de nascidos vivos (NV), o que tem evidenciado uma heterogeneidade espacial muito ampla. Entretanto, esta heterogeneidade poderia ser explicada, com base nesta estimativa de risco, pela diversidade na intensidade da transmissão do ZIKV na população geral, e, conseqüentemente, nas gestantes e seus conceptos. Poucos estudos têm analisado quais outros fatores poderiam estar associados a essas diferenças, a maioria deles de base individual (coortes). Nesses estudos foram analisadas as seguintes variáveis: incidência de Zika recente e de dengue em anos anteriores, renda familiar, densidade domiciliar, nível educacional materno, participação no programa bolsa família.

A partir de análises realizadas com dados do Registro de Eventos em Saúde Pública (RESP) sobre os achados clínicos, de exames de imagem das crianças com suspeita de SCZ e os resultados laboratoriais de identificação dos agentes etiológicos implicados, França *et al.* (2016) propuseram uma classificação de casos de infecção congênita pelo ZIKV (confirmado, altamente provável, moderadamente provável, suspeito e descartado). Com base nesse modelo, foi desenvolvido um algoritmo para classificação automatizada de casos de SCZ, baseado na técnica de aprendizagem de máquina, com uso de dados secundários (estruturados e não estruturados) indicando que seu uso pode ser muito útil na rede de serviços de saúde pública (VEIGA; SCHULER-FACCINI; FRANÇA; ANDRADE *et al.*, 2021).

Visando explorar fatores associados à distribuição espacial (nível municipal) da SCZ, foi realizada nova aplicação do algoritmo, ampliando-se o número de casos classificados como confirmados, para o período de 2015 a 2017. A partir da seleção desses casos, foram construídos indicadores de risco, utilizando-se no numerador o número de casos SCZ classificados como confirmados por município, e no denominador, as seguintes variáveis:

- a) Número acumulado de NV no período;
- b) Número acumulado de Mulheres em Idade Fértil/MIF (10-49 anos) notificadas para Zika;
- c) Número acumulado de gestantes notificadas para Zika no SINAN, 2015-2017.

Esses indicadores foram submetidos à análise uni e multivariada (Programa R) considerando as seguintes variáveis:

- a) Índice Brasileiro de Privação (IBP) desenvolvido por equipe de pesquisadores do Cidacs/Fiocruz Bahia em colaboração com outras instituições (ICHIHARA; RAMOS; REBOUÇAS; OLIVEIRA *et al.*, 2018; ALLIK; LEYLAND; ICHIHARA; DUNDAS, 2020);
- b) Incidência acumulada de Zika na população total e em MIF (2016 e 2017);
- c) Indicadores de renda familiar, densidade domiciliar e nível educacional materno.

Foram classificados 4.412 casos como confirmados (entre 372 classificados como casos prováveis no RESP, 146 foram descartados com o uso do algoritmo). Na análise univariada, foi observada associação positiva estatisticamente significativa com o IBP (total) para o indicador de risco construído, tendo como denominador o número de NV (risco estimado 0,2681). Não foi observada associação estatisticamente significativa deste indicador com IBP (decil) e IBP (quartil), bem como, nas análises realizadas com os indicadores em que foram utilizados no denominador os seguintes grupos: Gestantes notificadas com Zika e MIF notificadas com Zika.

Os resultados deste estudo indicam que a distribuição espacial heterogênea da SCZ está associada com indicadores de privação, sendo ainda necessário explorar se esta associação é independente da incidência de Zika na população, assim como outros fatores socioeconômicos.

## **Vigilância baseada em dados populacionais para Síndrome Congênita do Vírus Zika: uma análise de classe latente de casos registrados de 2015 a 2018<sup>5</sup>**

A emergência de grande número de casos de microcefalia relatados no Nordeste do Brasil em 2015, juntamente com a suspeita de uma relação causal com o vírus Zika (ZIKV), identificado pela primeira vez no país em maio do mesmo ano, levou o Ministério da Saúde a implantar o Registro de Eventos de Saúde Pública (RESP) para registrar dados sobre todos os casos suspeitos desta nova condição, mais tarde chamada de Síndrome Congênita do Zika (SCZ).

Os primeiros objetivos do RESP foram fornecer uma avaliação rápida da magnitude e localização geográfica dos casos e permitir uma compreensão do espectro clínico da síndrome. Por essas razões, os critérios iniciais de notificação foram mais amplos para capturar o maior número de casos possível. Estabelecer uma classificação baseada em critérios clínicos e epidemiológicos de uma nova entidade mórbida é sempre desafiador, principalmente antes de se adquirir conhecimentos suficientes sobre a sua história natural, em um cenário com limitada disponibilidade de testes diagnósticos específicos. A classificação dos casos suspeitos de SCZ baseou-se inicialmente na observação clínica e na análise de exames de imagem do cérebro. Os critérios de classificação de caso foram refinados várias vezes durante a epidemia de SCZ, à medida que lacunas no conhecimento foram sendo preenchidas. Os casos suspeitos de SCZ registrados no RESP foram posteriormente classificados como:

- a) casos confirmados por identificação etiológica (confirmação laboratorial de ZIKV por teste molecular, ou sorológico, de amostra da mãe, feto, ou recém-nascido), ou exame clínico e de imagem;
- b) casos prováveis - aqueles com sinais ou sintomas compatíveis com a síndrome, cujas mães não relataram sintomas de Zika durante a gravidez e apresentaram resultado de exame laboratorial específico para ZIKV negativo, ou inconclusivo.

---

5. Para ver o artigo publicado sobre essa análise, consultar PAIXÃO; RODRIGUES; COSTA; CARVALHO-SAUER *et al.*, 2022

c) casos descartados: diagnóstico de outras infecções congênitas (TORCHS: toxoplasmose, sífilis, varicela-zoster, parvovírus B19, rubéola, citomegalovírus, herpes); síndrome alcoólica fetal, dentre outras. Também poderiam não ter sido confirmados porque os critérios usados para definir microcefalia/SCZ tornaram-se mais rigorosos, não se adequando ao critério inicialmente estabelecido; ou porque investigação mais cuidadosa não encontrou resultados de imagens consistentes com o que estava sendo reconhecido como característico da SCZ;

d) Casos inconclusivos: ausência de informações suficientes para classificação adequada. Embora estudos anteriores tenham descrito as características de casos suspeitos registrados no RESP em 2015 e 2016, ainda não havia sido explorada a possível heterogeneidade existente entre os casos não confirmados.

Considerando que mesmo algumas dessas condições poderiam resultar em desfechos adversos para os RN, entendeu-se ser importante investigar a evolução, em médio prazo, do conjunto de casos suspeitos. Assim, este estudo teve como objetivo descrever os achados clínicos e determinar a sobrevida a médio prazo dos casos suspeitos de SCZ. Foi baseado em dados nacionais correspondentes a todos os casos suspeitos de SCZ no Brasil de 2015–2018, tendo dois focos principais: (1) descrever a antropometria ao nascimento, achados de imagem e mortalidade até 36 meses de vida, de acordo com a classificação clínico-epidemiológica final (confirmado/provável, inconclusivo ou não confirmado); (2) investigar padrões em uma série de características (perímetro craniano, peso ao nascer, exames de imagem, local e ano de nascimento) entre os casos não confirmados, para identificar grupos semelhantes e explorar a mortalidade por características de cada grupo até a idade de 36 meses.

Nossa hipótese foi que casos não confirmados, com medidas antropométricas abaixo do esperado ao nascimento e imagens anormais, têm um risco aumentado de morte, seja porque são casos de SCZ atípica, ou porque têm outras patologias graves. Ao assumir esta hipótese, nosso objetivo foi identificar padrões de características clínicas para melhor entender possíveis caminhos causais para resultados letais, por exemplo, reconhecendo características que estão fortemente associadas à mortalidade.



Assim, foi realizado um estudo retrospectivo de base populacional, incluindo todos os casos suspeitos de SCZ registrados no RESP e, como grupo de comparação, a população de nascidos vivos sem registro de SCZ no Brasil entre 01/01/2015 e 31/12/2018. Ambos os grupos foram acompanhados desde o nascimento até a morte, até 36 meses de vida, ou até 31/12/2018, considerando o evento que ocorreu primeiro. Em seguida, usou-se a análise de classes latentes para agrupar casos incluídos em classes com combinações semelhantes de antropometria ao nascimento, achados de imagem, relato materno de erupção cutânea, região e ano de nascimento. As curvas de Kaplan-Meier foram traçadas e os modelos de risco proporcional de Cox foram ajustados para determinar a mortalidade até 36 meses.

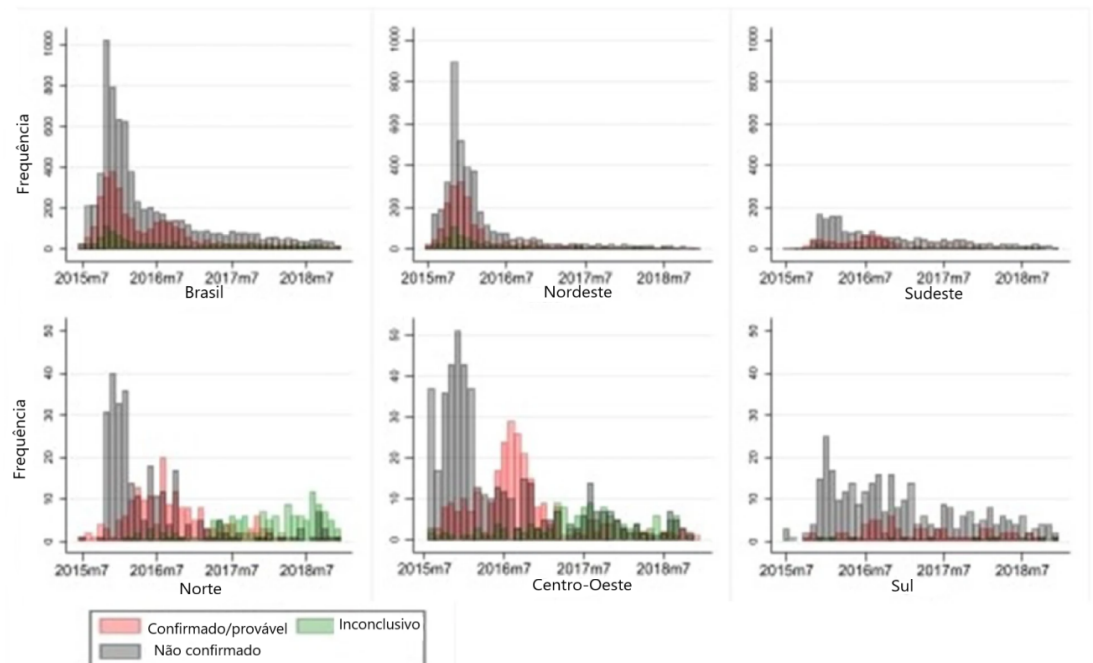
Foram acompanhados 11.850 casos suspeitos de SCZ, dos quais 28,3% confirmados/prováveis, 9,3% inconclusivos e 62,4% excluídos. O grupo de comparação constou de 11.724.061 nascidos vivos no mesmo período, sem registro vinculado de SCZ. Os casos confirmados/prováveis tiveram mortalidade quase duas vezes maior quando comparados aos casos excluídos.

Entre os casos excluídos, identificamos três classes distintas com diferentes trajetórias de mortalidade. O maior risco de mortalidade foi observado nos casos com achados de imagens anormais compatíveis com infecções congênicas (HR=12,6; IC95% 8,8-18,0) e outras anormalidades (HR=11,6; IC95% 8,6-15,6), em comparação com aqueles com achados de imagens normais. O risco foi alto naqueles com microcefalia grave (HR=8,2; IC95% 6,4-10,6) e macrocefalia (HR=6,6; IC95% 4,5-9,7) em comparação com aqueles cujo perímetro cefálico era normal.

A Figura 12 apresenta a distribuição temporal dos casos por mês de nascimento e classificação final, para o Brasil como um todo e para suas cinco regiões geográficas. No geral, observamos duas ondas de casos confirmados/prováveis de SCZ, ocorrendo de outubro de 2015 a janeiro de 2016 e de julho de 2016 a outubro de 2016. A primeira onda atingiu o pico em dezembro de 2015, com quase 400 casos registrados. O ápice da segunda onda, agosto e setembro de 2016, foi de cerca de metade dessa magnitude (menos de 200 casos).

A região Nordeste apresentou a maioria dos casos na primeira onda, enquanto a segunda onda atingiu principalmente as regiões Sudeste e Centro-Oeste, sendo que o maior número de casos confirmados/prováveis notificados nessas duas regiões e na região Norte, aconteceu de julho a setembro de 2016. A maioria dos casos notificados anteriormente a essa época, não foi confirmada após investigação epidemiológica.

Figura 12: Número de casos suspeitos de SCZ por mês e região de nascimento comparando casos confirmados, prováveis, não confirmados e inconclusivos Brasil, 2015-2018



\*O eixo Y difere por causa do pequeno número em algumas regiões.

Fonte: Registro de Eventos em Saúde Pública (RESP)

A análise de classe latente de casos não confirmados revelou três classes (ou grupos) distintas, baseadas na antropometria ao nascimento, exantema materno, achados de imagem, região e ano de nascimento. O grupo 1, que incluiu 16,0% destes casos, apresentou alta morbidade, com microcefalia grave e RN pequeno para idade gestacional

(PIG). Este grupo concentrou a maior proporção de casos não confirmados com achados de imagem (46,7%), mas as mães foram menos propensas a relatar erupções cutâneas durante a gravidez em comparação com os demais.

Os casos não confirmados nas classes 2 e 3 tendiam a perímetro cefálico normal ao nascer. Os casos do grupo 2 concentraram-se na região Sudeste em 2016, incluindo 33,7% dos casos não confirmados. Os casos do grupo 3 apresentaram menos achados anormais em exames de imagem, estavam concentrados na região Nordeste e reportados em 2015. Este último grupo, incluiu 50,3% dos casos não confirmados e maior proporção de mulheres que referiram erupção cutânea durante a gravidez (Tabela 1). A mortalidade por todas as causas não ajustada até 36 meses foi maior para a classe 1 do que para as classes 2 e 3, e ainda mais elevada que a dos casos confirmados. Já a mortalidade mais baixa foi observada para a classe 3, embora ainda maior do que o observado no grupo de comparação.

Para condições individuais entre casos não confirmados, os piores resultados foram observados naqueles com achados de imagem compatível com infecção congênita (HR = 12,6; IC95% 8,8–18,0), ou outras anormalidades (HR = 11,6; IC95% 8,6–15,6), em comparação com aqueles sem achados de imagem anormais. Microcefalia grave (HR = 8,2; IC95% 6,4–10,6) e macrocefalia (HR = 6,6; IC95% 4,5–9,7) também foram fortemente associadas à mortalidade, em comparação com a mortalidade daqueles com circunferência craniana normal. Entre os casos de SCZ não confirmados que eram PIG, a mortalidade foi 1,9 (HR = 1,9; IC95% 1,5–2,3), e daqueles que nasceram grandes para idade gestacional foi de 2,8 (IC95% 1,8–4,5), maior do que em casos com idade gestacional adequada (Figura 13). A erupção cutânea materna não foi associada à mortalidade infantil entre os casos não confirmados.

Estes achados revelaram que imagens com achados sugestivos e perímetro cefálico parecem ser os principais fatores do aumento da mortalidade entre os casos suspeitos de SCZ. Deste modo, sugere-se identificar crianças com maior probabilidade de morrer e que exigem assistência médica de melhor qualidade e maior complexidade, indicando a necessidade de aprimorar e otimizar os cuidados mediante alocação de recursos.

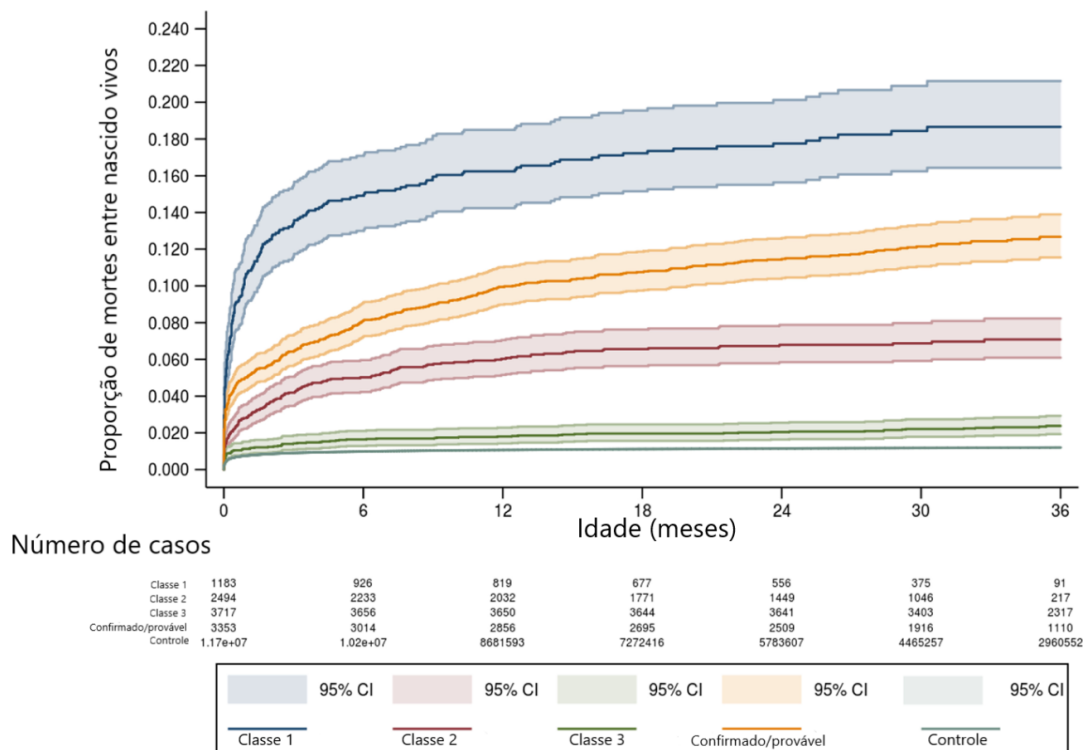
Tabela 1: Total de casos não confirmados, classes de casos não confirmados e casos sem informação de Síndrome Congênita do Zika, de acordo com perímetro cefálico, peso ao nascer, achados dos exames de imagem, exantemas durante a gestação, região e ano. Brasil, 2015-2018

Variável	Total de casos não confirmados	Clusters de casos não confirmados*			p-valor	Missing (%)		
		Classe 1* N=1183 (16,0%)	Classe 2** N=2494 (33,7%)	Classe 3*** N=3717 (50,3%)		Classe 1	Classe 2	Classe 3
<b>Perímetro cefálico</b>								
Normal	4102 (6,3)	-	1632 (72,9)	2470 (75,7)				
Microcefalia	1815 (27,6)	501 (46,2)	603 (27,0)	711 (21,8)	<0,001			
Microcefalia severa	493 (7,5)	423 (39,9)	-	70 (2,1)				
Macrocefalia	174 (2,6)	161 (14,8)	2 (0,1)	11 (0,3)		257 (10,3)	98 (8,3)	455 (12,2)
<b>Peso ao nascer</b>								
Adequado para idade gestacional	5787 (80,5)	371 (31,8)	2279 (93,1)	3137 (87,8)				
Pequeno para idade gestacional	1260 (17,5)	730 (62,6)	143 (5,8)	387 (10,8)	<0,001			
Grande para idade gestacional	140 (1,9)	65 (5,6)	27 (1,1)	48 (1,3)		45 (1,8)	17 (1,4)	145 (3,9)
<b>Exame de imagem</b>								
Sem alterações	2709 (54,3)	376 (41,8)	1076 (63,0)	1257 (52,7)				
Alteração compatível com infecção congênita	263 (5,3)	88 (9,8)	133 (7,8)	42 (1,8)	<0,001			
Outras alterações	743 (14,9)	332 (36,9)	326 (19,1)	85 (3,6)				
Não realizado	1276 (25,6)	103 (11,5)	172 (10,1)	1001 (42,0)		787 (31,6)	284 (24,0)	1332 (35,8)
<b>Relato de erupção cutânea</b>								
Sim	843 (18,1)	74 (8,1)	324 (15,1)	445 (27,7)	<0,001			
Não	3824 (81,9)	840 (91,9)	1822 (84,9)	1162 (72,3)		348 (13,9)	269 (22,7)	2110 (56,8)
<b>Região</b>								
Norte	284 (3,8)	26 (2,2)	151 (6,1)	107 (2,9)	<0,001			
Nordeste	4205 (56,9)	475 (40,1)	319 (12,8)	3411 (91,8)				
Sul	315 (4,3)	151 (12,8)	164 (6,6)	-				
Sudeste	2118 (28,6)	426 (36,0)	1690 (67,8)	2 (0,1)				
Centro-Oeste	472 (6,4)	105 (8,9)	170 (6,8)	197 (5,3)		-	-	-
<b>Ano</b>								
2015	2725 (36,8)	107 (9,0)	231 (9,3)	2387 (64,2)	<0,001			
2016	3173 (42,9)	545 (46,1)	1298 (52,0)	1330 (35,8)				
2017	955 (12,9)	323 (27,3)	632 (25,3)	-				
2018	541 (7,3)	208 (17,6)	333 (13,4)	-		-	-	-

Classe 1\* grupo com alta proporção de microcefalia; Classe 2\*\* alta proporção de perímetro cefálico e peso ao nascer normais e, segunda maior proporção de alterações nos exames de imagem; Classe 3\*\*\* alta proporção de perímetro cefálico e peso ao nascer normais e baixa proporção de alterações nos exames de imagem.

Fonte: Registro de Eventos em Saúde Pública (RESP)

Figura 13: Total de casos não confirmados, classes de casos não confirmados e casos sem informação de Síndrome Congênita do Zika de acordo com perímetro cefálico, peso ao nascer, achados dos exames de imagem, exantemas durante a gestação, região e ano. Brasil, 2015-2018



\*A Classe 1 caracteriza o grupo com alta proporção concomitante de microcefalia, pequeno para a idade gestacional e imagens anormais; Classe 2 - alta proporção de tamanho da cabeça e peso ao nascer normais e a segunda maior proporção de achados anormais nos testes de imagem; Classe 3 - alta proporção de tamanho da cabeça e peso ao nascer normais e baixa proporção de casos com achados anormais em testes de imagem.

Fonte: Registro de Eventos em Saúde Pública (RESP)

## **Marcadores de risco socioeconômico da Síndrome Congênita do Zika: estudo nacional baseado em registros administrativos, Brasil 2015-2018<sup>6</sup>**

Marcadores socioeconômicos estão associados a um maior risco de infecções por arboviroses urbanas, entretanto evidências sobre a relação entre fatores socioeconômicos e a Síndrome Congênita do Zika (SCZ) permanecem limitadas. Este estudo investigou a relação entre marcadores de risco socioeconômico e nascidos vivos com SCZ no Brasil. Foi realizado um estudo de base populacional com dados de todos os nascidos vivos registrados no Brasil pelo SINASC e vinculados ao RESP, de 1º de janeiro de 2015 a 31 de dezembro de 2018.

Utilizou-se modelos de regressão logística para estimar o *Odds Ratio* (OR) e os intervalos de confiança de 95% (IC 95%) da SCZ, com base em uma estrutura de três níveis. Em uma análise de 11.366.686 nascidos vivos, sendo 3.353 com SCZ, observamos que nascidos vivos de mães que se declararam como pretas ou pardas (OR=1,72; IC 95% 1,47-2,01 e OR=1,37; IC 95% 1,24-1,51, respectivamente) foram associados a maiores chances de SCZ. Nascidos vivos de mulheres solteiras em comparação com casadas e com menos de 12 anos de escolaridade, em comparação com mais de 12 anos de escolaridade, também apresentaram maiores chances de SCZ. Além disso, nascidos vivos com menor número de consultas de pré-natal tiveram maiores chances de SCZ ao se considerar o país como um todo (Tabela 2).

---

6. PAIXÃO; FERNANDES; CARDIM; PESCARINI *et al.*, 2022

Tabela 2: Total de casos não confirmados, classes de casos não confirmados e casos sem informação de Síndrome Congênita do Zika de acordo com perímetro cefálico, peso ao nascer, achados dos exames de imagem, exantemas durante a gestação, região e ano. Brasil, 2015-2018

Variáveis demográficas	Modelo 1 *		Modelo 2**		Modelo 3 ***	
	OR(CI 95%)	p-valor	OR(CI 95%)	p-valor	OR(CI 95%)	p-valor
<b>Ano de nascimento</b>						
2015	1		1		1	
2016	1,34(1,25-1,45)	0,000	1,31(1,21-1,42)	0,000	1,30(1,20-1,40)	0,000
2017	0,31(0,28-0,35)	0,000	0,30(0,27-0,34)	0,000	0,29(0,26-0,33)	0,000
2018	0,16(0,14-0,19)	0,000	0,16(0,14-0,19)	0,000	0,16(0,13-0,18)	0,000
<b>Sexo do RN</b>						
Masculino	1		1		1	
Feminino	1,19(1,11-1,28)	0,000	1,19(1,11-1,28)	0,000	1,20(1,11-1,28)	0,000
<b>Região de nascimento</b>						
Sul	1		1		1	
Sudeste	3,83(3,02-4,86)	0,000	3,40(2,66-4,33)	0,000	3,41(2,67-4,37)	0,000
Norte	3,19(2,44-4,18)	0,000	2,66(2,01-3,53)	0,000	2,71(2,03-3,61)	0,000
Nordeste	13,51(10,72-17,04)	0,000	11,07(8,66-14,13)	0,000	11,34(8,86-14,52)	0,000
Centro-Oeste	6,47(5,00-8,37)	0,000	5,43(4,15-7,10)	0,000	5,68(4,32-7,45)	0,000
<b>Raça/cor materna</b>						
Branca	1		1		1	
Preta			1,72(1,47-2,01)	0,000	1,47(1,25-1,73)	0,000
Parda			1,37(1,24-1,51)	0,000	1,21(1,09-1,34)	0,000
Asiática			1,33(0,75-2,36)	0,329	1,33(0,75-2,36)	0,327
Indígena			1,19(0,76-1,87)	0,442	0,98(0,61-1,57)	0,929
<b>Escolaridade materna (anos)</b>						
0 – 3					1,20(0,97 - 1,50)	0,094
4 – 7					1,30(1,13-1,50)	0,000
12 – 8					1,28(1,13-1,44)	0,000
> 12					1	
<b>Estado civil materno</b>						
Casada/união estável					1	
Solteira					1,38(1,28-1,49)	0,000
Viúva					0,88(0,33-2,34)	0,793
Divorciada					1,13(0,75-1,69)	0,557
<b>Número de consultas pré-natal</b>						
Nenhuma					2,84(2,12-3,80)	0,000
1 – 3					1,35(1,19-1,53)	0,000
4 – 6					1,10(1,02-1,20)	0,020
7+					1	
<b>Idade materna (anos)</b>						
< 20					1,12(0,97-1,28)	0,118
20-34					1,02(0,90-1,15)	0,778
35+					1	

\* Modelo 1 é minimamente ajustado - por ano de nascimento, região e sexo do recém-nascido.

\*\* Modelo 2 - ano de nascimento, região, sexo do recém-nascido e raça/cor materna

\*\*\* Modelo 3 - ano de nascimento, região, sexo do recém-nascido, raça/cor, educação materna, estado civil, idade materna, número de consultas pré-natal

Fonte: Registro de Eventos em Saúde Pública (RESP)



Entretanto, nas análises realizadas com dados da região Nordeste (onde a epidemia de microcefalia começou antes do estabelecimento do vínculo com o ZIKV e antes que as medidas preventivas fossem conhecidas, implementadas e/ou disseminadas), não foi encontrada associação estatística entre o número de consultas de pré-natal e a chance de SCZ. Este estudo mostrou que nascidos vivos das mulheres socialmente mais vulneráveis no Brasil tiveram maior chance de terem filhos com SCZ. Essa distribuição desproporcional de risco sobrecarrega ainda mais os grupos socioeconômicos já desfavorecidos, e as incapacidades ao longo da vida causadas por essa síndrome podem reforçar as desigualdades sociais e de saúde existentes.

## DESFECHOS ADVERSOS DA SÍNDROME CONGÊNITA DO ZIKA

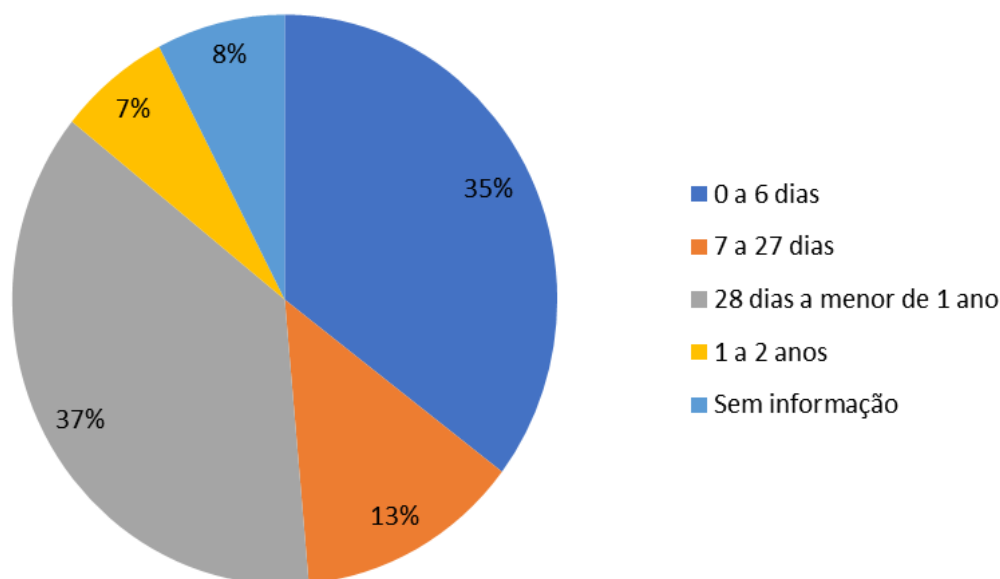
### Letalidade e Fatores Associados às mortes por Microcefalia/Síndrome Congênita do Zika - Brasil, 2015-2017<sup>7</sup>

A infecção congênita do vírus Zika (ZIKV) produz graves danos à saúde da criança afetada, aumentando sua vulnerabilidade à morbimortalidade, conforme já vinha sendo registrado pelos serviços de saúde, bem como, em algumas publicações científicas restritas a registros hospitalares (MENDES; MAIA; ZACARKIM; QUEIROZ; LABEAUD, 2018; CUNHA; MAGALHÃES-BARBOSA; LIMA-SETTA; MEDRONHO; PRATA-BARBOSA, 2016). Este é um estudo de abrangência nacional, realizado a partir da vinculação de três Bancos de Dados (SIM, SINASC e RESP), que teve como objetivo estimar a taxa de letalidade da microcefalia/síndrome congênita do Zika (SCZ), e identificar seus fatores associados, no Brasil, em 2015-2017. Dos 6.059 nascidos vivos, classificados como casos confirmados e possíveis (registrados como prováveis e inconclusivos) de microcefalia/SCZ incluídos no estudo, 603 evoluíram para óbito, dos quais 92,8% foram de menores de 1 ano, especialmente aqueles ocorridos antes de 28 dias de vida (52,6%) (Figura 14).

---

7. Para ver o artigo publicado sobre essa análise, consultar COSTA; CARDIM; TEIXEIRA; BARRETO *et al.*, 2020.

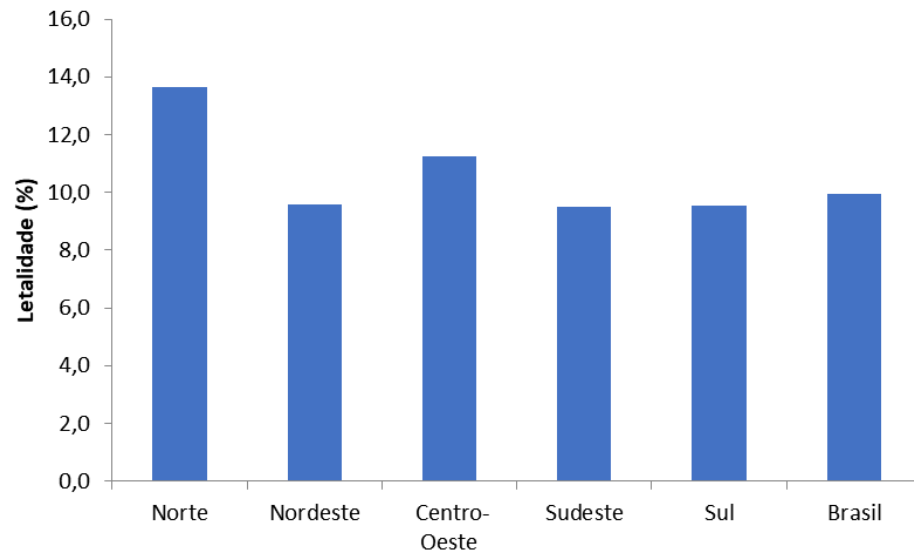
Figura 14: Proporção de óbitos de crianças até 2 anos de idade (n=603), nascidas com síndrome congênita (casos confirmados, prováveis e inconclusivos) do Zika, por faixa etária. Brasil, 2015-2017



Fonte: Elaborada pelos autores

A taxa de letalidade referente aos 603 casos confirmados, prováveis e inconclusivos para o país como um todo foi de 10%, sendo mais elevada nas regiões Norte (13,6%) e Centro-Oeste (11,3%), embora as regiões com maiores números de nascidos vivos com esta síndrome tenham sido Nordeste (3.818) e Sudeste (1.368) (Figura 15).

Figura 15: Taxa de letalidade (%) da Síndrome Congênita do Zika (casos confirmados, prováveis e inconclusivos) de crianças (n=603) até 2 anos de idade. Brasil e regiões geográficas, 2015-2017



Fonte: Elaborada pelos autores

Considerando-se apenas os casos confirmados, a taxa de letalidade foi de 9,4% (variação de 8,4% a 10,6%), ou seja, em média, morreram nove de cada 100 crianças com esta síndrome. Dentre os fatores associados investigados, destacaram-se Peso ao Nascer e índice Apgar<sup>8</sup>. Não se verificou associação entre essas mortes e as características sociodemográficas maternas analisadas. Apesar das limitações inerentes aos estudos realizados com dados secundários, devido à subnotificação, sub-registro e incompletude, que interferem negativamente na vinculação dos dados, os resultados desta investigação possibilitaram revelar e discutir as desigualdades sociais na distribuição dos leitos de Unidade de Tratamento Intensivo (UTI) Neonatal.

8. O índice de Apgar, é uma escala desenvolvida pela Dra. Virginia Apgar (1909 – 1974), que corresponde a um índice de vitalidade de 0 a 10 atribuído a, praticamente, todos os recém-nascidos com 1, 5 e 10 minutos de nascimento. A pontuação é baseada em medidas de frequência cardíaca, esforço respiratório, cor da pele, tônus muscular e irritabilidade reflexa.

As informações geradas neste estudo são importantes para os profissionais da rede de serviços de saúde do país que prestam cuidados às crianças nascidas com microcefalia/SCZ, visto que estas necessitam de atenção específica, especialmente nos primeiros dias de vida. Além disso, evidenciam a necessidade dos gestores dos três níveis de gestão do SUS, captarem e disporem de recursos para ampliar o número de UTIs Neonatais, e assim impedir a ocorrência de óbitos evitáveis.

Considerando-se que o risco de circulação do ZIKV permanece, quer na forma endêmica e/ou epidêmica, pelo fato de inúmeros centros urbanos do Brasil e de outros países das Américas continuarem infestados por mosquitos do gênero *Aedes*, principalmente *Aedes aegypti*, o presente estudo também alertou para a importância das orientações sobre a necessidade de manutenção das medidas de proteção individual, particularmente para gestantes e mulheres em idade fértil. Além disso, destacou a relevância do fortalecimento e aprimoramento das ações de controle dos *Aedes* visando à prevenção de Zika e, conseqüentemente, de novos casos de microcefalia/SCZ.

### **Mortalidade por Síndrome Congênita do Zika: Estudo de Coorte Nacional, Brasil 2015-2018<sup>9</sup>**

A exposição pré-natal ao vírus Zika tem potenciais efeitos teratogênicos, com amplo espectro de apresentação clínica que podem impactar a sobrevivência das crianças com SCZ. Entretanto, informações sobre o prognóstico destas crianças são limitadas. Este estudo teve como objetivos investigar o risco de mortalidade e as causas de óbitos entre nascidos vivos com essa síndrome, em comparação com aqueles sem a síndrome, além de examinar o papel da idade gestacional ao nascer e peso ao nascer como preditores de mortalidade entre estes dois grupos.

Foi realizado um estudo de coorte de base populacional, com dados coletados rotineiramente no Brasil, de janeiro de 2015 a dezembro de 2018, para estimar a mortalidade pela síndrome. Curvas de Kaplan-Meier e modelos de sobrevivência foram avaliados

---

9. Para ver o artigo publicado sobre essa análise, consultar PAIXÃO; CARDIM; COSTA; BRICKLEY *et al.*, 2022.

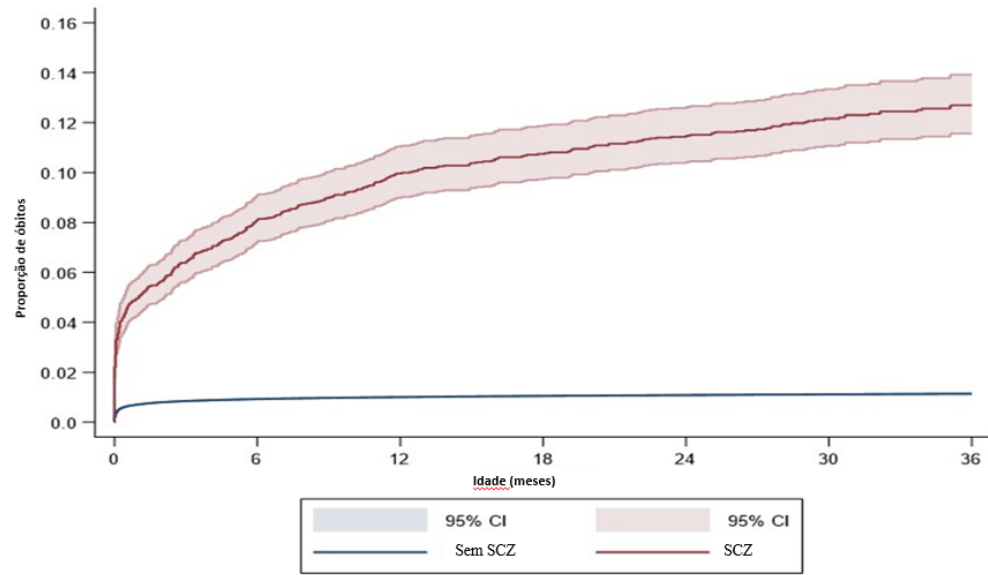
com ajuste para confundimento e com estratificação de acordo com a idade gestacional, peso ao nascer e *status* de ser pequeno para a idade gestacional.

Um total de 11.481.215 crianças nascidas vivas foram acompanhadas até os 36 meses de idade. A taxa de mortalidade foi de 52,6 mortes (IC 95% 47,6-58,0) por 1.000 pessoas-ano entre crianças nascidas vivas com SCZ, em comparação com 5,6 mortes (IC 95% 5,6-5,7) por 1.000 pessoas-ano entre aqueles sem a Síndrome. A razão da taxa de mortalidade entre os nascidos vivos com SCZ, em comparação com aqueles sem a Síndrome, foi 11,3 (IC 95% 10,2-12,4).

Entre os bebês nascidos antes de 32 semanas de gestação ou com peso de nascimento inferior a 1.500g, os riscos de morte foram semelhantes, independentemente do status da SCZ. Entre os bebês nascidos a termo, aqueles com a SCZ foram 14,3 vezes (IC 95% 12,4-16,4) mais prováveis de morrer do que aqueles sem a Síndrome (taxa de mortalidade, 38,4 vs. 2,7 mortes por 1.000 pessoas-ano). Entre os recém-nascidos com peso de nascimento igual ou superior a 2.500g, aqueles com a SCZ foram 12,9 vezes (IC 95% 10,9-15,3) mais prováveis de morrer do que aqueles sem a Síndrome (taxa de mortalidade, 32,6 vs. 2,5 mortes por 1.000 pessoas-ano).

A carga de anomalias congênitas, doenças do sistema nervoso e doenças infecciosas como causas básicas de mortes foram maiores entre os nascidos vivos com SCZ do que entre aqueles sem a síndrome. Assim concluiu-se que o risco de morte foi maior entre os nascidos vivos com SCZ do que entre aqueles sem a Síndrome e persistiu durante todos os três primeiros anos de vida. Os autores destacaram a necessidade de instituir protocolos pós-natais bem estabelecidos, incluindo intervenção precoce, que podem ajudar a reduzir sequelas físicas e de desenvolvimento neurocognitivo e melhorar a sobrevivência dessas crianças, bem como, o aperfeiçoamento da vigilância e controle da Zika, com especial destaque à proteção das gestantes e mulheres em idade fértil.

Figura 16: Curvas de Kaplan-Meier mostrando a sobrevida comparando nascimentos com SCZ com nascimentos sem SCZ até 36 meses de idade



Fonte: Registro de Eventos em Saúde Pública (RESP) e Sistema de Informação sobre Nascidos Vivos (SINASC)

## Principais Causas de morte de Crianças com Síndrome Congênita do Zika no Brasil, 2015-2018<sup>10</sup>

A infecção intrauterina pelo ZIKV pode danificar gravemente o sistema nervoso central e periférico do concepto, aumentando o risco de morte da criança nascida com a SCZ (MOORE; STAPLES; DOBYNS; PESSOA, *et al.*, 2017; KARWOWSKI; NELSON; STAPLES; FISCHER, *et al.*, 2016). Todavia, a sequência das causas (básicas, intermediárias, imediatas e contribuintes) que resultaram na morte dessas crianças ainda não estava estabelecida. Visando preencher essa lacuna do conhecimento, este estudo de base populacional descreveu a sequência de causas registradas nas Declarações de Óbito

10. Para ver o artigo publicado sobre essa análise, consultar COSTA; CARDIM; MOORE; JESUS *et al.*, 2023.

(DO) dessas crianças até os 36 meses de idade, bem como, sua probabilidade de morrer devido a cada uma dessas causas, no Brasil, de 2015 a 2018.

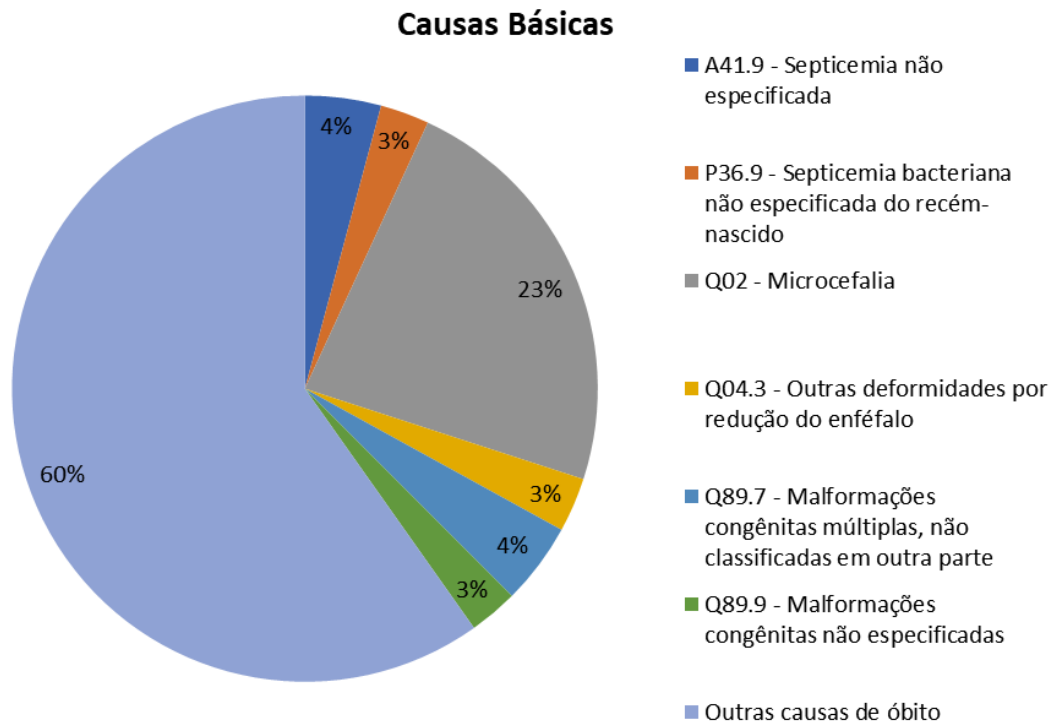
Para esse estudo, foram vinculados dados administrativos sobre nascidos vivos, óbitos e casos confirmados e prováveis de SCZ registrados de 2015 a 2018, respectivamente, no SINASC, RESP e SIM, sendo encontrados 403 óbitos da faixa etária de interesse, que passaram a constituir a população deste estudo. A seguir, estimou-se a proporção (%) de óbitos dessas crianças por faixa etária e a mortalidade e proporcional (MP %) para cada doença e/ou circunstância que constituiu a sequência de causas que determinaram essas mortes.

Esse indicador também foi calculado para essas mesmas causas para as crianças que nasceram com anomalias congênitas do sistema nervoso central, porém não relacionadas ao ZIKV (grupo de comparação), após o cálculo da razão de mortalidade proporcional por causa (RMPc). Este indicador é obtido a partir da divisão entre a MP (%) desses dois grupos de crianças. No presente estudo, RMPc indica a probabilidade da criança com SCZ morrer por uma dada causa, quando comparada com aquelas do grupo com anomalias congênitas do sistema nervoso central, porém não relacionadas ao Zika.

Das 403 crianças com SCZ confirmada e provável, cujo óbito ocorreu até os 36 meses de idade, 81,9% tinham menos de 12 meses de vida. Malformações congênitas múltiplas não classificadas em outra parte e septicemia não especificada foram as condições mais registradas pelos médicos como causas básicas de morte, com respectivamente 18 (MP=4,5%) e 17 (MP=4,2%) óbitos (Figura 17).



Figura 17: Principais causas básicas de óbitos de crianças com até 36 meses de vida, nascidas com síndrome congênita do Zika. Brasil, 2015-2018



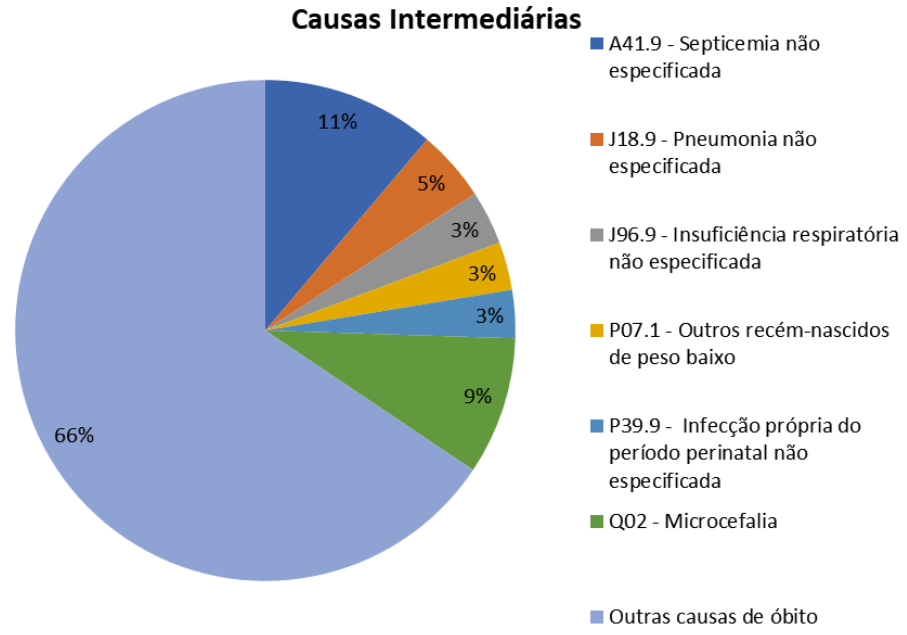
Obs: Embora Microcefalia tenha sido registrada como causa de morte, ela não é uma doença e, portanto, não deveria ter sido atestada como tal; trata-se de um sinal resultante dos danos produzidos pelo vírus Zika ao sistema nervoso central da criança.

Fonte: Elaborada pelos autores

Quando comparadas com o grupo sem a SCZ, as causas básicas com maior probabilidade de levar à morte em crianças com SCZ foram: feto e recém-nascido acometidos por doenças infecciosas e parasitárias da mãe (RMPc=17,0), seguidas de desnutrição proteico-calórica grave não especificada; pneumonia não especificada e asfixia não especificada ao nascimento, cada uma com RMPc=10,0.

Septicemia não especificada e pneumonia não especificada foram relatadas como as causas intermediárias de morte mais frequentes entre crianças com SCZ (29 e 12 causas de morte registradas e MP de 8,9% e 4,6%, respectivamente (Figura 18).

Figura 18: Principais causas intermediárias de óbitos de crianças com até 36 meses de vida, nascidas com síndrome congênita do Zika. Brasil, 2015-2018

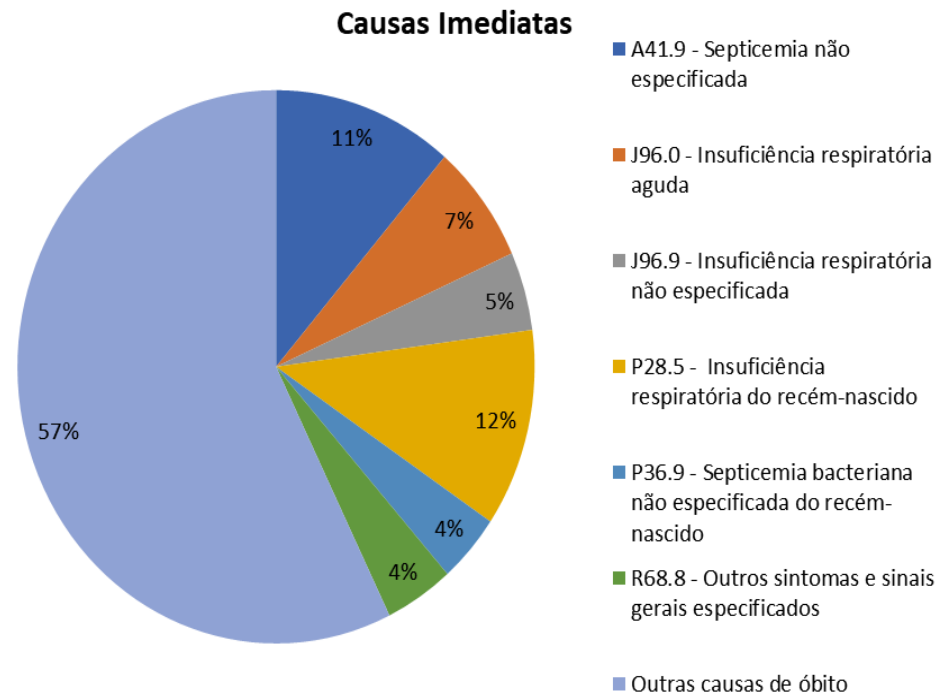


Obs.: Embora Microcefalia tenha sido registrada como causa de morte, ela não é uma doença e, portanto, não deveria ter sido atestada como tal; trata-se de um sinal resultante dos danos produzidos pelo vírus Zika ao sistema nervoso central da criança.

Fonte: Elaborada pelos autores.

Entre as causas imediatas mais atestadas, destacaram-se insuficiência respiratória do recém-nascido (40 óbitos; MP=12,1%), septicemia não especificada (39 óbitos; MP=11,8%) e insuficiência respiratória aguda (24 óbitos; MP=7,3%). Outras causas mal definidas e não especificadas de mortalidade (RMPc=12,0); microcefalia (RMPc=9,0); broncopneumonia não especificada (RMPc=7,5) contribuíram para um aumento relativo  $\geq 7$  vezes como causa imediata de morte entre crianças com SCZ, do que entre crianças com anormalidades do SNC, não relacionadas ao Zika (Figura 19).

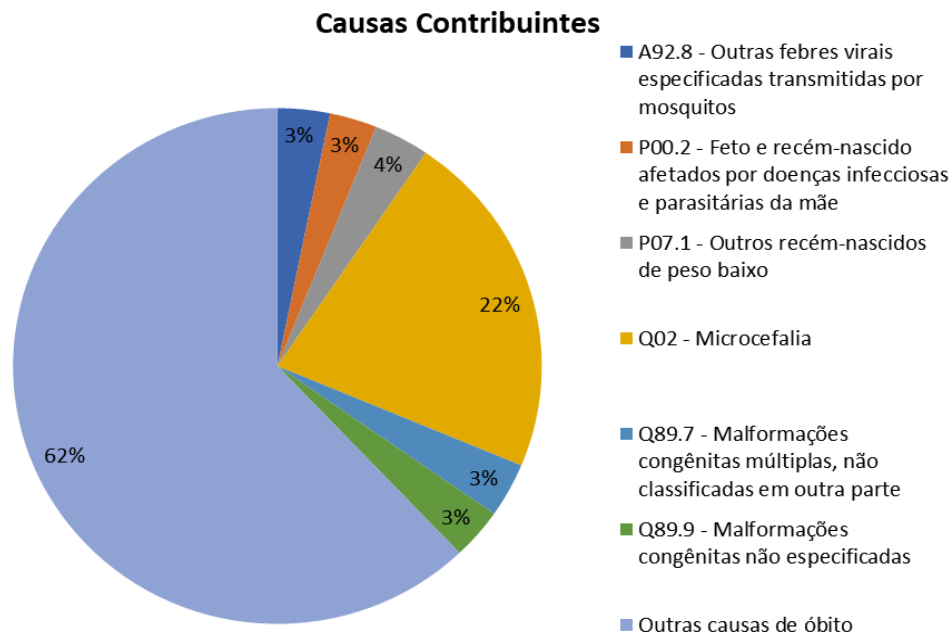
Figura 19: Principais causas imediatas de óbitos de crianças com até 36 meses de vida, nascidas com síndrome congênita do Zika. Brasil, 2015-2018



Fonte: Elaborada pelos autores

Outros recém-nascidos de baixo peso ao nascer e malformações congênitas múltiplas não classificadas em outra parte, foram as causas contribuintes de morte mais atestadas para crianças com SCZ (Figura 20).

Figura 20: Principais causas contribuintes de óbitos de crianças com até 36 meses de vida, nascidas com síndrome congênita do Zika. Brasil, 2015-2018



Obs.: Embora Microcefalia tenha sido registrada como causa de morte, ela não é uma doença e, portanto, não deveria ter sido atestada como tal; trata-se de um sinal resultante dos danos produzidos pelo vírus Zika ao sistema nervoso central da criança.

Fonte: Elaborada pelos autores

Entre suas limitações, encontram-se o fato de que essa investigação se baseou em dados de registro, e informações clínicas adicionais não estavam disponíveis. Além disso, subnotificação de casos no RESP é possível, principalmente entre os fetos expostos ao ZIKV no pré-natal, mas sem malformações detectáveis imediatamente após o nascimento. O processo de vinculação pode ter introduzido viés de classificação incorreta devido a erro de vinculação, quando crianças com SCZ não foram vinculadas a registros de óbitos e, portanto, contribuir para a subnotificação de óbitos. Também houve limitações laboratoriais para confirmação de casos e uso de diversos critérios de confirmação, principalmente no início da epidemia dessa síndrome, além de erros

no preenchimento das causas de óbito na DO. Salienta-se que embora “microcefalia” tenha sido uma das condições mais atestadas como causa básica nas DO das crianças analisadas, aqui ela não está referida pelo fato desta condição não ser uma causa de óbito e sim um sinal resultante das lesões do SNC produzidas pelo agente infeccioso.

Apesar destas limitações, os resultados deste estudo revelam aspectos importantes da SCZ e destacam a maior vulnerabilidade dessas crianças a condições infecciosas e respiratórias. Ao identificar as principais causas que constituem a cadeia de causas de morte das crianças com SCZ nascidas no Brasil, esta investigação aporta subsídios que devem ser considerados no desenvolvimento de protocolos de saúde para que incluam procedimentos e condutas que possam minimizar os efeitos desfavoráveis produzidos pelas principais condições clínicas que elevam a mortalidade dessas crianças.

## PRODUTOS TECNOLÓGICOS DESENVOLVIDOS

### **Estratégia para aperfeiçoar a identificação de casos de SCZ na rede do SUS: Algoritmo de classificação da síndrome congênita do Zika<sup>11</sup>**

Logo após a eclosão da epidemia de microcefalia/SCZ, a SVS/MS deu início à investigação epidemiológica e adaptou imediatamente o RESP para coletar os dados sobre os casos notificados desta nova condição nosológica. Neste formulário eletrônico foram registrados os dados de todas as crianças nascidas com perímetro cefálico reduzido, de acordo com os critérios estabelecidos inicialmente pelo Ministério da Saúde. Embora, no início da epidemia a rede de serviços de saúde não contasse com exames diagnósticos específicos para ZIKV em recém-nascidos, as observações clínicas de casos com perímetro cefálico abaixo do valor de referência revelavam que o SCZ apresentava características diferentes das produzidas por outras infecções (infecções

---

11. Para ver o artigo publicado sobre essa análise, consultar VEIGA; SCHULER-FACCINI; FRANÇA; ANDRADE *et al.*, 2021.

TORCH: toxoplasmose, sífilis, varicela-zoster, parvovírus B19, rubéola, citomegalovírus, herpes). Assim, foi possível estabelecer uma classificação com base em critérios clínico-epidemiológicos. Esta iniciativa foi fundamental para determinar a magnitude da epidemia, sua distribuição geográfica, o espectro clínico da Síndrome, bem como, para planejar as necessidades de saúde das crianças afetadas.

Até agosto de 2017, 14.951 notificações de casos suspeitos de SCZ já haviam sido registradas no RESP. No entanto, cerca de 33,7% destas ainda permaneciam sem classificação final, principalmente devido ao número insuficiente, na rede de serviços de saúde, de testes laboratoriais específicos para Zika e para as TORCH, bem como, de especialistas capazes de discutir e aplicar para cada caso todos os critérios de classificação, que são complexos e abrangem uma grande variedade de informações.

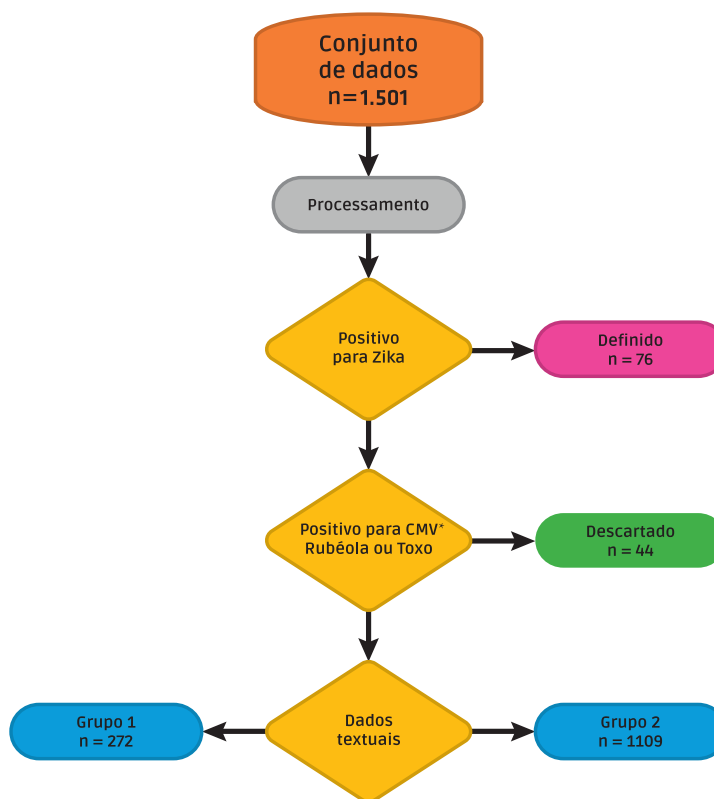
Grande parte das informações usadas para classificar os casos suspeitos de SCZ pelos especialistas baseia-se em dados contidos em textos oriundos de laudos médicos e de exames de imagens. Esses textos apresentam muitas abreviaturas, termos técnicos e diferem muito na forma de escrita entre os médicos. Estas dificuldades demandam sistemas de apoio à decisão, que possibilitem classificar o microcefalia/SCZ sem a necessidade de experiência humana.

Técnicas de Aprendizado de Máquina (AM) podem ser utilizadas na criação de sistemas computacionais que aprendam por meio de treinamento, sendo estes capazes de tomar decisões com o menor erro possível. Inclusive, a AM tem a capacidade de uso de textos por meio da construção de um sistema autônomo para tomada de decisão. Este desafio técnico foi enfrentado por meio da construção de um sistema autônomo capaz de classificar casos suspeitos de SCZ contidos no RESP.

A classificação utilizada foi proposta e definida em trabalho científico anterior (FRANÇA; SCHULER-FACCINI; OLIVEIRA; HENRIQUES *et al.*, 2016): casos descartados; casos pouco prováveis; casos moderadamente prováveis; casos altamente prováveis; e casos definitivos. Na referida publicação, um total de 1.501 casos foram previamente classificados por especialistas em acordo com o algoritmo descrito no esquema

mostrado na Figura 21, onde: 120 (8%) casos foram classificados deterministicamente em descartados e definitivos; 272 (18%) classificados como Grupo 1, não tinham informações textuais e foram classificados como descartados e pouco prováveis, e 1.109 (74%) classificados como Grupo 2, tinham informações textuais.

Figura 21: Fluxograma que descreve o processo de classificação dos casos de Síndrome Congênita do Zika



\*CMV é sorologia positiva para citomegalovírus e TOXO é sorologia positiva para *Toxoplasma gondii*

Fonte: Elaborada pelos autores.

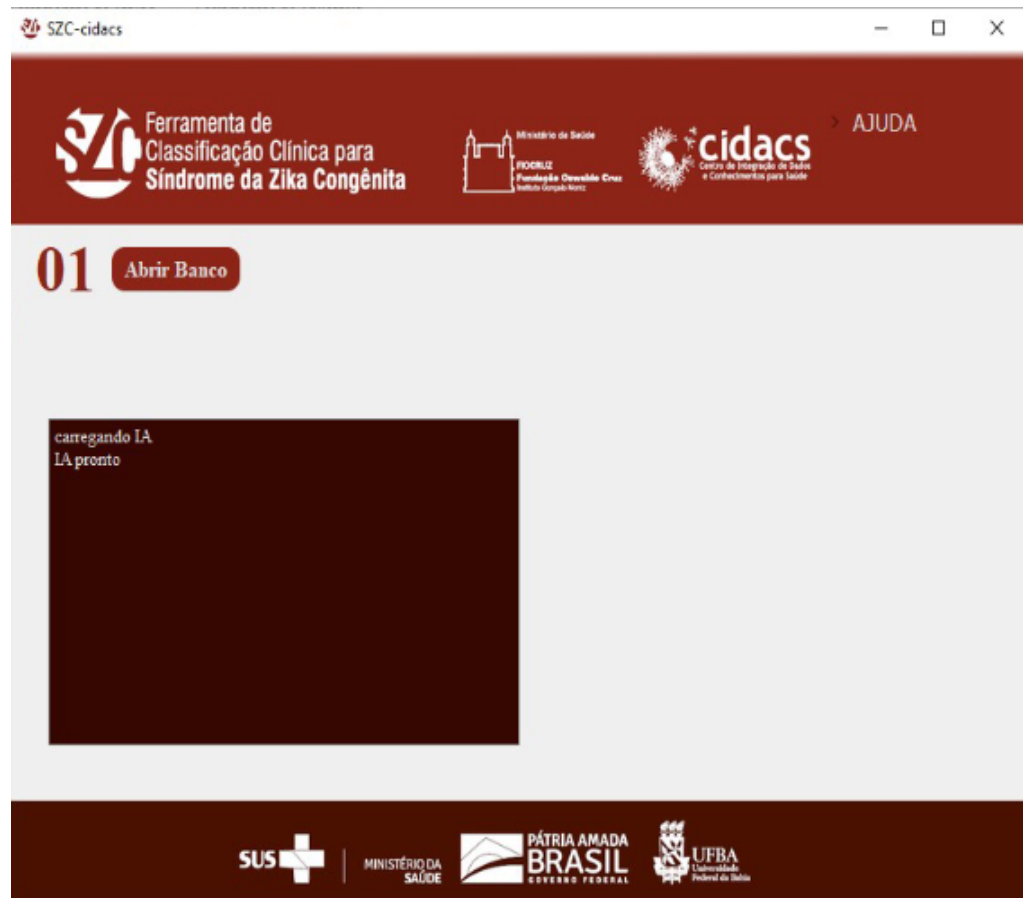


O algoritmo que desenvolvemos foi aplicado aos 1.501 casos e foi capaz de replicar os resultados obtidos pelos especialistas com 97% de acurácia, sendo estimada uma acurácia igual ou superior a 76% para novos casos no Grupo 1 (sem informações textuais), e 83% para novos casos no Grupo 2 (com informações textuais). O algoritmo foi capaz de acertar mais de 90% de casos descartados para novos casos, que constituem a maioria (60%). Assim, foi constatado que este algoritmo, além de ter uma boa acurácia para classificar o risco de desenvolver a SCZ, também poderia ser usado para triagem de casos descartados, a fim de reduzir a demanda de casos a serem avaliados por especialistas humanos.

Durante este estudo, foi possível desenvolver um programa/aplicativo denominado SZC-cidacs com interface gráfica, a fim de auxiliar os gestores e profissionais de saúde a utilizarem este algoritmo (Figura 22). A aplicação deste programa permite:

- Abrir e ler bancos de dados em diferentes formatos;
- Aplicar o programa SZC-cidacs por grupo de risco e realizar uma triagem entre casos confirmados / descartados;
- Obter para cada registro individual o grupo mais provável, e o nível de certeza que o algoritmo tem desta classificação;
- Obter a probabilidade de ser um caso descartado;
- Obter gráficos simples dos resultados da aplicação do algoritmo SZC-cidacs;
- Salvar e exportar o banco de dados contendo os resultados das análises geradas pelo programa SZC-cidacs em diversos formatos.

Figura 22: Exemplo da interface gráfica disponível no aplicativo SZC-Cidacs



Fonte: [bit.ly/SCZ\\_CIDACS](https://bit.ly/SCZ_CIDACS)

Além disso, um tutorial no formato pdf e um vídeo tutorial foram elaborados a fim de auxiliar nas etapas, desde a instalação do aplicativo até o seu uso, bem como, nas possíveis dúvidas que possam vir a existir a respeito do uso do programa<sup>12</sup>.

12. Os tutoriais e o programa encontram-se disponíveis no link: [bit.ly/Tutorial\\_SZC](https://bit.ly/Tutorial_SZC)

Este programa pode ser distribuído aos responsáveis pela vigilância desta síndrome, por consideramos ser da maior importância divulgar este conhecimento para os três níveis do SUS<sup>13</sup>. A metodologia pode ser aplicada aos dados coletados no nível municipal e estadual para imediata classificação dos casos de SCZ.

### **Coorte de Nascimentos do Centro de Integração de Dados e Conhecimentos para Saúde (Cidacs/Fiocruz Bahia)<sup>14</sup>**

Para acompanhar a população de crianças acometidas por Microcefalia/SCZ, foi construída uma coorte de nascimentos usando o banco de dados identificados do SINASC fornecido pela SVS, adotando-se os protocolos éticos e de segurança vigentes, previstos na Lei de Acesso à Informação e Lei Geral de Proteção de Dados Pessoais Conforme apresentado no Capítulo 1. Essa coorte epidemiológica foi concebida para estudar as crianças com SCZ, mas pode ser utilizada para inúmeras outras exposições durante o período pré-natal, nascimento e ao longo da vida. Importante destacar que essa coorte está vinculada à Coorte de 100.000 milhões de brasileiros do Cidacs (BARRETO; ICHIHARA; PESCARINI; ALI *et al.*, 2022).

A Coorte de Nascimentos do Cidacs/Fiocruz Bahia é constituída por uma de base populacional derivada de dados nacionais vinculados, desenvolvida para investigar as relações entre eventos pré-natais e precoces e desfechos relacionados à saúde. A população desta Coorte é composta por 28.631.390 de nascidos vivos no Brasil entre 1º de janeiro de 2001 e 31 de dezembro de 2018, que se vinculou a partir de informações maternas à linha de base da Coorte de 100 Milhões de Brasileiros. As informações referem-se às crianças desde o nascimento até os 17 anos de vida, e já estão sendo atualizadas até o ano de 2020.

O *linkage* foi realizado por meio de uma ferramenta denominada CIDACS-RL (ALMEIDA; GORENDER; ICHIHARA; SENA *et al.*, 2020), utilizando o nome município e data de

13. Mais detalhes relacionados ao algoritmo e sua validação podem ser encontrados no link: [https://github.com/cidacslab/CIDACS-Zika\\_Classification](https://github.com/cidacslab/CIDACS-Zika_Classification)

14. Para ver o artigo publicado sobre essa análise, consultar PAIXÃO; CARDIM; FALCÃO; ORTELAN *et al.*, 2021.

nascimento, ou idade da mãe. O *link* entre o arquivo composto SINASC-100 Million Brazilian Cohort com outros conjuntos de dados (mortalidade, hospitalização, doenças infecciosas e estado nutricional) fornecerá o acompanhamento da Coorte e informações sobre desfechos.

A construção desta Coorte demonstra o potencial do Cidacs/Fiocruz Bahia para promover o uso das estratégias de vinculação de bancos de dados, implementadas para explorar os efeitos a longo prazo da microcefalia/SCZ na morbidade e mortalidade das crianças afetadas, bem como, a aplicação para estudos de outros problemas de saúde pública.

### **Ferramenta para extrair, transformar e analisar dados de sistemas de informação em saúde do Brasil: pyBIS<sup>15</sup>**

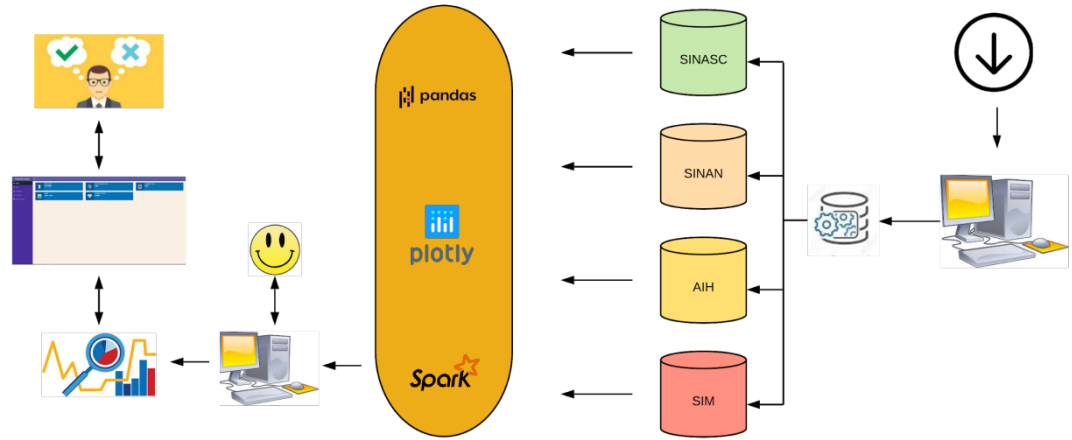
Um dos objetivos da Vigilância em Saúde é monitorar a situação epidemiológica das doenças de importância para a saúde pública nos municípios, bem como, estudar seus determinantes sociais. Visando agilizar e aperfeiçoar este processo, foi desenvolvido um *software* denominado pyBIS, que possibilita a articulação de um conjunto de funções, e permite editar e transformar dados secundários de diferentes bases públicas disponibilizados pelo governo brasileiro, de maneira fácil e rápida, por profissionais que não são especialistas em programação (Figura 23).

As funções do pyBIS foram desenhadas para trabalhar com dados tabulares disponíveis nos principais sistemas de informação em saúde do Brasil (SIM, SINASC, AIH, SINAN) e demográficas (IBGE). Ademais, outras bases de dados oriundas de sistemas de informações, oficiais e não oficiais, poderão ser facilmente incorporadas em sua rotina. Este *software* fornece um conjunto de operações capazes de realizar de maneira eficiente e automatizada, a extração, transformação e carregamento de grandes volumes de dados, além de possibilitar a integração de diferentes bases através da utilização de algoritmos determinísticos.

---

15. Este produto encontra-se em fase de teste e pode ser acessado em [https://github.com/cidacslab/pyBIS\\_cidacs](https://github.com/cidacslab/pyBIS_cidacs). Um breve tutorial sobre como instalar e usar o software encontra-se em <https://bit.ly/instalacaopybis>

Figura 23: Representação esquemática das fases de funcionamento do pyBIS



Fonte: [https://github.com/cidacslab/pyBIS\\_cidacs](https://github.com/cidacslab/pyBIS_cidacs)

O processo de funcionamento do pyBIS pode ser realizado em três etapas. A primeira consiste na leitura de dados, que é feita por meio da conexão ao serviço de transferência de arquivos abertos disponibilizados pelo Datasus, ou outros repositórios de dados públicos. Além disso, é oferecido ao usuário a utilização de dados próprios, desde que se encontre em formato tabular. A segunda etapa consiste na manipulação desses dados. Para isso, os dados ingeridos são convertidos em *dataframes* e manipulados via *pandas* (baixo volume de dados) ou *spark* (*big data*). Durante esse processo, é permitido ao usuário interagir com os dados de maneira que consiga executar tarefas como filtrar e/ou selecionar variáveis. Já o terceiro processo consiste na análise dos dados selecionados. Até o momento são oferecidas duas análises para o usuário. Uma análise temporal, que permite verificar alterações na dinâmica temporal de um determinado desfecho, e uma análise espaço-temporal, que possibilita verificar variações no tempo e no espaço de um determinado desfecho.

Tendo em vista que esta versão já é passível de ser operada, a equipe do Cidacs/Fiocruz Bahia apresentou a metodologia para técnicos e dirigentes da Secretaria de Vigilância em Saúde, ao tempo em que disponibilizou acesso público à ferramenta.

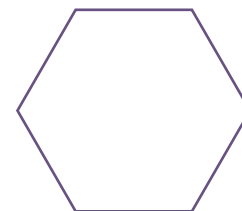
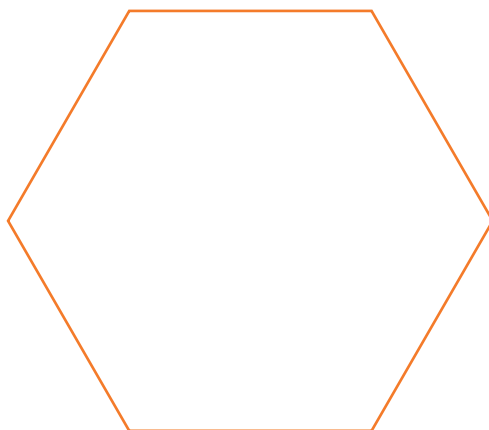
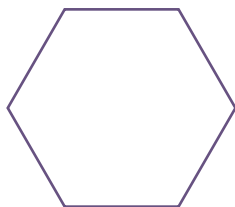
Importante destacar que uma das suas aplicações práticas é facilitar e aprimorar a realização das análises epidemiológicas das arboviroses, doenças mais afeitas aos objetivos desta Plataforma. Contudo, pode ser aplicado a qualquer doença ou agravo, pois tais análises fazem parte das atividades rotineiras das equipes técnicas da SVSA e das Secretarias de Saúde Estaduais e Municipais, inclusive para serem divulgadas nos respectivos Boletins Epidemiológicos. Assim, este *software* constitui uma ferramenta muito especial e útil ao SUS, na medida em que, a partir dos algoritmos, torna muito mais rápida a realização destas análises.

## CONCLUSÕES

Os trabalhos desenvolvidos pelos pesquisadores da Plataforma Zika não deixam dúvidas quanto a sua contribuição para a ampliação dos conhecimentos sobre a epidemia da Zika e seus desdobramentos no Brasil. Foi evidenciado que no curso dessa epidemia, os casos notificados de Zika estavam subregistrados, visto terem sido confundidos com dengue. Devido à eclosão da epidemia de microcefalia/SCZ, foi realizado o aprimoramento do instrumento de registro de informações sobre nascidos vivos, iniciativa que impactou positivamente na notificação de todas as anomalias congênitas (AC), demonstrando que a prevalência nacional destes agravos vinha sendo subestimada. Estudos igualmente relevantes demonstraram que imagens anormais do cérebro e tamanho reduzido do perímetro cefálico parecem ser os principais fatores para o aumento da mortalidade das crianças com SCZ, pelo menos nos três primeiros anos de vida, sendo o risco de morte 11,3 vezes maior entre os nascidos vivos com esta síndrome, do que entre aqueles sem esta condição, possivelmente devido à sua maior vulnerabilidade à desnutrição e enfermidades infecciosas e respiratórias, as quais constituíram nas principais causas de morte destas crianças.

Portanto, não restam dúvidas de que, além de cumprir o seu objetivo, a Plataforma Zika deixou um importante legado para subsidiar os profissionais de saúde na elaboração de protocolos de vigilância e de atenção à saúde das crianças com SCZ. Acresce-se a isto, o fato de que, o desenvolvimento de um programa para aprimoramento do processo de classificação de casos desta síndrome representa um instrumento sensível para o seu diagnóstico logo após o nascimento e, assim, também contribui para fortalecer o alerta sobre a necessidade de adoção das medidas de prevenção e controle da Zika, bem como, para a atenção voltada para minimizar os efeitos adversos na qualidade de vida e saúde dos portadores desta grave síndrome.

Em que pesem os esforços que a comunidade científica vem empreendendo para o avanço dos conhecimentos sobre estes problemas de saúde emergentes, considera-se da maior relevância a continuidade de estudos longitudinais de base populacional, por maior período de tempo, visando conhecer a história natural da SCZ ao longo da vida, como também, para contribuir para definição de ações governamentais e não governamentais de saúde e sociais a serem desenvolvidas para apoiar as crianças acometidas e suas famílias.





## ARTIGOS CIENTÍFICOS E PRODUTOS RESULTANTES

CENTRO DE INTEGRAÇÃO DE DADOS E CONHECIMENTOS PARA SAÚDE. pyBIS: **Uma interface gráfica para trabalhar com dados dos Sistemas de informação Brasileiros**. Salvador, 2022. Disponível em GitHub - cidacslab/pyBIS\_cidacs. Acesso em 24 Jul 2023.

COSTA, LC; VEIGA, RV; OLIVEIRA; JF, RODRIGUES; MS *et al.* New insights on the zika virus arrival in the Americas and spatiotemporal reconstruction of the epidemic dynamics in Brazil. **Viruses**, v.13, n.1, p. 1-11, 2021.

COSTA, MCN; CARDIM, LL; MOORE, CA; JESUS, ES; CARVALHO-SAUER, R; BARRETO, ML *et al.* (2023) Causes of death in children with congenital Zika syndrome in Brazil, 2015 to 2018: A nationwide record linkage study. **PLoS Med** v.20, n.2, p.1-17, 2023. <https://doi.org/10.1371/journal.pmed.1004181>

COSTA, MCN; CARDIM, LL; TEIXEIRA, MG; BARRETO, ML *et al.* Case fatality rate related to microcephaly congenital zika syndrome and associated factors: A nationwide retrospective study in Brazil. **Viruses**, v.1, n.11, p.1–12, 2020.

FERNANDES, QHRF; PAIXÃO, ES; COSTA, MCN; TEIXEIRA, MG. Temporal trends in prevalence and infant mortality of birth defects in Brazil, from 2001 to 2018. **Ciênc. saúde coletiva**, v. 28, n. 4, p. 969-979, 2023.

OLIVEIRA, JF; RODRIGUES, MS; SKALINSKI, LM; SANTOS, AES *et al.* Interdependence between confirmed and discarded cases of dengue, chikungunya and Zika viruses in Brazil: A multivariate time-series analysis. **PLoS One**, v.15, n. 2, p. 1–13, 2020. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0228347>

PAIXAO, ES; RODRIGUES, MS; CARDIM, LL; OLIVEIRA, JF *et al.* Impact evaluation of Zika epidemic on congenital anomalies registration in Brazil: An interrupted time series analysis. **PLoS Negl. Trop. Dis.**, v. 13, n. 9, p. 1-12, 2019.

PAIXÃO, ES; CARDIM, LL; FALCÃO, IR; ORTELAN, N *et al.* Cohort Profile: Centro de Integração de Dados e Conhecimentos para Saúde (CIDACS) Birth Cohort. **Int J Epidemiol.**, v. 50, n.1, p.37–38, 2021.

PAIXÃO, ES; CARDIM, LL; COSTA, MCN; BRICKLEY, EB *et al.* Mortality from Congenital Zika Syndrome - Nationwide Cohort Study in Brazil. **N. Engl. J. Med.**, v. 386, n. 8, p.757–67, 2022.

PAIXÃO, ES; FERNANDES, QHRF; CARDIM, LL; PESCARINI, JM *et al.* Socioeconomic risk markers of congenital Zika syndrome: a nationwide, registry-based study in Brazil. **BMJ Global Health** v. 7, p.1–10, 2022. <http://dx.doi.org/10.1136/bmjgh-2022-009600>.

PAIXÃO, ES; RODRIGUES, LC; COSTA, MCN; CARVALHO-SAUER, RCO *et al.* Population-based surveillance for congenital Zika virus syndrome: a latent class analysis of recorded cases from 2015–2018. **BMC Pregnancy Childbirth**, v. 22, n. 1, p.1–10, 2022.

VEIGA, RV; SCHULER-FACCINI, L; FRANÇA, GVA; ANDRADE, RFS *et al.* Classification algorithm for congenital Zika Syndrome: characterizations, diagnosis and validation. **Sci. Rep.**, v. 11, n.1, p.1–7, 2021.

## REFERÊNCIAS

ALLIK, M; LEYLAND, A; ICHIHARA, MY; DUNDAS, R. Creating small-area deprivation indices: A guide for stages and options. **J. Epidemiol. Community Health**, v. 74, n. 1, p. 20–5, 2020.

ALMEIDA, D; GORENDER, D; ICHIHARA, MY; SENA, S *et al.* Examining the quality of record linkage process using nationwide Brazilian administrative databases to build a large birth cohort. **BMC Med Inform Decis Mak.**, v.20, n.1, p.1–9, 2020.

BARRETO, ML; ICHIHARA, MY; PESCARINI, JM; ALI, MS *et al.* Cohort Profile: The 100 Million Brazilian Cohort. **Int. J. Epidemiol.**, v. 51, n.2, e27–e38, 2022.

BRASIL. Agência Nacional de Vigilância Sanitária. **Regulamento Sanitário Internacional – RSI 2005**. Versão em Português, 2009. Disponível em: <https://www.gov.br/anvisa/pt-br/assuntos/paf/regulamento-sanitario-internacional/arquivos/7181json-file-1>. Acesso em: 01 nov 2022.

BRASIL. Ministério da Saúde. Portaria nº 1.813, de 11 de novembro de 2015. **Declara emergência em saúde pública de importância nacional por alteração de ocorrência de microcefalia no Brasil.** Disponível em: [https://bvsmms.saude.gov.br/bvs/saudelegis/gm/2015/prt1813\\_11\\_11\\_2015.html](https://bvsmms.saude.gov.br/bvs/saudelegis/gm/2015/prt1813_11_11_2015.html). Acesso em: 01 nov 2022.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Secretaria de Atenção à Saúde. **Orientações integradas de vigilância e atenção à saúde no âmbito da Emergência de Saúde Pública de importância nacional:** procedimentos para o monitoramento das alterações no crescimento e desenvolvimento a partir da gestação até a primeira infância, relacionadas à infecção pelo vírus Zika e outras etiologias infecciosas dentro da capacidade operacional do SUS. Brasília: Ministério da Saúde. 2017. 158p. Disponível em: <http://portalarquivos.saude.gov.br/images/pdf/2016/dezembro/12/orientacoes-integradas-vigilancia-atencao.pdf>. Acesso em: 27 nov. 2022.

BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Análise em Saúde e Vigilância de Doenças Não Transmissíveis. **Saúde Brasil 2020/2021:** anomalias congênitas prioritárias para a vigilância ao nascimento. Brasília: Ministério da Saúde, 2021. 414p.

CDC. Centers for Disease Control and Prevention. **Update on overall prevalence of major birth defects-Atlanta, Georgia, 1978-2005.** MMWR Morb Mortal Wkly Rep, 57, (1, p.1), p.5-11 Jan. 2008. PMID: 18185492. Disponível em: <https://www.cdc.gov/mmwr/preview/mmwrhtml/mm5701a2.htm>. Acesso em: 20 ago 2022.

CUNHA, AJLA; MAGALHÃES-BARBOSA, MC; LIMA-SETTA, F; MEDRONHO, RA; PRATA-BARBOSA, A. Microcephaly case fatality rate associated with Zika virus infection in Brazil. **Pediatr. Infect. Dis J.**, v. 36, n. 5, p.528–530, 2016.

FARIA, NR; AZEVEDO, RSS; KRAEMER, MUG; SOUZA, R *et al.* Zika virus in the Americas: Early epidemiological and genetic findings. **Science**, v. 352, n. 6283, p. 345–9, 2016.

FARIA, NR; QUICK, J; CLARO, IM; THÉZÉ, J *et al.* Establishment and cryptic transmission of Zika virus in Brazil and the Americas. **Nature**, v. 546, n. 7658, p. 406–10, 2017.

FOY, BD; KOBYLINSKI, KC; FOY, JLC; BLITVICH, BJ, *et al.* Probable Non-Vector-borne Transmission of Zika Virus. **Emerg. Infect Dis.**, Colorado, USA, v. 17, n. 5, p. 880–2, 2011.

FRANÇA, GVA; SCHULER-FACCINI, L; OLIVEIRA, WK; HENRIQUES, CMP *et al.* Congenital Zika virus syndrome in Brazil: a case series of the first 1501 livebirths with complete investigation. **Lancet**, v. 388, n. 10047, p. 891–7, 2016.

GRUBAUGH, ND; LADNER, JT; KRAEMER, MUG; DUDAS, G *et al.* Genomic epidemiology reveals multiple introductions of Zika virus into the United States. **Nature**, v. 546, n. 7658, p.401–405, 2017.

HEANG, V; YASUDA, CY; SOVANN, L; HADDOW, AD *et al.* Zika virus infection, Cambodia, **Emerg. Infect. Dis.**, v. 18, n. 2, p. 349–51, 2012.

HEUKELBACH, J; ALENCAR, CH; KELVIN, AA; OLIVEIRA, WK; GÓES, LP. Zika virus outbreak in Brazil. **J. Infect. Dev. Ctries**, v. 10, n.2, p.116–20, 2016.

ICHIHARA, MY; RAMOS, D; REBOUÇAS, P; OLIVEIRA, FJ *et al.* Area deprivation measures used in Brazil: A scoping review. **Rev. Saúde Pública**, v. 52, n. 83, p.1–14, 2018.

KARWOWSKI, MP; NELSON, JM; STAPLES, JE; FISCHER, M *et al.* Zika virus disease: A CDC update for pediatric health care providers. **Pediatrics**, v. 137, n. 5, p.1–13, 2016.

MENDES NETO, NN; MAIA, JTS; ZACARKIM, MR; QUEIROZ, I; LABEAUD, AD. Perinatal Case Fatality Rate Related to Congenital Zika Syndrome in Brazil: A Cross-Sectional **Study. Pediatr. Neurol.**, v. 81, p. 47-48, 2018.

METSKY, HC; MATRANGA, CB; WOHL, S; SCHAFFNER, SF *et al.* Zika virus evolution and spread in the Americas. **Nature**, v. 546, n. 7658, p. 411–5, 2017.

MOORE, CA; STAPLES, JE; DOBYNS, WB; PESSOA, A *et al.* Characterizing the pattern of anomalies in congenital zika syndrome for pediatric clinicians. **JAMA Pediatr**, v. 171, n. 3, p.288–95, 2017.

PAIXÃO, ES; BARRETO, F; TEIXEIRA, MG; COSTA, MCN; RODRIGUES, LC. History, epidemiology, and clinical manifestations of Zika: A systematic review. **Am. J. Public Health**, v. 106, n. 4, p.606–12, 2016.

RASMUSSEN, SA; JAMIESON, DJ; HONEIN, MA; PETERSEN, LR. Zika Virus and Birth Defects – Reviewing the Evidence for Causality. **N. Engl. J. Med.**, v. 374, n. 20, p.1981–7, 2016.

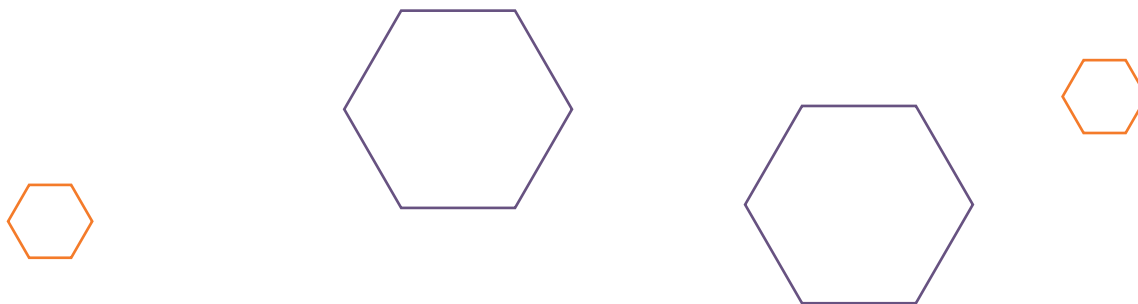
SHERIDAN, E; WRIGHT, J; SMALL, N; CORRY, PC *et al.* Risk factors for congenital anomaly in a multiethnic birth cohort: An analysis of the Born in Bradford study. **Lancet**, v. 382, n. 9901, p.1350–9, 2013.

TEIXEIRA, MG; COSTA, MCN; OLIVEIRA, WK; NUNES, ML; RODRIGUES, LC. The epidemic of Zika virus-related microcephaly in Brazil: Detection, control, etiology, and future scenarios. **Am. J. Public Health**, v. 106, n. 4, p. 601–5, 2016.

THÉZÉ, J; LI, T; DU PLESSIS, L; BOUQUET, J *et al.* Genomic Epidemiology Reconstructs the Introduction and Spread of Zika Virus in Central America and Mexico. **Cell Host Microbe**, v. 23, n. 6, p. 855-864, 2018.

WHO. World Health Organization. **Statement on the 2nd meeting of IHR Emergency Committee on Zika virus and observed increase in neurological disorders and neonatal malformations.** 2016. Disponível em: <https://www.who.int/news/item/08-03-2016-who-statement-on-the-2nd-meeting-of-ihf-emergency-committee-on-zika-virus-and-observed-increase-in-neurological-disorders-and-neonatal-malformations>. Acesso em: 15 set. 2020.

WHO. World Health Organization. **Congenital anomalies.** [s.d.]. Disponível em: [https://www.who.int/health-topics/congenital-anomalies#tab=tab\\_1](https://www.who.int/health-topics/congenital-anomalies#tab=tab_1). Acesso em: 15 set. 2020.





# Prospecção das iniciativas de pesquisa sobre controle e tratamento de arboviroses

---

*Ricardo Barros Sampaio  
Bruno Justino Garcia Praciano  
Bruno Santos Ferreira  
Fábio Castro Gouveia  
Marcos Aires  
Micael Filipe Ribeiro de Lima*

Este componente dedica-se à prospecção de pesquisas, o que está acontecendo de mais atual sobre o tema, fornece elementos aos epidemiologistas da Plataforma Zika para aprimoramento das hipóteses, e apoia na identificação de grupos de colaboração.

O objetivo desta meta foi avaliar a produção científica e demais iniciativas de pesquisa nas áreas de controle e tratamento das arboviroses. Para isso, foi realizada uma análise dos pesquisadores envolvidos com o tema dentro do contexto brasileiro e da produção científica nacional e internacional. As competências e ferramentas disponibilizadas permitem uma avaliação detalhada do universo de interesse.

A proposta de pesquisa foi centrada no uso da Cientometria como ferramenta para o mapeamento da pesquisa em Zika e dos pesquisadores envolvidos no tema da tríplice epidemia com foco em Zika. As bases de dados utilizadas para as pesquisas foram as bases internacionais de publicação científica, com acesso pelo portal de Periódicos Capes, e a Plataforma Lattes como base curricular.

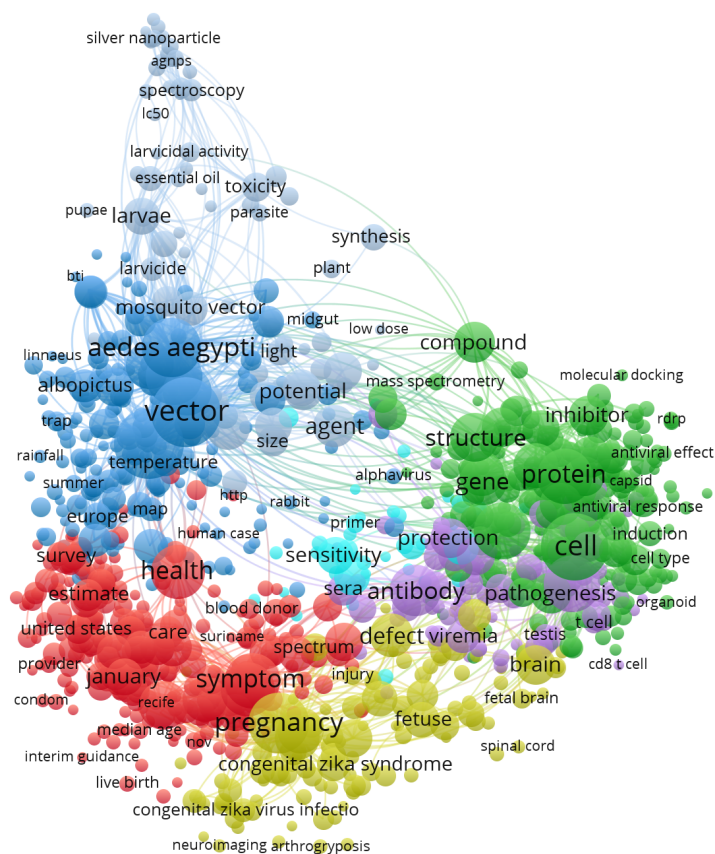
Além da pesquisa, foi realizado o desenvolvimento da ferramenta e-Lattes, e estabelecidas parcerias com instituições de pesquisa que trabalham com métodos similares, resultando no uso e na construção de novas bases.

## **ANÁLISE DA PRODUÇÃO CIENTÍFICA**

Os estudos cientométricos são um campo de pesquisa para diversas áreas do conhecimento, e têm sido utilizados com sucesso para analisar a construção do conhecimento, a prática e a divulgação científica. Neste estudo, buscamos verificar como a epidemia de Zika refletiu-se nos artigos científicos publicados em todo o mundo, analisando colaborações internacionais e o panorama do conhecimento ao longo do tempo, bem como, tópicos de pesquisa e envolvimento de pesquisadores e de diferentes países.

Foram trabalhados vários aspectos da produção científica em Zika com a geração de mapas como o apresentado na Figura 1, além de pesquisa sobre a produção científica internacional e nacional. Abaixo, apresentamos alguns dos nossos achados começando pelos resultados de artigo científico publicado (OLIVEIRA *et. al.*, 2020), seguido de outras produções e contribuições em eventos científicos.

Figura 1: Mapa do conhecimento da Zika, produzido a partir dos títulos e resumos dos artigos científicos da base *Web of Science*



O mapa da Figura 1, produzido pelo *software* VOS-viewer<sup>1</sup>, é um mapa do conhecimento que concatena em uma mesma figura os termos mais relevantes dos títulos e resumos de artigos científicos encontrados sobre o tema. As diferentes cores são atribuídas por um algoritmo de agrupamento. Onde há maior interação (número de linhas entre dois nós aqui representando palavras), grupos se formam representando áreas específicas de conhecimento.

Fonte: Elaboração dos autores.

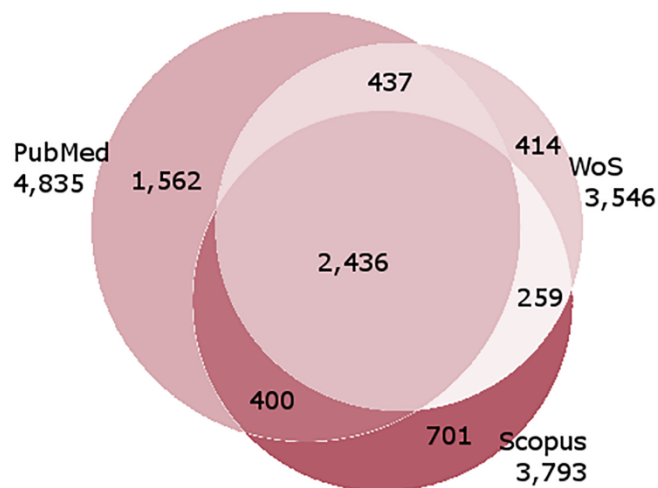
1. Disponível em <http://www.vosviewer.com>.



## A resposta global da pesquisa científica à emergência de saúde pública resultante da infecção pelo vírus Zika

Pesquisamos na *Web of Science* (WoS), Scopus e PubMed por estudos publicados até 31 de dezembro de 2018 sobre Zika, usando os termos de pesquisa “zika”, “zkv” ou “zikh”. Analisamos a produção científica sobre quais países mais publicaram, sobre quais tópicos, bem como, a colaboração de cada país. Realizamos uma análise cientométrica de pesquisas sobre Zika com foco no mapeamento do conhecimento e na trajetória da pesquisa científica ao longo do tempo e do espaço. A Figura 2 apresenta um diagrama de Venn com a sobreposição das amostras coletadas das três diferentes bases.

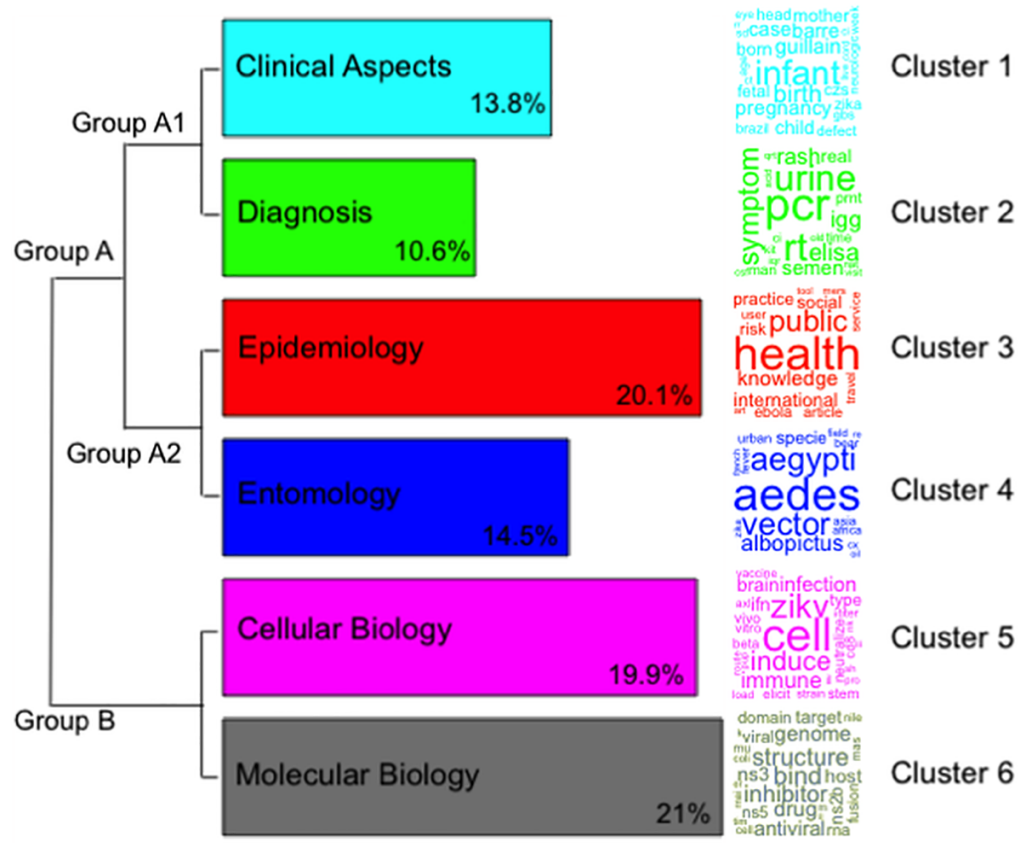
Figura 2: Artigos sobre Zika identificados na base de dados PubMed, WoS e Scopus publicados de 1945 até o final de dezembro de 2018



Fonte: Elaboração dos autores.

Como descobertas, elencamos que duas áreas de pesquisa bem definidas foram encontradas, divididas em três subtópicos, representando seis agrupamentos (*clusters*). Na Figura 3 temos o dendrograma obtido por intermédio da ferramenta IRAMuTeQ<sup>2</sup> através da qual estas seis temáticas foram identificadas.

Figura 3: Dendrograma dos *clusters* obtidos com IRAMuTeQ



Fonte: Elaboração dos autores.

2. Disponível em <http://www.iramuteq.org>

Com relação à análise dos países, EUA e Brasil foram aqueles com maior número de publicações sobre Zika. A China entrou como um novo participante com foco em áreas específicas de pesquisa. Na Tabela 1, é apresentada a quantidade de artigos por país de autoria e o percentual destes produzidos em colaboração, ou não, com outros países.

Tabela 1: Publicações nacionais e internacionais sobre a infecção pelo vírus Zika nos 10 países mais produtivos

Countries	Articles Total*		Articles with single country author(s)**		Articles with international collaboration)**	
	N	%	N	%	N	%
<b>USA</b>	1718	48.52	983	57.22	735	42.78
<b>Brazil</b>	648	18.27	307	47.38	341	52.62
<b>China</b>	279	7.87	128	45.88	151	54.12
<b>France</b>	264	7.44	70	26.52	194	73.48
<b>England</b>	236	6.66	36	15.25	200	84.75
<b>Germany</b>	186	5.24	47	25.27	139	74.73
<b>Italy</b>	176	4.98	43	24.43	133	75.57
<b>India</b>	151	4.28	55	36.42	96	63.58
<b>Australia</b>	148	4.17	67	45.27	81	54.73
<b>Canada</b>	142	4.02	45	31.7	97	68.3

\* The column “Articles Total” represents the total number of articles published by a country and the percentage of these productions over the total number of articles about Zika found in WoS until 2018.

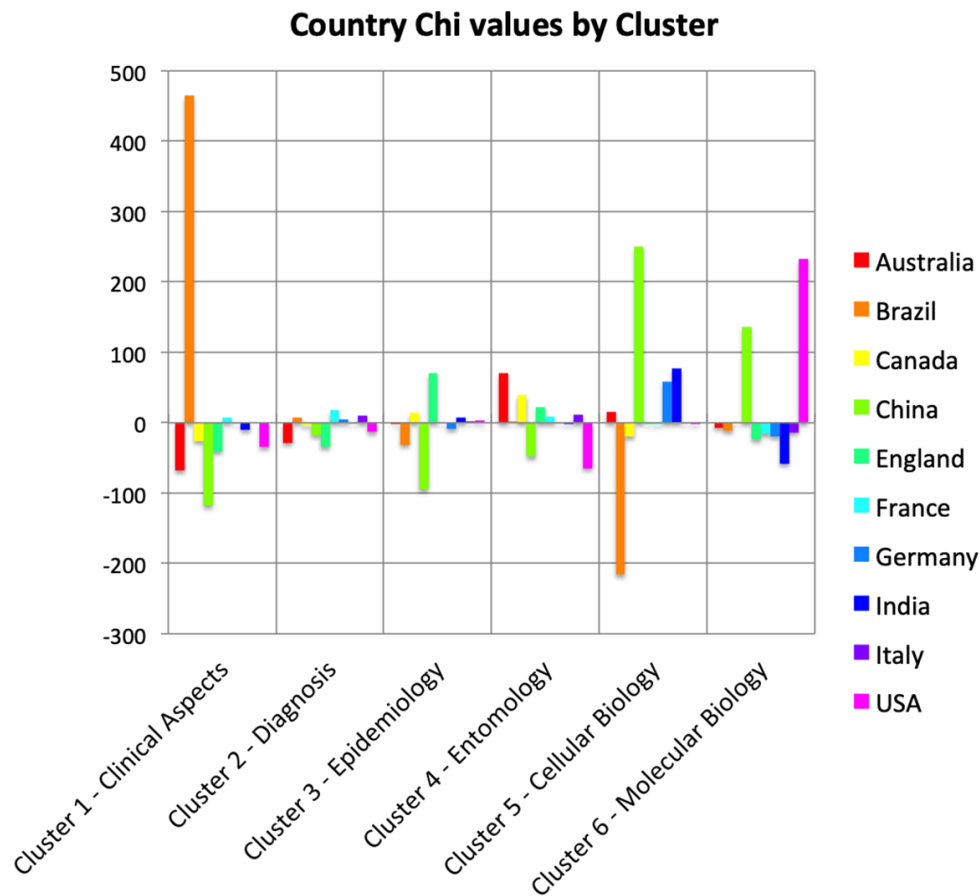
\*\* Among these values, we see the number of articles produced by the country itself and the proportion produced with international collaboration.

Fonte: Elaboração dos autores.

Quando levamos em consideração as epidemias e os casos notificados, Brasil e França foram os principais países de pesquisa em temas relacionados. Quanto à colaboração internacional, destacam-se os EUA seguidos da Inglaterra e da França como os principais concentradores (*hubs*). Anteriormente a 2015 até o início de 2016, as áreas de pesquisa mais publicadas incluíram tópicos relacionados à saúde pública, seguidos por um aumento de tópicos relacionados aos aspectos clínicos da doença em 2016, e

o surgimento da pesquisa laboratorial em 2017 e 2018. Essa distribuição por foco temático pode ser verificada nos valores de Chi para os países em cada tema conforme Figura 4. Os valores de Chi se referem a um teste estatístico em que quanto maior o valor positivo, mais forte é a relação entre a temática e o país em questão, enquanto que valores mais negativos indicam uma relação menor.

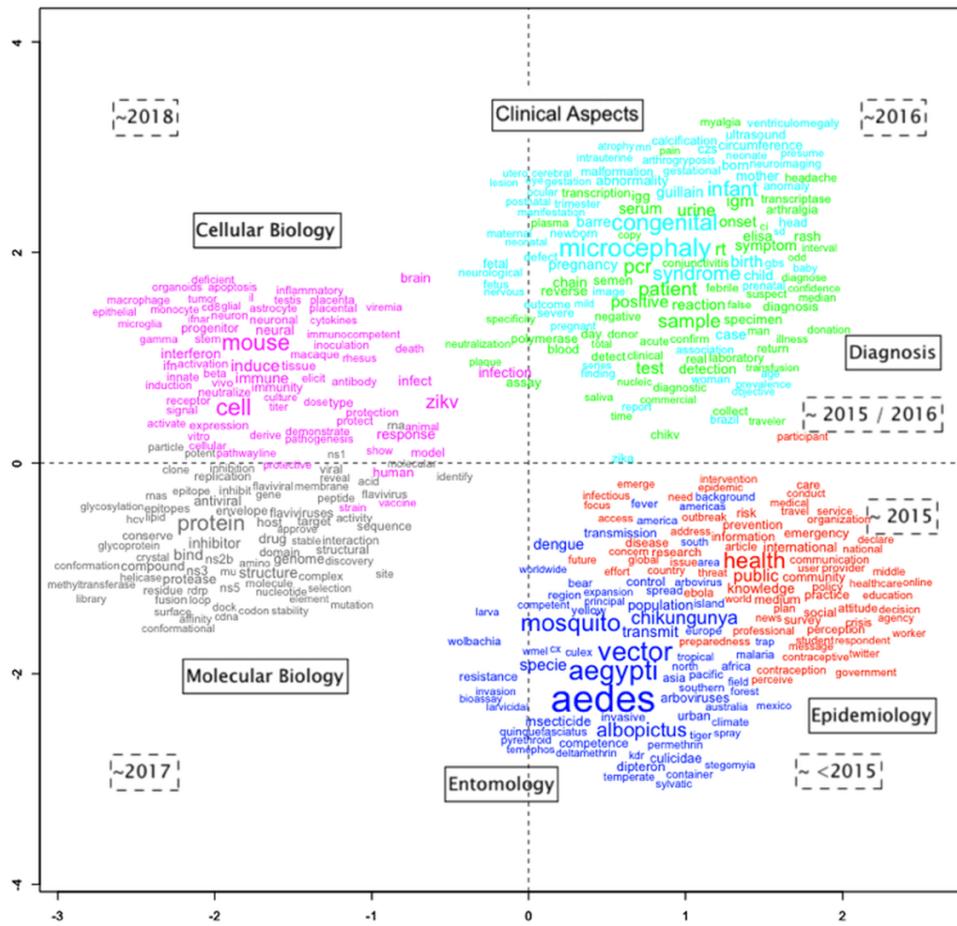
Figura 4: Valores de Chi para os 10 principais países de cada agrupamento temático



Fonte: Elaboração dos autores.

Ao fazermos um mapeamento das temáticas em um plano bidimensional, temos a distribuição apresentada na Figura 5, onde demarcamos os períodos de maior ênfase nos estudos por tema, e observamos as proximidades entre os campos de Diagnóstico e Aspectos Clínicos, Epidemiologia e Entomologia, respectivamente.

Figura 5: Mapa de conhecimento de palavras extraídas de títulos e resumos de artigos na WoS relacionados à infecção pelo vírus Zika, de janeiro de 1945 a dezembro de 2018



Fonte: Elaboração dos autores.

O mapeamento da resposta ao Zika, uma emergência de saúde pública, demonstrou um padrão claro da participação dos países nos avanços científicos. O padrão de produção de conhecimento encontrado neste estudo representou diferentes perspectivas por país, capacidade de pesquisa e interesses, baseados primeiro em seu nível de exposição à epidemia e, segundo, em suas posições no financiamento à ciência.

### **Uma análise cientométrica de pesquisadores do Zika Vírus com base em bases de dados curriculares científicos brasileiros: um estudo de caso do software e-Lattes**

A plataforma Lattes, que pode ser acessada por meio de seu site ([lattes.cnpq.br](http://lattes.cnpq.br)), permite que o usuário faça uma busca utilizando o nome de um pesquisador, ou palavras específicas. Uma busca feita em 1º de fevereiro de 2019, usando o operador booleano (or) com as palavras “zika”, “zikh” e “zkh” retornou um total de 9.178 nomes. Desses, focamos primeiro nos 3.188 que tinham doutorado. Na etapa seguinte foram direcionados aqueles que possuíam pelo menos uma publicação bibliográfica em seu currículo, podendo ser um artigo, um trabalho de congresso científico, um livro ou um capítulo de livro, resultando num total de 1.594 CVs. Como o CV Lattes está aberto para edição dos próprios pesquisadores, eles podem incluir qualquer informação que julguem importante.

Uma vez definido o universo de nossa pesquisa, decidimos utilizar todos os dados disponíveis sobre esses Currículos Lattes incluindo Dados Gerais, Áreas de Pesquisa, Publicações, Orientações Acadêmicas e Redes de Colaboração. Para este grupo específico não encontramos patentes, ou outros desenvolvimentos tecnológicos registrados em seus currículos. Na Tabela 2, temos a área, subárea e especialidade de pesquisa principal elencada pelo pesquisador em seu CV Lattes.

Tabela 2: Número de pesquisadores por área, subárea e especialidade de pesquisa segundo a Plataforma Lattes

<i>Area</i>	#	<i>Subarea</i>	#	<i>Specialty</i>	#
Medicine	359	Molecular Biology	187	Virology	110
Public Health	159	Public Health	186	Infections Disease	83
Microbiology	126	Epidemiology	181	Medical Microbiology	33
Biochemistry	101	Medical Clinic	175	Paediatrics	29
Immunology	87	Biology and Microorganisms	122	Neurology	24
Genetics	85	Maternal and child health	108	Proteins	21
Nursing	51	Molecular Genetic	92	Molecular Biology	17
Parasitology	43	Cellular Immunology	76	Animals Infections Disease	15
Physiotherapy	37	Applied Immunology	66	Human Protozoology	13
Pharmacy	36	Applied Microbiology	64	Neurophysiology	13

Fonte: Elaboração dos autores.

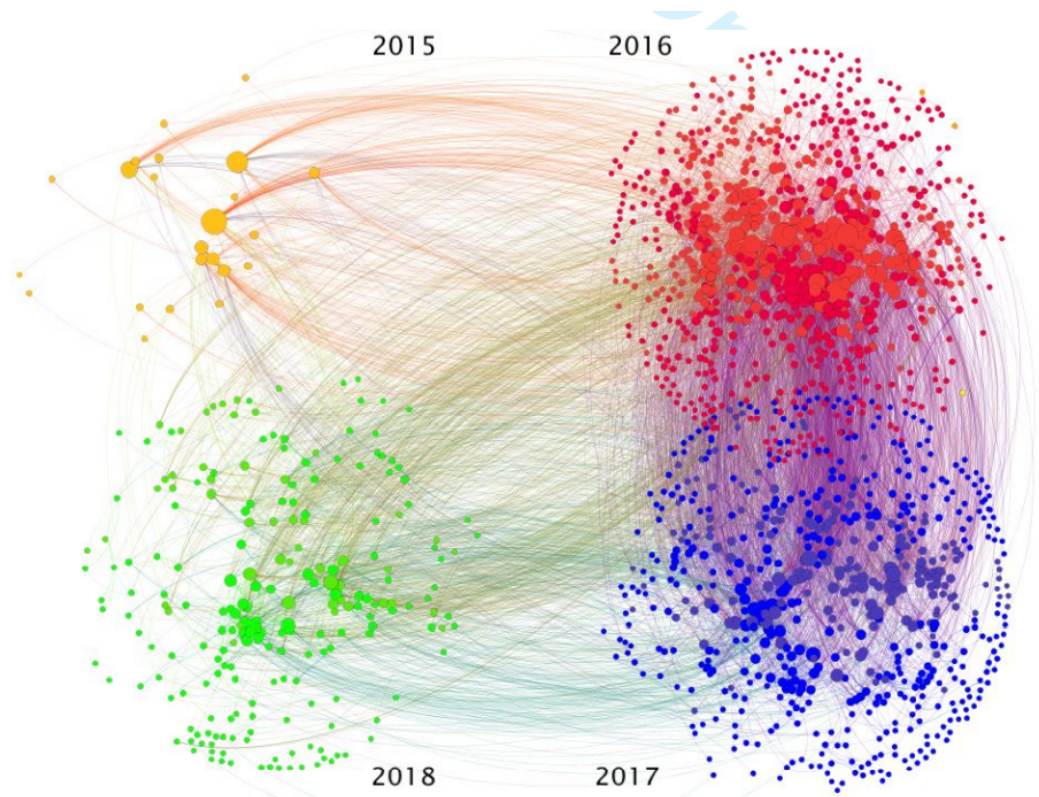
Para mapear a evolução da colaboração por coautoria, dividimos os pesquisadores de Zika em 4 grupos. Os grupos foram definidos de acordo com o ano de publicação de seu primeiro artigo sobre Zika: até 2015, 2016, 2017 e 2018. O número de pesquisadores para cada ano foi de 26, 748, 603 e 219, respectivamente. A Figura 6 mostra a rede de coautoria de cada autor com seus parceiros no mesmo ano de primeira produção, e com parceiros nos anos seguintes.

Analisando a Figura 6, podemos ver que os primeiros pesquisadores a produzir um artigo até 2015, continuaram a produzir e colaborar com grupos nos anos seguintes. O grupo de 2016, além de ser o maior de todos os quatro, teve uma taxa média de colaboração interna (entre pesquisadores do mesmo grupo) maior do que a dos demais, com exceção daqueles até 2015. Cada ponto corresponde a um pesquisador e cada linha uma coautoria entre eles. Quanto maior o ponto, maior o número de colabora-



ções. As cores e a posição descrevem o período quando cada pesquisador publicou seu primeiro artigo sobre Zika.

Figura 6: Colaboração científica entre pesquisadores.



Fonte: Elaboração dos autores.

### **Mapeamento dos projetos de pesquisa submetidos e contemplados na chamada MCTIC/FNDCT-CNPq/ MEC-CAPES/ MS-Decit / N° 14/2016 - Prevenção e Combate ao vírus Zika**

Foi efetuado o mapeamento de todos os projetos submetidos e aceitos na chamada Zika, bem como, todos os seus metadados como área temática, subáreas, regiões,





## PLATAFORMA E-LATTES

Um dos focos da Meta 2, foi o desenvolvimento e implementação da ferramenta e-Lattes na infraestrutura institucional e a geração de resultados de análise, com vistas a apoiar as metas 1 e 3 com informações sobre os pesquisadores brasileiros em Zika e a evolução do conhecimento na área, por meio da análise da produção científica. Na execução desta meta, foram identificadas as bases de dados para realização da prospecção da pesquisa, além de desenvolvimento de instrumentos e metodologia para realização da busca e disponibilização dos dados.

### *Bases de dados utilizadas na prospecção*

Plataforma Lattes, Diretório de Grupos de Pesquisa (DGP), Biblioteca Digital Brasileira de Teses e Dissertações (BDTD), projetos submetidos à chamada MCTIC/FNDCT-CNPq, bases científicas internacionais da *Web of Science* já tratadas (CWTS) e base de dados da altmetric.com.

### *Tecnologias e metodologias desenvolvidas*

Foram desenvolvidas tecnologias para captura das informações nas bases de dados referidas, denominada “Ferramenta e-Lattes”, para extração e análise dos currículos e aplicação da ferramenta de visualização VOSviewer do CWTS. Também foi desenvolvida metodologia de análise cientométrica para elaboração do mapa do conhecimento e análise da estrutura de poder na ciência.

## Metodologia/Estratégias Operacionais

Um primeiro levantamento foi realizado sobre quais pesquisadores possuíam como palavra-chave, ou produção científica, o nome “zika” em seus currículos. Um total de 2.653 pesquisadores doutores foram encontrados e os seus currículos analisados.

Esse procedimento pode ser replicado para o Ministério da Saúde e outras organizações acompanharem as demandas de pesquisa e o desenvolvimento das áreas de interesse fomentadas.

O sistema de busca avançada desenvolvido para a plataforma e-Lattes e alinhado com as demandas do próprio CNPq, permite encontrar perfis dos mais diversos com capacidade de análise pela própria ferramenta. As áreas de conhecimento são um componente importante, mas também são utilizadas análises textuais e de colaboração entre os pares, para que se possa chegar em um conjunto adequado de pesquisadores.

## **Produtos Tecnológicos**

Os produtos aqui descritos dizem respeito à última etapa do projeto, correspondendo às entregas dos pesquisadores vinculados ao projeto.

### *Desenvolvimento do e-Lattes*

Esse produto tem como objetivo principal elaborar um ambiente integrado de análise de dados com interface de recuperação e extração de informação sobre a ciência, aplicando modelos preditivos estatísticos que identifiquem inovações e concentração de especialidades da tríplice epidemia (Zika/dengue/chikungunya) para a integração de conhecimentos da coorte epidemiológica. Com o propósito de alcançar o objetivo principal, foram definidos os seguintes objetivos específicos:

1. Conceituar os artefatos tecnológicos necessários para a produção do ambiente integrado de análises de dados científicos;
2. Projetar interface de recuperação e extração de informação (*Information Extraction and Retrieval*), a fim de localizar temas, especialistas e áreas de conhecimento específicos dentro dos textos e dados das iniciativas de pesquisa da saúde;

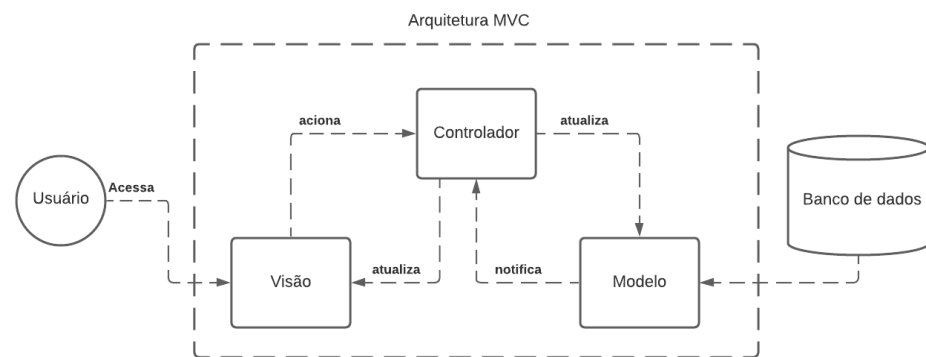
3. Desenvolver ambiente integrado de análise de dados sobre a ciência, aplicando modelos estatísticos que identifiquem inovações e concentração de especialidades da tríplice epidemia.

Para que isso seja possível, é necessário caracterizar e explicar os aspectos formativos e profissionais de especialistas relacionados com a tríplice epidemia em larga escala, por meio do tratamento, transformação, análise, visualização do currículo Lattes e de textos acadêmicos por eles produzidos.

O e-Lattes é um dispositivo tecnológico baseado em arquitetura *web*, isto é, para acessá-la é necessário um navegador, assim como possuir uma conexão à internet. O sistema pode ser acessado a partir do endereço [www.eLattes.com.br](http://www.eLattes.com.br), por meio de um computador ou aparelhos móveis.

De um ponto de vista técnico, o dispositivo foi desenvolvido utilizando as linguagens de programação: PHP 7.2, Javascript (CHAPMAN, 2018), Html5 (MCLAUGHLIN, 2011) e CSS3; baseando-se em uma metodologia padrão para desenvolver aplicações *web* com acesso a banco de dados, chamada de MVC (BUSCHMANN, 1996). A Figura 8 ilustra esse modelo padrão de codificação de sistemas.

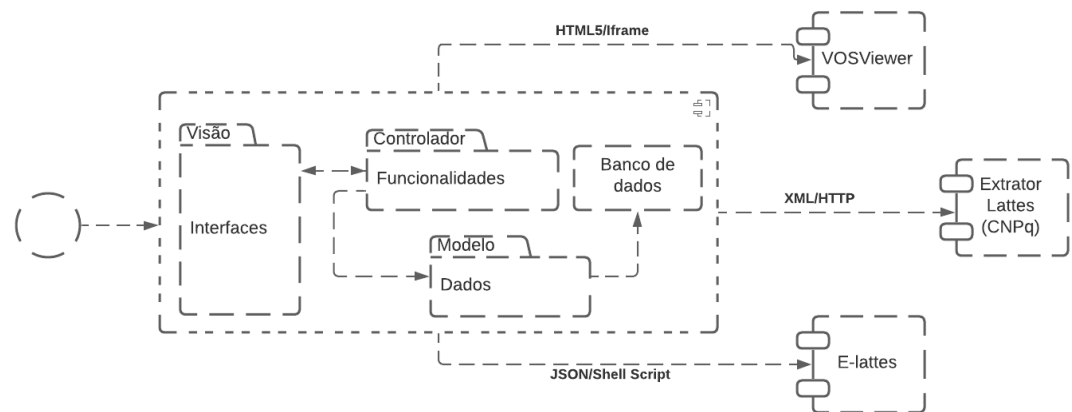
Figura 8: Estrutura padrão de codificação do e-Lattes *WEB*.



Fonte: Elaboração dos autores.

O e-Lattes foi desenvolvido para interoperar com outras soluções tecnológicas de manipulação de CVs Lattes: a coleta sistematizada dos CVs Lattes é feita por meio do serviço web Extrator Lattes (CNPq, 2021), e a mineração e classificação inicial dos dados curriculares é feita a partir da solução e-Lattes (SAMPAIO *et al.*, 2018). Ademais, o e-Lattes *WEB* integra-se com o *software* VOSviewer (VAN ECK; WALTMAN, 2018), a fim de criar visualizações e exploração por mapas baseados em dados de redes. A Figura 9 ilustra a interoperabilidade do e-Lattes *WEB* com essas diversas soluções citadas.

Figura 9: Arquitetura do *software* e-Lattes *WEB*.



Fonte: Elaboração dos autores.

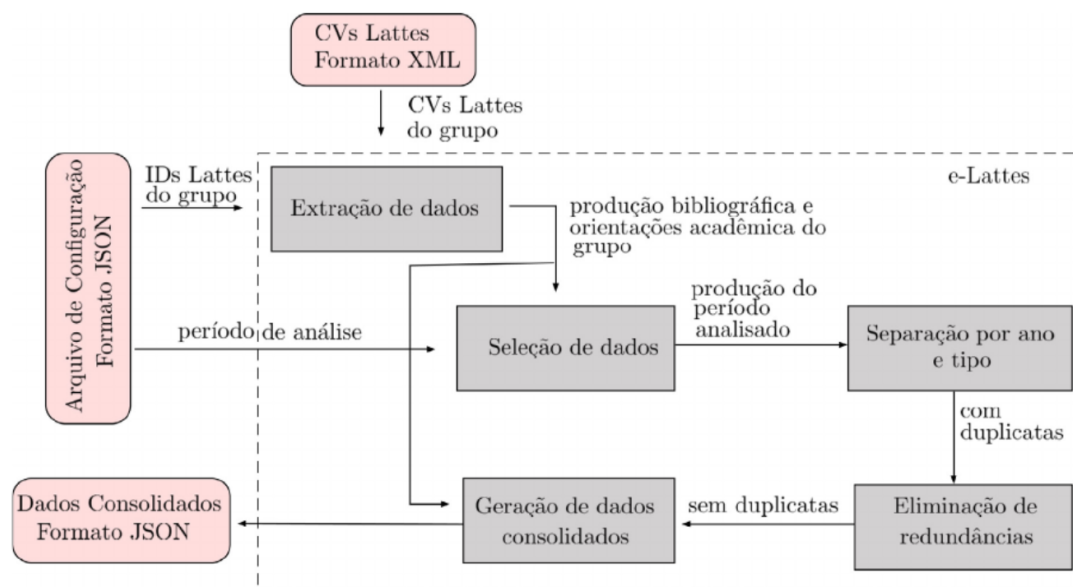
### *Coleta dos CVs Lattes: integração com o Lattes Extrator do CNPq*

O serviço Lattes Extrator (CNPQ, 2021) é uma solução de extração de CVs Lattes, voltada para a integração de sistemas, desenvolvida pelo próprio CNPq (ALVES *et al.*, 2012). Esta solução, conhecida pela expressão *Web Service* (AL-MASRI; MAHMOUD, 2007), permite que aplicações tecnológicas diferentes possam trocar informações de forma normalizada, por meio de uma linguagem universal, o XML. No caso do Lattes Extrator, este viabiliza que aplicações externas ao CNPq (ex. e-Lattes *WEB*) consuma as informações dos CVs Lattes diretamente do banco de dados da Plataforma Lattes.

## Mineração e processamento dos Currículos Lattes: integração com a solução de processamento de currículos em linguagem R

Após obter os CVs Lattes em formato XML, o e-Lattes aciona e monitora, via Shell Script, a solução desenvolvida por Sampaio *et al.* (2018), que por sua vez: 1) extrai os dados gerais dos pesquisadores e toda a sua produção bibliográfica e acadêmica; 2) filtra os dados com base em um período que pode ser personalizado; 3) elimina todas as redundâncias existentes; e, por fim 4) organiza as informações por tipo (publicações e orientações acadêmicas) e as estrutura no formato JSON (PEZOA *et al.*, 2016). A Figura 10 ilustra o processo executado pelo e-Lattes.

Figura 10: Processo de execução e-Lattes

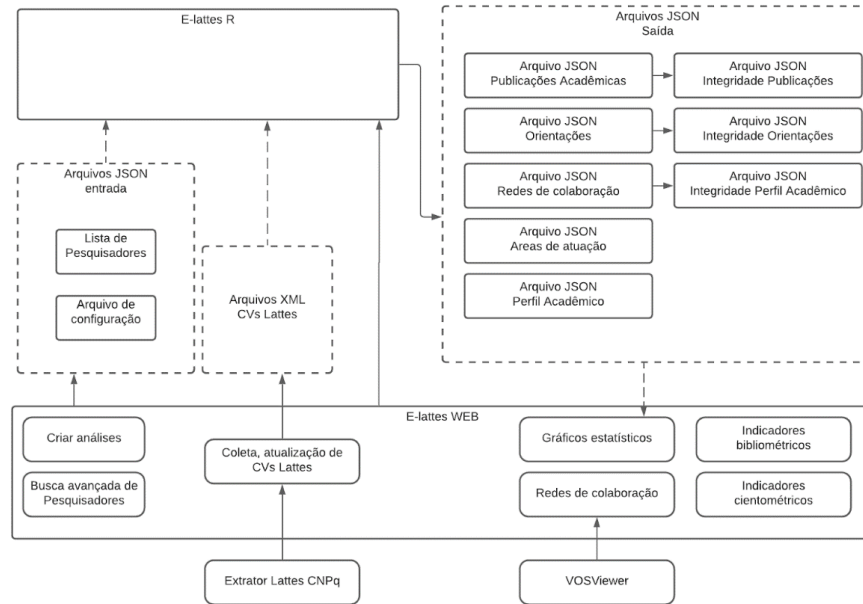


Fonte: Sampaio *et al.* (2018).

### Análises bibliométrica e cientométricas

A partir das informações estruturadas nos arquivos JSON, o e-Lattes realiza diversas análises bibliométricas e cientométricas e gera indicadores, totalizadores, gráficos estatísticos e visualizações de redes de colaboração, conforme ilustrado pela Figura 11.

Figura 11: Componentes do e-Lattes WEB



Fonte: Elaboração dos autores.

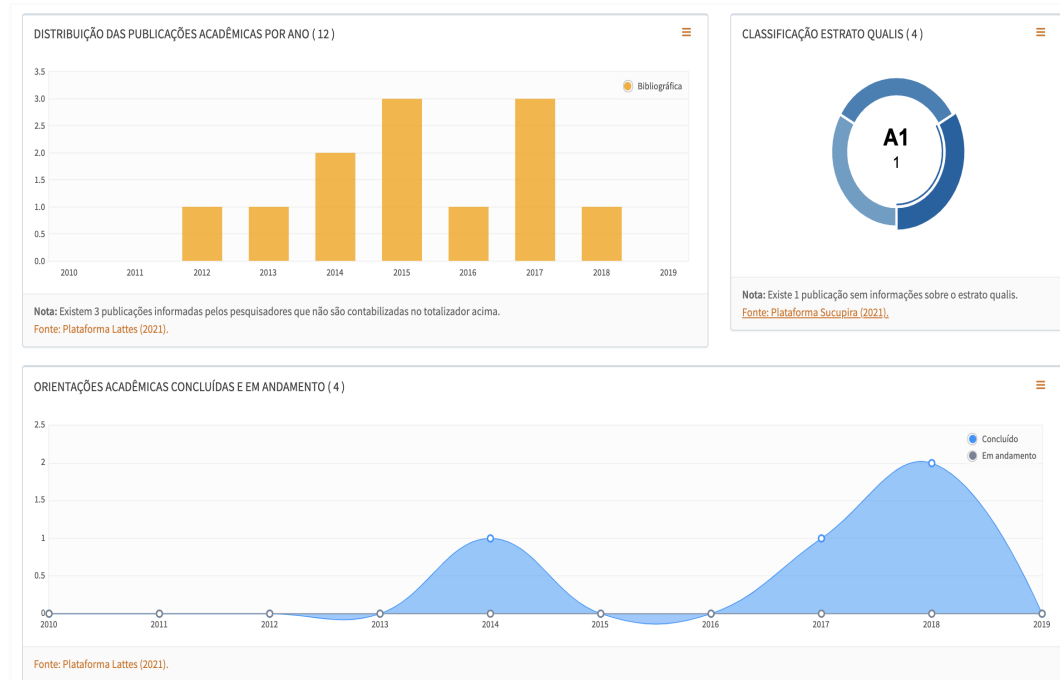
Por bibliometria, entendemos a utilização de métodos estatísticos e matemáticos específicos no estudo e avaliação de produções bibliográficas, isto é, livros, artigos ou comunicações científicas (MOREIRA *et al.*, 2020). Por cientometria, entendemos o estudo quantitativo da atividade científica, visto como atividade social, isto é, cujo objeto de análise consiste, não exclusivamente no produto científico (artigos e comunicações científicas), mas se estende ao produtor da ciência (o pesquisador), bem como, à circulação e ao consumo da produção científica (DIAS *et al.*, 2009).



## Resultados

As análises geradas pelo e-Lattes comportam métricas importantes definidas dentro do contexto da bibliometria e cientometria, obtendo-se os seguintes resultados, conforme as figuras<sup>3</sup>.

Figura 12: Indicadores de produção bibliográfica



Fonte: Plataforma Lattes, 2021.

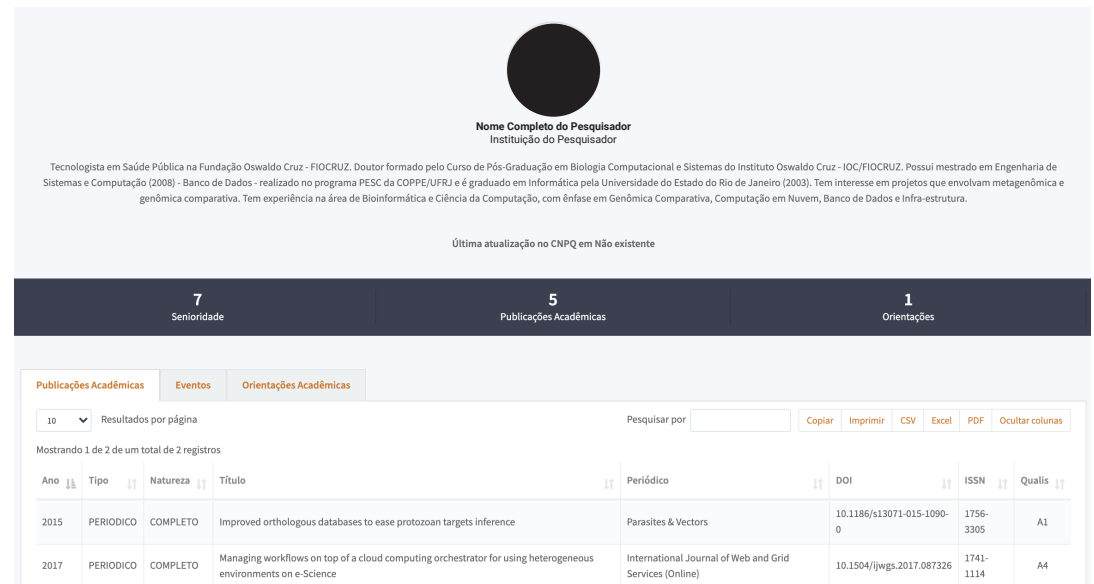
3. Análise dos currículos dos servidores da unidade de coordenação geral administração da Fiocruz. Os dados são coletados levando em consideração a data de entrada do servidor na instituição.

Figura 13: Indicadores de produção científica



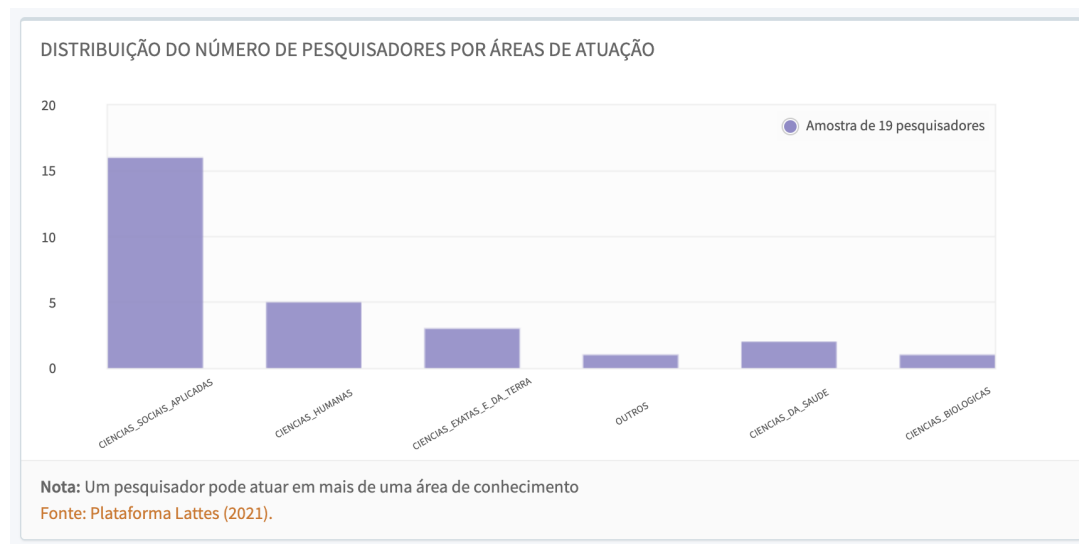
Fonte: Plataforma Lattes, 2021.

Figura 14: Exemplo de perfil acadêmico de pesquisador



Fonte: Plataforma Lattes, 2021.

Figura 15: Intercessão de áreas de atuação



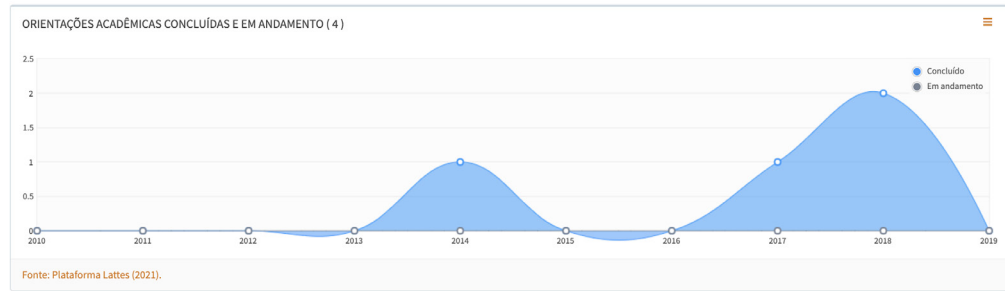
Fonte: Plataforma Lattes, 2021.

Figura 16: Produção bibliográfica estratificada em: livros, capítulos de livros, artigos, jornais/revistas e patentes



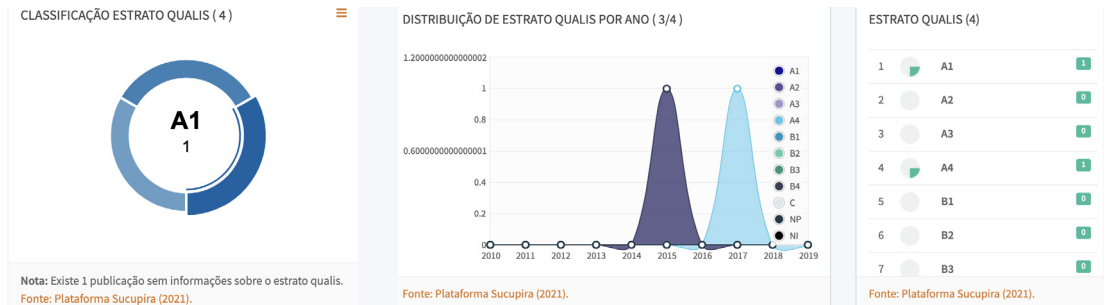
Fonte: Plataforma Lattes, 2021.

Figura 17: Orientações acadêmicas de um grupo



Fonte: Plataforma Lattes, 2021.

Figura 18: Classificação segundo o extrato Qualis Capes



Fonte: Plataforma Sucupira, 2021.

Figura 19: Índice de senioridade acadêmica calculado segundo Akazaki *et al.* (2016).

Pesquisador	Lattes id	Última atualização	Senioridade
<p><b>Paulo Roberto de Souza Lopes</b> Possui graduação em Economia pela Universidade Gama Filho (1990). Atualmente é assistente em c&amp;t da Fundação Oswaldo Cruz. Tem experiência na área de Economia, com ênfase em Contábil...</p> <p><b>Áreas de Atuação:</b> CIENCIAS_SOCIAIS_APLICADAS / Administração Pública</p> <p><b>Instituição:</b> Fundação Oswaldo Cruz / Presidência da Fiocruz</p>	3472311805512716	Não existente	7
<p><b>Nelson Peixoto Kotowski Filho</b> Tecnologista em Saúde Pública na Fundação Oswaldo Cruz - FIOCRUZ. Doutor formado pelo Curso de Pós-Graduação em Biologia Computacional e Sistemas do Instituto Oswaldo Cruz - IOC/FIOCRUZ. Possui mestrado em Engenharia de Sistemas e Computação (2008) - Banco de Dados - realizado no programa PESC da COPPE/UFRJ e é graduado em Informática pela Universidade do Estado do Rio...</p> <p><b>Áreas de Atuação:</b> CIENCIAS_BIOLÓGICAS / Genômica Comparativa</p> <p><b>Instituição:</b> Fundação Oswaldo Cruz / Diretoria de Administração</p>	8751616380292831	Não existente	7
<p><b>Saulo José da Silva</b> Possui graduação - BACHAREL EM ADMINISTRAÇÃO - FACULDADES INTEGRADAS SIMONSEN (2003). Especializado em Gestão Pública pela Universidade Cândido Mendes, e em Gestão de Organizações de Ciência e Tecnologia em Saúde pela Escola Nacional de Saúde Pública (ENSP/FIOCRUZ). Atualmente é Analista de Gestão em Saúde da Fundação Oswaldo Cruz. Experiência profissional na Gestão Pública, ...</p> <p><b>Áreas de Atuação:</b> CIENCIAS_SOCIAIS_APLICADAS</p> <p><b>Instituição:</b> Fundação Oswaldo Cruz / Diretoria de Administração</p>	0096264742018045	Não existente	0

**Nelson Peixoto Kotowski Filho**  
Fundação Oswaldo Cruz / Diretoria de Administração

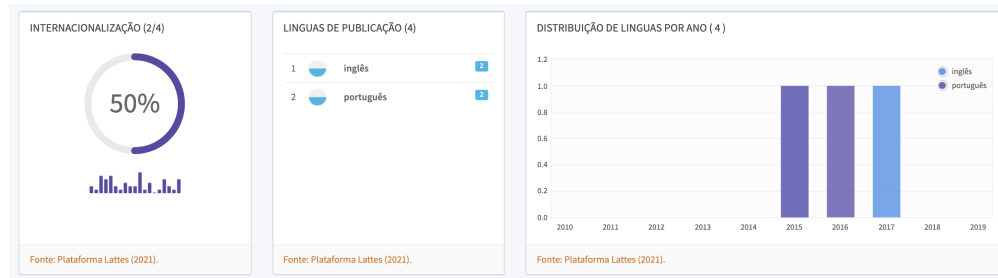
Tecnologista em Saúde Pública na Fundação Oswaldo Cruz - FIOCRUZ. Doutor formado pelo Curso de Pós-Graduação em Biologia Computacional e Sistemas do Instituto Oswaldo Cruz - IOC/FIOCRUZ. Possui mestrado em Engenharia de Sistemas e Computação (2008) - Banco de Dados - realizado no programa PESC da COPPE/UFRJ e é graduado em Informática pela Universidade do Estado do Rio de Janeiro (2003). Tem interesse em projetos que envolvam metagenômica e genômica comparativa. Tem experiência na área de Bioinformática e Ciência da Computação, com ênfase em Genômica Comparativa, Computação em Nuvem, Banco de Dados e Infra-estrutura.

Última atualização no CNPQ em Não existente

<b>7</b> Senioridade	<b>5</b> Publicações Acadêmicas	<b>1</b> Orientações
-------------------------	------------------------------------	-------------------------

Fonte: Plataforma Lattes, 2021.

Figura 20: Indicadores de internacionalização



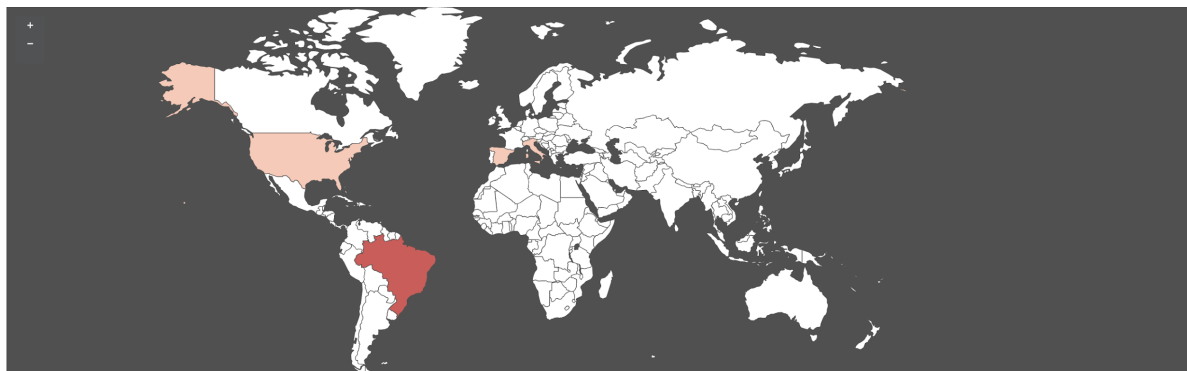
Fonte: Plataforma Lattes, 2021.

Figura 21: Distribuições das publicações acadêmicas por ano



Fonte: Plataforma Lattes, 2021.

Figura 22: Distribuição geográfica de publicações e participações em eventos, por país e por ano



Fonte: Plataforma Lattes, 2021.

Por intermédio destas informações e indicadores apresentados, é possível medir a produção científica e a sua importância, seja referente a um autor (indivíduo, departamento, grupo de investigação, instituição, ou país), através da sua produtividade (número total de artigos), seja referente a um artigo, através da sua relevância, influência em relação a outros autores investigadores (número total de citações), ou ainda, referente ao fator de impacto de jornal/revista/evento acadêmico onde foram publicados os artigos.

Dentro do escopo dos estudos bibliométricos e cientométricos, temos os de cálculo de indicadores e a análise de redes de colaboração científica, também conhecida como a exploração de redes sociais dentro do âmbito científico, ou redes de co-autoria (DIAS *et al.*, 2009). Para apresentar visualizações de redes de colaboração entre os pesquisadores, o e-Lattes integra-se com o *software* VOSviewer (VAN ECK; WALTMAN, 2018) um *software* de criação, visualização e exploração de dados baseado em redes (MOREIRA *et al.*, 2020). Segundo seus criadores (VAN ECK & WALTMAN, 2018), o VOSviewer pode ser utilizado em qualquer conjunto de dados de redes e adota o método conhecido como VOS (Visualization of Similarities). Todos os dados gerados

pelo e-Lattes estão disponíveis para visualização e download em diversos formatos: JSON, CSV, XLS, ZIP, PDF etc.

### **Integração do e-Lattes**

Este produto teve como objetivo principal, elaborar análises detalhadas visando a compreensão de como se agrupam os especialistas relacionados com a tríplice epidemia (Zika/dengue/chikungunya), porém pode ser aplicado para qualquer outra área científica. Para que isso seja possível, é necessário caracterizar e explicar os aspectos formativos e profissionais de especialistas relacionados com a tríplice epidemia, em larga escala e por meio do tratamento, transformação, análise, visualização do currículo Lattes e de textos acadêmicos por eles produzidos.

Com o propósito de alcançar o objetivo principal, foram definidos os seguintes objetivos específicos:

1. Modelo de Dados disponível via protocolo REST;
2. Desenvolvimento do pré-processamento de dados de forma automatizada em ambiente Docker;
3. Integração da base de dados desenvolvida em conformidade com o padrão BR-CRIS (IBICT).

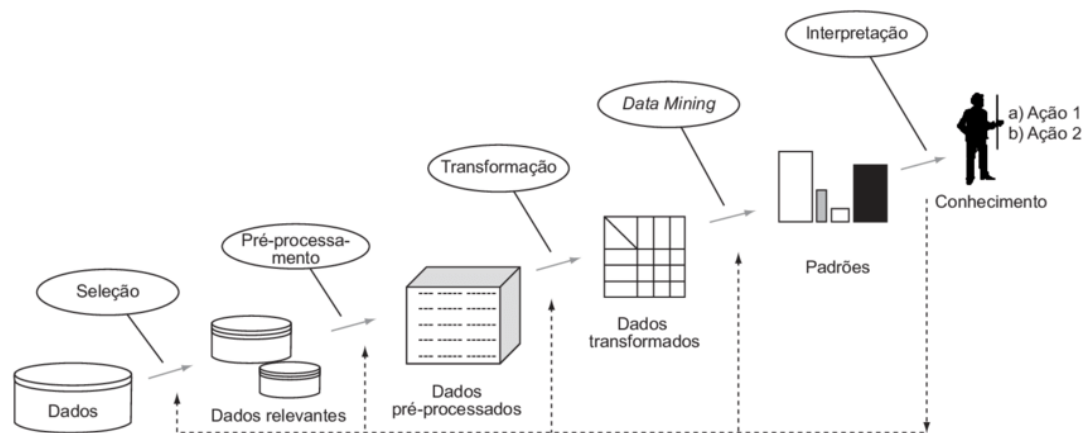


## Metodologia

Para a extração de conhecimentos em textos diversos, é necessário a utilização de metodologias de mineração de textos (*Text mining*), que de acordo com Aranha e Passos (2006), pode ser definida como uma metodologia para extração de informações a partir de dados textuais. Sua implementação tem como base o uso de técnicas de processamento de linguagem natural, recuperação de informação e aprendizado de máquina para tratar os dados, a fim de filtrar apenas informações que possam ser úteis para o usuário final. Assim sendo, esse processo de mineração pode ser resumido em quatro etapas fundamentais:

1. Coleta de documentos: esta etapa tem como objetivo a coleta de documentos que irão compor o banco de textos a ser analisado;
2. Pré-processamento: nesta etapa os documentos coletados são estruturados de forma padronizada para que os algoritmos que serão utilizados posteriormente, sejam capazes de fazer a manipulação de todos os documentos. Também são definidos os termos e caracteres especiais que não possuem significado relevante (preposições, artigos, pontuação, cabeçalhos, etc.), que serão removidos do conjunto de textos;
3. Extração de conhecimento: esta etapa consiste na extração do conhecimento por meio da aplicação de algoritmos de extração automática. Os termos são agrupados conforme sua similaridade e interação entre si no conjunto de textos;
4. Avaliação e interpretação dos resultados: por fim, os resultados obtidos são avaliados a fim de determinar se os algoritmos revelaram informações relevantes sobre o conjunto de textos.

Figura 23: Etapas do processo de extração de conhecimento de texto



Fonte: Fayyad *et al.*, 1996

A mineração de texto ou *text mining* é uma área multidisciplinar que envolve estatística, mineração de dados e aprendizado de máquina, e está voltada para a análise de dados não estruturados textuais, como por exemplo, postagens em redes sociais, *reviews* de produtos, etc. Trata-se de uma área que vem se tornando cada vez mais importante, pois pode, por exemplo, auxiliar empresas em tomadas de decisões rápidas e pontuais, baseando-se na avaliação de um produto por seus usuários, ou na identificação de algum acontecimento através de publicações nas redes sociais, entre outras aplicações (PRACIANO *et al.*, 2018). No entanto, tratar tamanha quantidade de dados não é uma tarefa trivial. Nesse sentido, faz-se necessário utilizar outras técnicas relacionadas ao aprendizado de máquina (*Machine learning*), para que possam ser criados modelos preditivos para um estudo mais robusto do conjunto de documentos analisados.

## *Machine Learning*

Com o alto volume de informações geradas no dia a dia, torna-se cada vez mais necessária a utilização de algoritmos que sejam capazes de identificar padrões, pois diversas empresas e órgãos do governo não são capazes de analisar todas as informações existentes, ou entregar de maneira célere (NASRABADI, 2007). Portanto, para realizar esse tipo de atividade, é necessário entender o problema existente e ter um volume de dados razoável para realizar o treinamento do modelo. Com a técnica de Aprendizado de Máquina (AM) é possível automatizar a construção de sistemas inteligentes que podem ser ajustados de acordo com a necessidade de cada tarefa (BONNACCORSO, 2017). Em outras palavras, AM é um conjunto de regras que possibilitam a uma máquina tomar decisões baseadas em experiências anteriores, diferentemente de um *software*, o qual é previamente definido para tratar cada tipo de regra. Além disso, modelos de AM podem ser aperfeiçoados e apresentar melhor desempenho quando expostos a novos dados (ZURADA, 1995).

O conceito de AM pode ser sintetizado como a capacidade de um programa de computador aprender com a experiência (E) relacionada a alguma classe de tarefas (T), baseada em uma medida de desempenho (P). Dessa forma, o desempenho em tarefas (T), quando medido por (P), melhora com a experiência em (E) (MITTCHEL, 2017). Essas técnicas têm sido amplamente utilizadas para a resolução de diversos problemas. Atualmente, o assunto está em alta e existem diversas oportunidades para colocar em prática a matemática desenvolvida para a criação desses algoritmos. Gigantes da indústria têm investido severamente para o desenvolvimento de novas tecnologias, como os veículos autônomos, robôs advogados, veículos não tripulados e na identificação de doenças.

Existem distintos paradigmas utilizados para o ensinamento de máquinas. Entretanto, este trabalho focou apenas nos algoritmos de AM utilizados no desenvolvimento do *framework*, ou seja, não foram abordados temas como aprendizado estatístico, ou redes bayesianas.

1. Aprendizagem supervisionada. O aprendizado supervisionado é uma das formas de ensinar uma tarefa para a máquina, onde existe uma entrada e uma saída desejada que já foram anteriormente rotuladas. Esse processo pode ser feito de forma manual, ou por meio de dicionários, quando a informação de entrada é um texto. Com os dados detalhados, a máquina é capaz de classificar novas entradas a partir de experiências antigas (MITCHELL, 1999). No aprendizado supervisionado, todos os dados de entrada do modelo devem ser rotulados. No exemplo, são fornecidos três dados ao modelo, sendo dois rotulados como “cão” e um rotulado como “gato”. Tendo por base as características, ou *features* de cada dado de entrada, o modelo passa por uma fase de treinamento:

a. O k-NN ou *K-Nearest Neighbors* é um classificador onde o aprendizado é baseado “no quão similar” é um dado (um vetor) do outro. O treinamento é formado por vetores de  $n$  dimensões.

b. O SVM ou *Support Vector Machine* é um classificador onde o hiperplano é “cortado”, a fim de separar em duas classes o conjunto de dados.

c. O Naive Bayes utiliza a teoria de Bayes de probabilidade, e é chamada de Naive (ingênua) por partir do pressuposto de que cada *feature* é independente uma da outra.

2. Aprendizagem não supervisionada. Nesse tipo de aprendizado, não existe a necessidade de um vetor contendo os rótulos de cada medida, ou seja, não é necessário informar à máquina o tipo de cada dado. Essa não inserção de rótulos faz com que o algoritmo aprenda de acordo com as características dos dados de entrada. Esse tipo de paradigma não tem certos benefícios que existem no aprendizado supervisionado. Para o seu funcionamento, o modelo propõe hipóteses a partir do que foi inserido na entrada (BONNACORSO, 2017). Esse tipo de algoritmo é muito interessante em bases de dados que ainda não foram exploradas, quando se necessita visualizar os dados agrupados. O conceito de *clusters* (agrupamentos) é utilizado nesse tipo de aprendizado, que é o método quando as variáveis que possuem ten-

dências, ou padrões semelhantes, agrupam-se para formar grupos que possuam as mesmas características. Alguns dos algoritmos mais populares que resolvem este tipo de problema são:

- a. O algoritmo *k-Means*, usado geralmente em problemas de *hard clustering*, ou seja, o dado observado só pode pertencer a um *cluster*.
- b. O algoritmo *Fuzzy C Means*, usado geralmente em problemas de *soft clustering*, ou seja, o dado observado pode pertencer a um ou mais clusters.
- c. O *Hierarchical Clustering* agrupa objetos semelhantes em grupos chamados *clusters*. Ao final, tem-se um conjunto de *clusters*, em que cada um é distinto do outro e os objetos em cada *cluster* são amplamente semelhantes entre si.

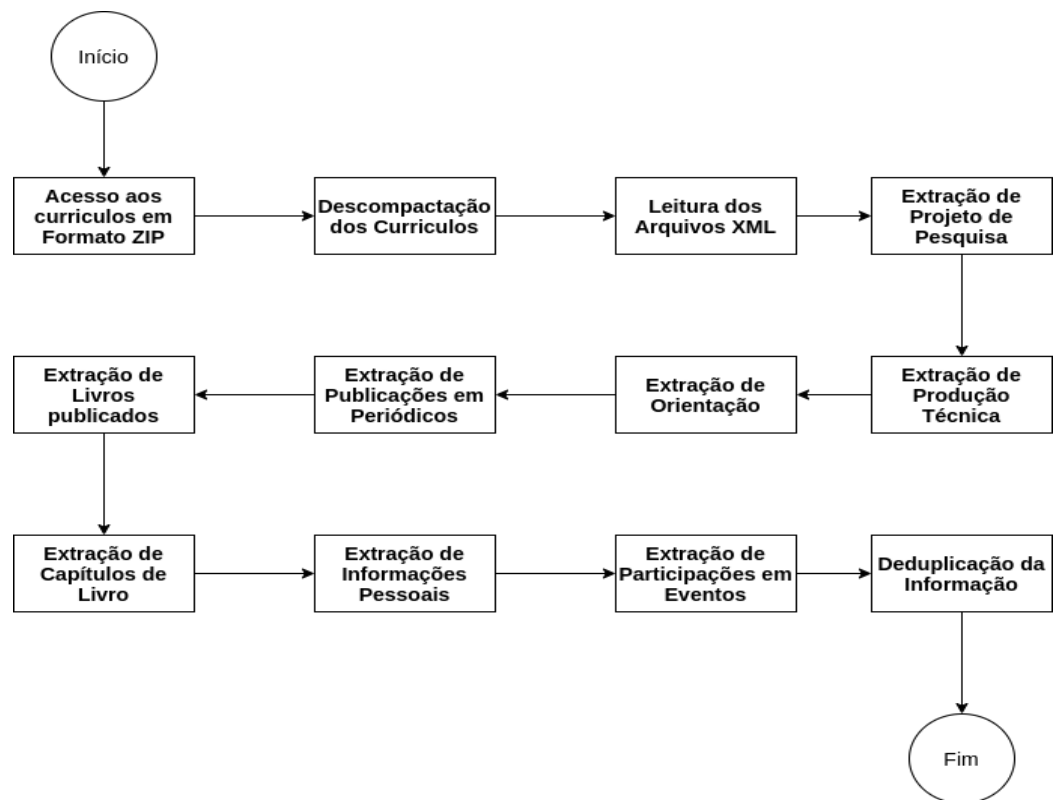
3. Aprendizagem por reforço. O último paradigma abordado nesse trabalho, conhecido como aprendizado por reforço, baseia-se na forma com a qual o modelo reage de acordo com os dados passados a ele pelo ambiente, ou seja, o modelo aprende continuamente. Esse aprendizado é realizado através das ações que são executadas. Caso seja a decisão correta, o modelo ganha uma recompensa, caso contrário pode ser imposta algum tipo de penalidade (ZURADA, 1999). Nesse paradigma de aprendizado, é necessário ajustar os parâmetros de forma contínua com o intuito de maximizar ou minimizar um determinado índice, e sempre há um “treinador” nessa fase, verificando se os estímulos proporcionados estão tendo a saída correta. Caso esses dados não estejam balanceados, eles podem alterar a performance do algoritmo (MAZUROWSKI, 2008).

Dentre os três paradigmas apresentados, os tipos de aprendizagens 1 e 2, são geralmente mais utilizados em contextos de mineração de texto, sendo cada um aplicado em seu escopo de estudo respectivo.

## Resultados

Visto que o objetivo geral é analisar detalhadamente os especialistas relacionados com a tríplice epidemia, iniciamos pela aplicação da metodologia *Text mining*. Primeiramente obtivemos o desenvolvimento de um novo interpretador para leitura dos arquivos de origem da Plataforma Lattes (Figura 24).

Figura 24: Diagrama de blocos do novo interpretador de currículos Lattes



Fonte: Elaborado pelos autores.

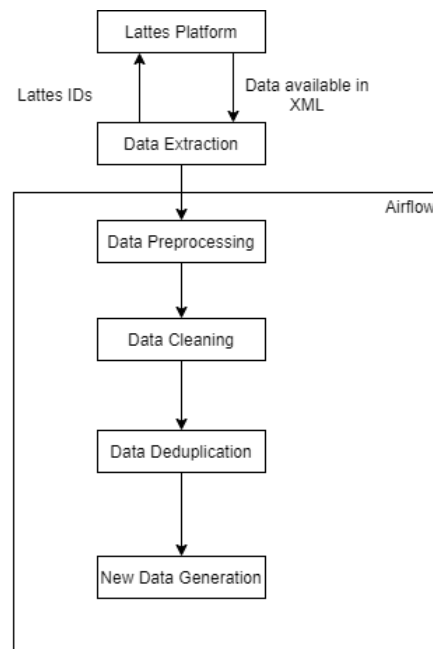


A Figura 26 mostra em detalhes as etapas do *framework* proposto. Cada caixa disponível é um *script* Python, que auxilia no uso dos dados durante todo o processo, e a caixa principal usa o Airflow como orquestrador de código.

Este pacote é responsável por analisar os dados coletados pela Plataforma Lattes. Ele pode extrair os detalhes através da *Extensible Markup Language* (XML), e isso é possível porque cada pesquisa possui sua chave de identificação. Portanto, o *framework* proposto pode extrair os detalhes dos pesquisadores, organizá-los e colocá-los em uma pasta local.

A estrutura proposta usa *Airflow*, um gerenciador de *pipeline* de código aberto. Com esta biblioteca, é possível orquestrar quando cada *script* deve ser executado. Ele contém *scripts* responsáveis pela limpeza de dados, de duplicação de dados e geração de novos dados (transformação de dados).

Figura 26: Esquema de execução do *framework* proposto



Fonte: Elaborado pelos autores.



O processamento dos dados é a etapa responsável por descompactar os dados coletados pela Plataforma Lattes e disponibilizá-los para as próximas etapas do *pipeline* de ingestão de dados. Ele também verifica a chave de identificação de cada currículo e a quantidade de dados disponíveis naquele contêiner local.

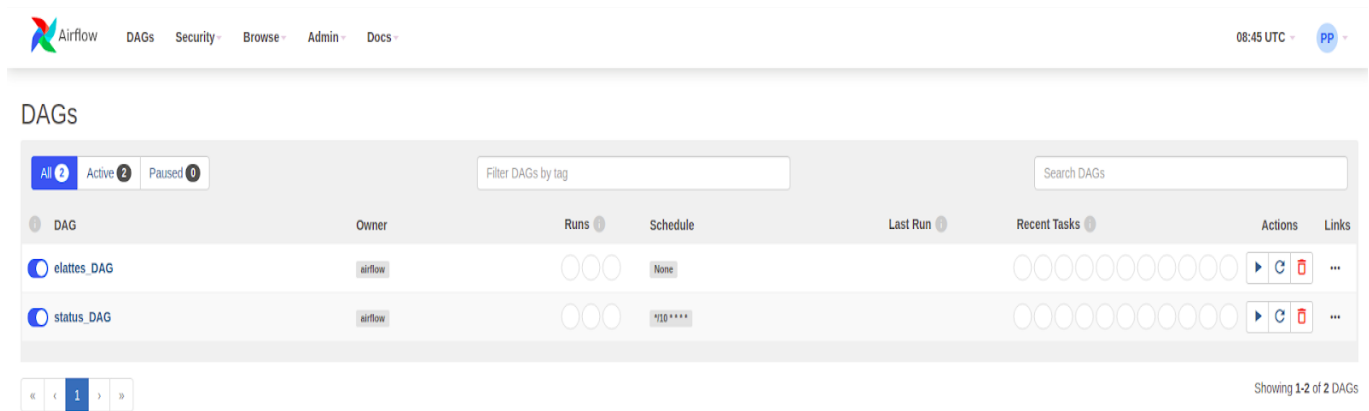
Devido à grande quantidade de dados disponíveis em português, o processo de limpeza de dados é necessário para remover alguns caracteres incomuns e normalizar os dados para as próximas etapas. Nesse processo, as palavras irrelevantes e a pontuação também são removidas. Outra ação importante é disponibilizar o texto em UTF-8, assim que os dados estiverem disponíveis em um banco de dados para consultas futuras.

Como a colaboração acadêmica é uma realidade nas universidades e os pesquisadores brasileiros devem ter um perfil na Plataforma Lattes, as informações sobre periódicos acadêmicos, conferências e projetos podem aparecer duas vezes na etapa de extração de dados. Para superar esse problema, a deduplicação de dados é importante, mas é necessário rastrear os participantes e criar um grafo de colaboração entre eles.

Para tanto, a biblioteca deduplicação oferece suporte à deduplicação de texto utilizando lógica *fuzzy* em seu núcleo. A decisão de remover os dados extras é feita com base na ordem dos autores deste artigo. Com o apoio de outra Biblioteca Python, chamada Pandas, os dados duplicados são removidos e a ordem dos autores de cada artigo é mantida.

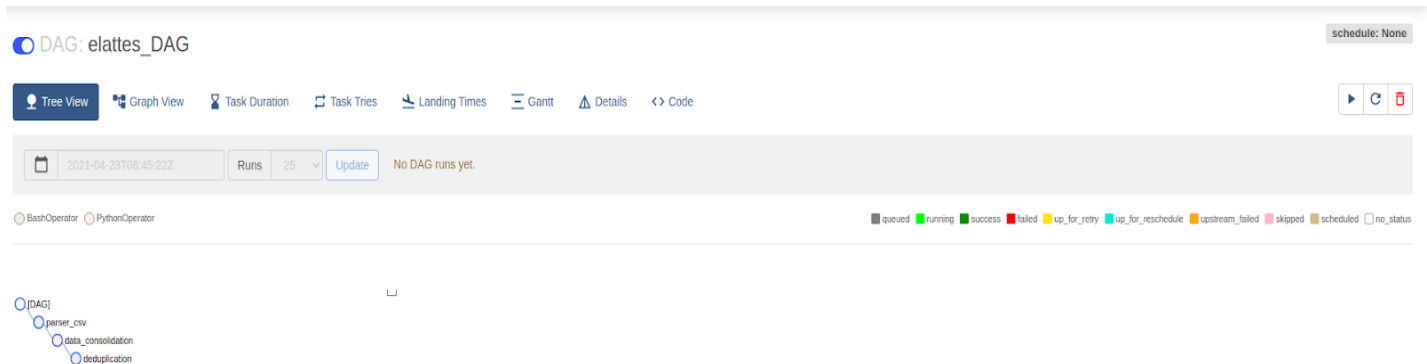
Na última etapa, os dados estão disponíveis para uso futuro. A saída principal é em valores separados por vírgula (CSV) e consiste em seis arquivos que contêm informações que estavam disponíveis apenas em XML. Outra opção é persistir os dados em um banco de dados, e isso é o fim do processo.

Os artefatos gerados pelo *Airflow* podem ser visualizados nas Figuras 27 e 28. Na Figura 27 é apresentada toda a parte de informações referentes às execuções, onde também é possível realizar seu monitoramento.

Figura 27: *Dashboard Airflow*Fonte: *Software Airflow*

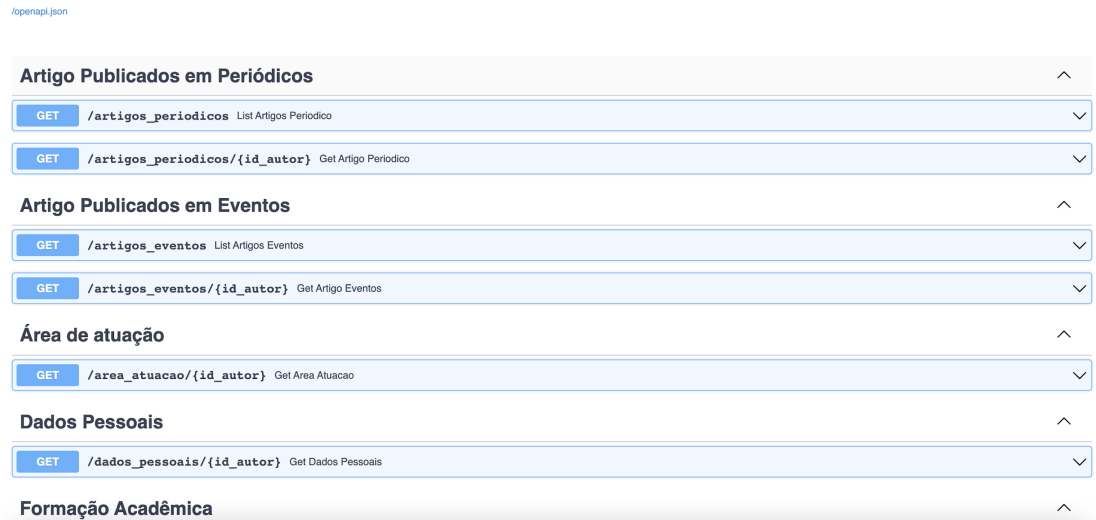
A Figura 28 representa o estado de execução de cada estágio e, com isso, é possibilitado ao operador identificar se houve algum problema na etapa de Extração, Transformação e Carga (ETL).

Figura 28: Etapas do processo de ETL da plataforma e-Lattes

Fonte: *Software Airflow*

Em relação ao desenvolvimento do serviço *web* através do protocolo REST, este se deu por intermédio do micro *framework* FastAPI, onde é possível incluir diversas lógicas de negócio de maneira escalável e com uma interface gráfica chamada Swagger (Figura 29), onde se pode visualizar todos os endereços para realizar futuras integrações.

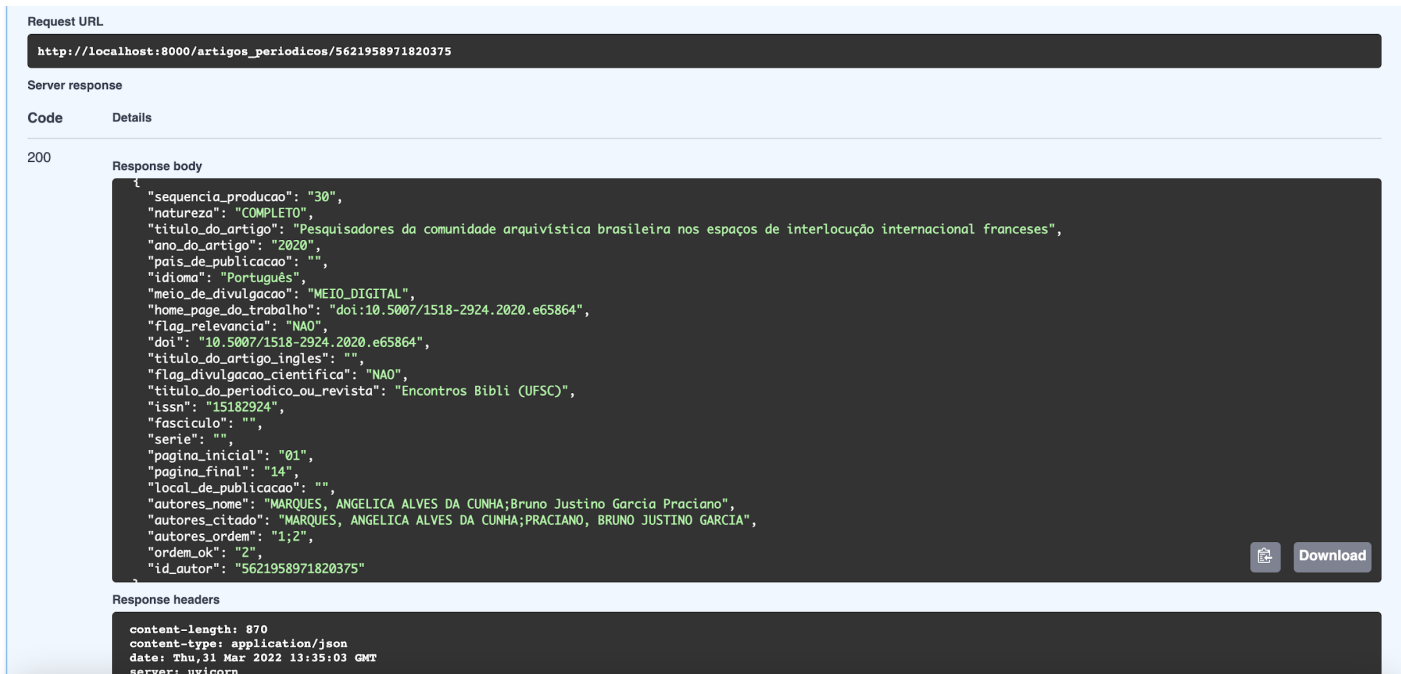
Figura 29: Tela principal da interface do serviço *web*



Fonte: Interface Swagger - FastAPI

Na Figura 30 é disponibilizado o formato do JSON referente à produção bibliográfica baseada no currículo Lattes. Esse formato foi definido em algumas reuniões com o grupo de pesquisa do BrCris.

Figura 30: JSON de artigos publicados em periódicos



```
Request URL
http://localhost:8000/artigos_periodicos/5621958971820375

Server response
Code 200
Details

Response body
{
  "sequencia_producao": "30",
  "natureza": "COMPLETO",
  "titulo_do_artigo": "Pesquisadores da comunidade arquivística brasileira nos espaços de interlocução internacional franceses",
  "ano_do_artigo": "2020",
  "pais_de_publicacao": "",
  "idioma": "Português",
  "meio_de_divulgacao": "MEIO_DIGITAL",
  "home_page_do_trabalho": "doi:10.5007/1518-2924.2020.e65864",
  "flag_relevancia": "NAO",
  "doi": "10.5007/1518-2924.2020.e65864",
  "titulo_do_artigo_ingles": "",
  "flag_divulgacao_cientifica": "NAO",
  "titulo_do_periodico_ou_revista": "Encontros Bibli (UFSC)",
  "issn": "15182924",
  "fasciculo": "",
  "serie": "",
  "pagina_inicial": "01",
  "pagina_final": "14",
  "local_de_publicacao": "",
  "autores_nome": "MARQUES, ANGELICA ALVES DA CUNHA;Bruno Justino Garcia Praciano",
  "autores_citado": "MARQUES, ANGELICA ALVES DA CUNHA;PRACIANO, BRUNO JUSTINO GARCIA",
  "autores_ordem": "1;2",
  "ordem_ok": "2",
  "id_autor": "5621958971820375"
}

Response headers
content-length: 870
content-type: application/json
date: Thu, 31 Mar 2022 13:35:03 GMT
server: uvicorn
```

Elaborado pelos autores

Outro resultado desse produto é a base de dados em PostgreSQL contendo mais de milhões de currículos Lattes já processados e disponibilizados em formato de base de dados relacional. Na Figura 31, é possível visualizar os dados das tabelas que contêm as informações referentes a publicações em periódicos.

Figura 31: Tabela contendo as informações referentes a publicações em periódicos

Grande	TITULO_DO_TEXTO	ANO_DO_TEXTO	PAIS_DE_PUBLICACAO	IDIOMA	MEIO_DE_DIVULGACAO	HOME_PAGE_DO_TRABALHO
1	RQUIVIVOS BRASILEIROS DE CARDIOLOGIA	2009	Brasil	Português	VARIOS	
2	ireito à leitura: desafios e contradições	2011	Brasil	Português	VARIOS	
3	iblioteconomia como opção profissional	2016	Brasil	Português	VARIOS	
4	ual melhor capim?	2006	Brasil	Português	IMPRESSO	revistaruraldobrasil.com.br
5	tervalo entre partos: Importante na produção	2006	Brasil	Português	IMPRESSO	revistaruraldobrasil.com.br
6	uidado na alimentação de animais de elite	2008	Brasil	Português	IMPRESSO	
7	imentação de vacas no período de déficit ei	2006	Brasil	Português	VARIOS	www.revistaruraldobrasil.com.br
8	ma "homepage" voltada para a Medicina. H	1996	Brasil	Português	IMPRESSO	
9	so de máscaras é obrigatório a partir de hoj	2020	Brasil	Português	MEIO_DIGITAL	https://www.em.com.br/app/noticia/gerais/
10	oronavírus: os detalhes das sete vacinas em	2020	Brasil	Português	MEIO_DIGITAL	https://www.em.com.br/app/noticia/gerais/
11	andemia de COVID-19 põe em risco o comb.	2020	Brasil	Português	VARIOS	https://www.em.com.br/app/noticia/gerais/
12	pecialistas divergem sobre uso de máscara	2020	Brasil	Português	MEIO_DIGITAL	https://www.em.com.br/app/noticia/gerais/
13	NVESTIGANDO CIÊNCIAS NO MUNICÍPIO D	2004	Brasil	Português	IMPRESSO	http://www.pucminas.br/graduacao/cursos
14	engenharia genética e a biotecnologia na pi	2007	Brasil	Português	IMPRESSO	
15	terferência da dieta de alto grão sobre as ci	2013	Brasil	Português	MEIO_DIGITAL	
16	inética de fermentação ruminal in vitro, com	2013	Brasil	Português	MEIO_DIGITAL	
17	gronegócio : Ovinocultura é a aposta da pec	2007	Brasil	Português	IMPRESSO	http://folhadaestancia.com.br/modules/nev
18	elhoramento florestal: seleção	2005	Brasil	Português	IMPRESSO	www.sif.org.br
19	esprestígio dos professores: política educac	2019	Brasil	Português	VARIOS	http://revistacipa.com.br/a-revista/
20	iferenças de Genero, Sexo e Suceptibilidade	2018	Moçambique	Português	IMPRESSO	
21	eminário de Biodiesel - Combustível para a	2010	Brasil	Português	IMPRESSO	http://www.gazetanortemineira.com.br/nov
22	Círculo de Meio Ambiente	2010	Brasil	Português	IMPRESSO	http://www.gazetanortemineira.com.br/nov
23	pecialidades Ampliam Leque de Atuação Pr	2017	Brasil	Português	IMPRESSO	
24	CIDADE DE LONDRES NO SÉCULO XIX: UM	2008	Brasil	Português	MEIO_DIGITAL	http://eduep.uepb.edu.br/alpharrabios/v2-r
25	CINE SÃO JOSÉ COMO ESPAÇO DE LAZER	2011	Brasil	Português	VARIOS	http://mhn.uepb.edu.br/revista_tarairiu/h2/
26	RISPR: Editando o DNA como se edita um te	2018	Brasil	Português	VARIOS	https://biomedinforma.files.wordpress.com
27	estes Rápidos: Como funcionam?	2019	Brasil	Português	NAO_INFORMADO	https://biomedinforma.files.wordpress.com
28	esign Magazine	2016	Brasil	Português	VARIOS	designmagazine.com.br
29	efinição do Padrão de Qualidade dos Gên	1995	Brasil	Português	IMPRESSO	

Fonte: DBeaver 22.0.0

## Modelagem de Tópicos para detecção de especialistas

Esse produto tem como objetivo principal: a) o desenvolvimento de técnica para modelagem de tópicos para detecção de temas específicos de pesquisa nas áreas de saúde incluindo a Tríplice Epidemia e b) utilização de técnicas de mineração de texto com aprendizado de máquina para busca de especialistas. Com o propósito de alcançar o objetivo principal, foram definidos os seguintes objetivos específicos:

1. Analisar a qualidade dos agrupamentos das informações conforme sua similaridade e interação no conjunto de textos;
2. Inferir se os algoritmos de extração automática são capazes de revelar informações relevantes.

## Metodologia

### *Mineração de Textos*

Foram seguidas as etapas necessárias para mineração de texto abordadas a seguir e conforme a Figura 32.

Figura 32: Etapas do *text mining*



Fonte: Adaptado de Aranha e Passos, 2006

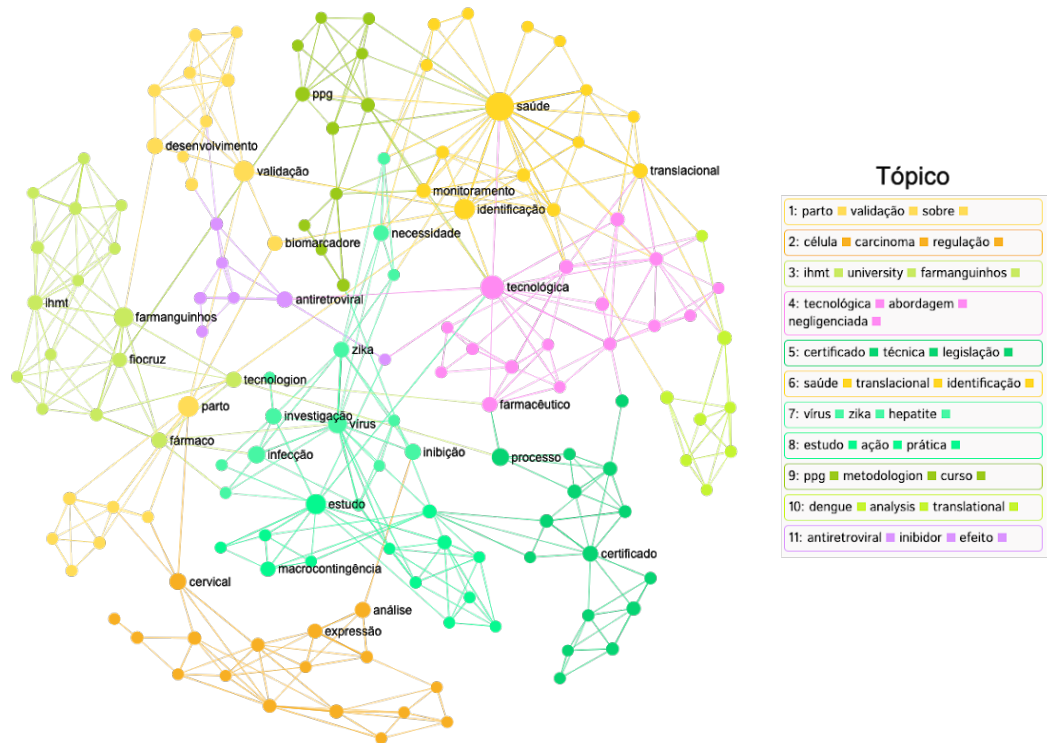
## Resultados

Foi proposta uma abordagem baseada na modelagem de tópicos, utilizando dados relativos aos periódicos e as teses e dissertação. Em razão da quantidade de dados a serem analisados, a modelagem de tópicos se mostra uma estratégia eficiente no que diz respeito ao descobrimento de novas informações.

Utilizamos o modelo LDA (*Latent Dirichlet Allocation*) para verificar se esse tipo de estratégia é adequado para o que se propõe. O LDA é um modelo bayesiano completo e se baseia na geração dos tópicos como distribuições de Dirichlet (BLEI; NG; JORDAN, 2003). Esse modelo é capaz de classificar documentos não conhecidos (documentos que não foram utilizados no treinamento) e utilizar informações *a priori*. Por essas características, o LDA tem influenciado uma grande quantidade de trabalhos e se tornou a base dos modernos modelos estatísticos de aprendizado de máquina, resultando em uma nova classe de modelos estatísticos chamados Modelos Probabilísticos de Tópicos.

A partir de um modelo LDA foi possível gerar visualizações como a da Figura 33.

Figura 33: *Cluster* de concentração de termos em função dos tópicos encontrados



Fonte: Elaborado pelos autores

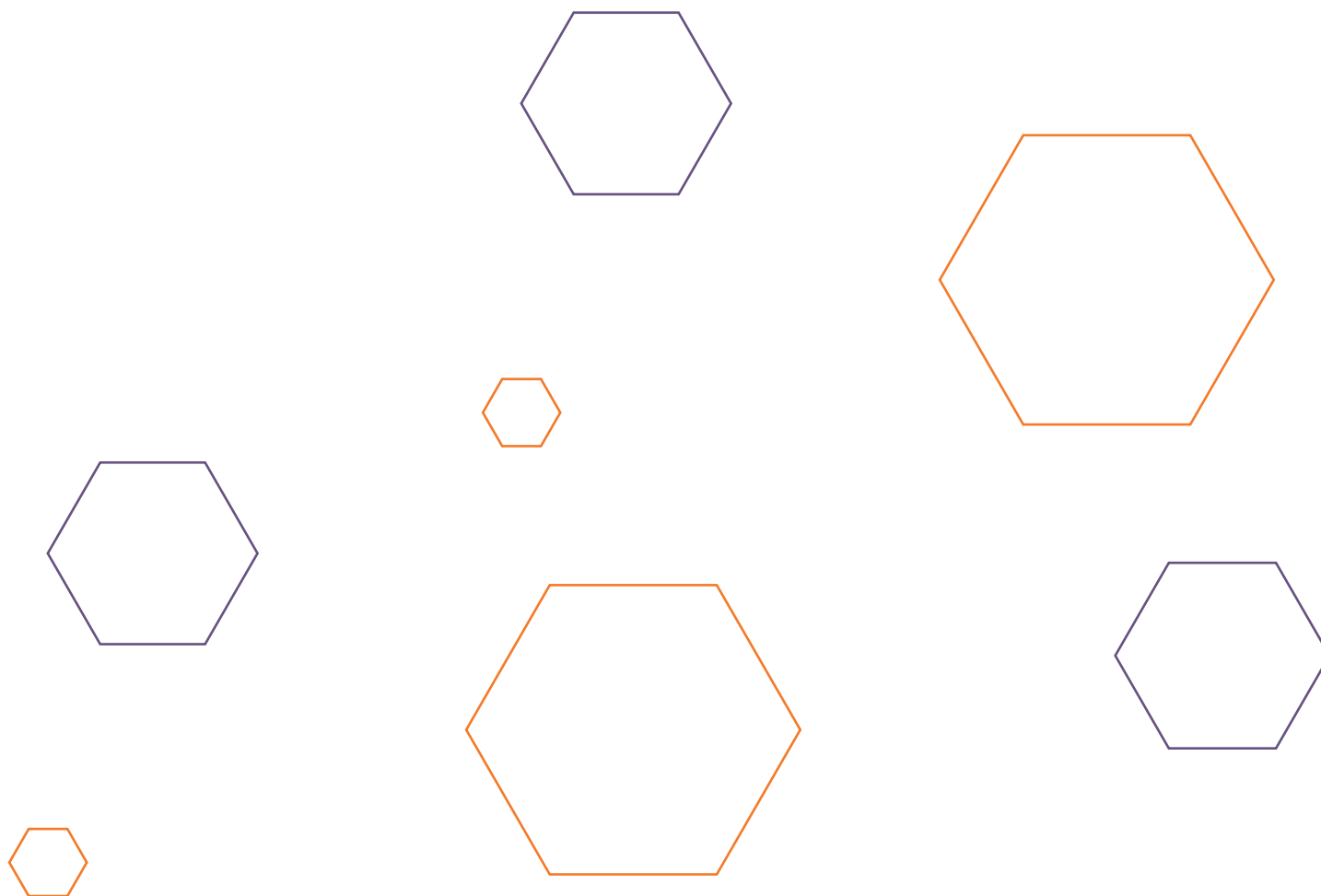
Tabela 3: Concentração de pesquisadores em função dos tópicos-*clusters* encontrados

PESQUISADOR	CLUSTER
Pesquisador A	6 - [saúde,translacional,identificação]
Pesquisador A	1 - [parto,validação,sobre]
Pesquisador A r	11 - [antiretroviral,inibidor,efeito]
Pesquisador A	7 - [vírus,zika,hepatite]
Pesquisador B	2 - [célula,carcinoma,regulação]
Pesquisador B	7 - [vírus,zika,hepatite]
Pesquisador B	8 - [estudo,ação,prática]
Pesquisador C	5 - [certificado,técnica,legislação]
Pesquisador C	3 - [ihmt,university,farmanguinhos]
Pesquisador C	9 - [ppg,metodologion,curso]
Pesquisador C	6 - [saúde,translacional,identificação]
Pesquisador C	10 - [dengue,analysis,translational]
Pesquisador C	4 - [tecnológica,abordagem,negligenciada]
Pesquisador C	11 - [antiretroviral,inibidor,efeito]
Pesquisador D	1 - [parto,validação,sobre]
Pesquisador D	8 - [estudo,ação,prática]
Pesquisador D	2 - [célula,carcinoma,regulação]

Fonte: Elaborado pelos autores



A partir dessas visualizações, podemos verificar que de fato a modelagem de tópicos atende às necessidades da proposta, visto que é possível identificar claramente os tópicos principais e suas relações, bem como, o relacionamento de pesquisadores intra-tópicos (Tabela 3). Além disso, é possível detectar o surgimento de novos tópicos de pesquisa de determinada comunidade de pesquisadores.



## ARTIGOS CIENTÍFICOS E PRODUTOS RESULTANTES

ÁGORA FIOCRUZ. **Ágora Elattes** – ferramenta de pesquisa e análise de redes de pesquisadores brasileiros. [Brasil, s.d.]. Disponível em: <https://agora.fiocruz.br/elattes/>. Acesso em 26 jul. 2023.

OLIVEIRA, J. F. D. *et al.* The global scientific research response to the public health emergency of Zika virus infection. **PLOS ONE**, v. 15, n. 3, 2020. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0229790>. Acesso em: 19/10/2022.

SAMPAIO, R. B. *et al.* Scientometric Analysis of Research Output from Brazil in Response to the Zika Crisis Using e-Lattes. **Journal of Data and Information Science**, v. 5, n. 4, p. 137–146, 1 nov., 2020.

SAMPAIO, R.; BATISTA-JR, A.; MENA-CHALCO, J. P. e-Lattes: Um novo arcabouço em linguagem R para análise do currículo Lattes. In: Encontro Brasileiro de Bibliometria e Cientometria, 6., (EBBC), 2018, Rio de Janeiro. **Anais eletrônicos** [...] Disponível em: <http://hdl.handle.net/20.500.11959/brapci/117661>. Acesso em: 23 jul. 2022.

## REFERÊNCIAS

AKAZAKI, J. M.; BUENO, L. R.; MENA-CHALCO, J. P. Caracterização de Pesquisadores Utilizando Indicadores Bibliométricos de Senioridade e Inatividade: Um Estudo de Caso dos Bolsistas de Produtividade 1A-CNPq, 2016. In: ENCONTRO BRASILEIRO DE BIBLIOMETRIA E CIENTOMETRIA, 2016. **Anais eletrônicos** [...] v. 5, n. , 2016. Disponível em: <http://hdl.handle.net/20.500.11959/brapci/45988>. Acesso em: 19 dez. 2022.

AL-MASRI, E.; MAHMOUD, Q. H. Discovering the best web service. In: 16TH INTERNATIONAL WORLD WIDE WEB CONFERENCE (WWW 2007), Banff, Alberta, 2007. **Proceedings** [...] Banff, Alberta. Disponível em: <https://thewebconf.org/www2007/proceedings.html>. Acesso em: 13/10/2022.

ALVES, A. D.; YANASSE, H. H.; SOMA, N. Y. LattesMiner: uma linguagem de domínio específico para extração automática de informações da Plataforma Lattes. In: XII WORKSHOP DE COMPUTAÇÃO APLICADA, 2012, São José dos Campos. **Anais [...]** São José dos Campos, MCTI/INPE, 2012.

ARANHA, C.; PASSOS, E. A Tecnologia de Mineração de Textos. **Revista Eletrônica de Sistemas de Informação**, v. 5, n. 2, p.1-8, 31 ago. 2006.

BLEI, D. M.; NG, A. Y.; JORDAN, M. I. Latent dirichlet allocation. **Journal of Machine Learning Research**, v. 3, p. 993-1022, mar. 2003.

BONACCORSO, G. **Machine learning algorithms**. Birmigham UK: Packt Publishing Ltd., 2017.

BUSCHMANN, F. **Pattern-Oriented Software Architecture**. [S.L] Wiley.1996. 476 p.

CHAPMAN, S. **Introduction to JavaScript**. 7 de abril de 2018

DIAS, T.M. R.; MOITA, G. F.; DIAS, P.M. Adoção da plataforma lattes como fonte de dados para caracterização de redes científicas. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, [S. l.], v. 21, n. 47, p. 16–26, 2016.

MAZUROWSKI, M. A. *et al.* Training neural network classifiers for medical decision making: The effects of imbalanced datasets on classification performance. **Nural Networks**, v. 21, n. 2–3, p. 427–436, mar. 2008.

MCLAUGHLIN, B. **What is HTML5**. [S.L], O'Reilly Media, 2011.

MITCHELL, T. M. **Machine Learning**. New York: McGraw-Hill, 1997.432 p.

MITCHELL, T. M. Machine learning and data mining. **Communications of the ACM** v. 42, n. 11, p. 30-36, 1999.

MOREIRA, P. S. D. C.; GUIMARÃES, A. J. R.; TSUNODA, D. F. Qual ferramenta bibliométrica escolher? um estudo comparativo entre softwares. **P2P E INOVAÇÃO**, v. 6, p. 140–158, 31 mar. 2020.

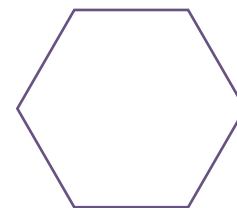
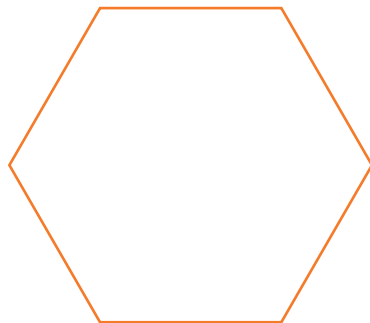
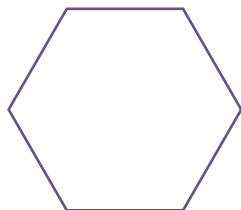
NASRABADI, N. M. Pattern recognition and machine learning. **Journal of electronic Imaging**, v.16, n.4, 2007. Disponível em: <https://spie.org/Publications/Journal/10.1117/1.2819119?SSO=1>. Acesso em: 13/10/2022.

PEZOA, F. *et al.* Foundations of JSON schema. In: 25th International World Wide Web Conference (WWW), Montreal, 2016. **Proceedings[...]**Montreal. Disponível em: <https://dl.acm.org/doi/proceedings/10.5555/2872427>. Acesso em 27/09/2022.

PRACIANO, B. J. C. *et al.* Spatio-temporal trend analysis of the Brazilian elections based on Twitter data. 2018. In: IEEE International Conference on Data Mining Workshops (ICDMW) 2018, Singapore. **Conference proceedings**. Disponível em: <https://www.computer.org/csdl/proceedings/icdmw/2018/18jXxWg9hao>. Acesso em 13/10/2022.

VAN ECK N. J.; WALTMAN, L. V. A new method for visualizing similarities between objects. In: Decker R, Lenz HJ, editors. **Advances in Data Analysis. Studies in Classification, Data Analysis, and Knowledge Organization**. Springer, Berlin: Heidelberg, p. 299–306.,2018.

ZURADA, J. M.; MARKS, R. J.; ROBINSON, J. Review of computational intelligence: imitating life. **IEEE Transactions of Neural Networks**, v. 6, n. 6, p. 1563, 1995.





# Feiras de Soluções para a Saúde

---

*Wagner de Jesus Martins  
Márcio Aldrin França Cavalcante  
Mônica Angélica Carreira Fragoso*

A constituição dos novos arranjos institucionais e de pesquisa, desenvolvimento e inovação buscam a maximização dos recursos disponíveis e a integração de competências, conhecimentos e atores, estimula a inovação, assim como, seu potencial transformador de bens e serviços em prol da sociedade. Uma base para a consolidação de uma ciência cidadã.

A legislação brasileira fornece os subsídios legais para definir na estrutura do sistema nacional de saúde, ações governamentais para incrementar o desenvolvimento científico e tecnológico por meio da promoção e do incentivo à pesquisa e à capacitação tecnológica. O Ministério da Saúde (MS) e as agências estaduais de fomento à pesquisa atuam diretamente no financiamento da capacidade produtiva do setor, assim gerando a formação de redes cooperativas entre as diversas instâncias de geração de conhecimento: decisória, normativa e regulatória; produtiva e desenvolvimentista. Ações geradas neste âmbito tiveram por objetivo implementar e aperfeiçoar os processos de absorção de conhecimento científico pelas indústrias, pelos serviços de saúde e pela sociedade, estando contidas na Meta 3 - Fortalecer propostas de colaboração/cooperação na rede do complexo produtivo da saúde (Instituições de Ciência e Tecnologia - ICT/empresas) - do projeto “Plataforma de Vigilância de Longo Prazo para Zika Vírus e Microcefalia no âmbito do SUS”.

Nesse contexto, as Feiras de Soluções formaram um programa com os seguintes objetivos:

- Identificar e mapear os atores, tecnologias nacionais/internacionais e horizontes tecnológicos:
  - Áreas envolvidas: biotecnologia, química, fármacos e medicamentos, equipamentos, assistência à saúde, tecnologias sociais, diagnósticos, controle e combate de vetor, equipamentos de proteção individual, tecnologias assistivas, comunicação/informação, patentes, artigos e mídias sociais;
- Potencializar a geração de produtos, processos e articulações intra e extra setores;



- Promover:
  - a educação em saúde e estimular a construção de consciência sanitária relacionada às arboviroses;
  - a cooperação em redes sociotécnicas;
  - parcerias institucionais para Ciência, Tecnologia & Inovação em Saúde (CT&IS);
  - a translação do conhecimento e pesquisa translacional.

No âmbito desta Plataforma, foram realizadas feiras tecnológicas, denominadas de Feiras de Soluções para a Saúde, em diferentes regiões do Brasil, de modo a identificar e disseminar inovações tecnológicas em saúde e contribuir, não somente com o enfrentamento das arboviroses, como também, que estivessem em consonância com a Agenda de Prioridades em Pesquisa do Ministério da Saúde, com os Objetivos de Desenvolvimento Sustentável (ODS) e a Agenda 2030 da Organização das Nações Unidas.

Foram realizadas quatro feiras de 2017 a 2020, sendo três presenciais: Salvador - BA (08 a 10/08/2017), Bento Gonçalves-RS (22 a 25/04/2019) e Fortaleza - CE (16 a 18/10/2019), e uma na modalidade virtual, devido à Emergência de Saúde Pública imposta pelo Vírus SARS-Cov2 (09 a 11/12/2020)<sup>1</sup>.

Tais iniciativas tiveram como objetivo implementar e aperfeiçoar os processos de absorção de conhecimento científico pelas indústrias, pelos serviços de saúde e pela sociedade. Esse arranjo institucional inovador atuou como apoio à gestão estratégica de redes cooperativas no contexto da saúde pública e contribuiu para a consolidação do Sistema de Ciência, Tecnologia e Inovação em Saúde, de acordo com as quatro características de mudanças organizacionais em curso e traduzidas da seguinte forma: 1) integração organizacional com configurações descentralizadas e interconectadas em

---

1. Disponível em: <https://solucoes.agora.fiocruz.br/>

rede; 2) capacidade de realizar melhorias contínuas, com permanente capacitação de modo a enfrentar e resolver os problemas; 3) agilidade e adaptabilidade nos processos produtivos, para responder às necessidades da demanda segmentada; 4) organizações estruturadas como sistemas abertos, criando vínculos sistêmicos de cooperação e redes interorganizacionais.

As organizações constituem-se em redes, unidades e atores interconectados, cada vez mais dependentes do estabelecimento de uma relação comunicacional entre usuários, fornecedores e concorrentes, formando assim, uma rede de informação/comunicação para realização de ações pactuadas.

Guiados pelos eixos da Política Nacional de Ciência, Tecnologia e Inovação em Saúde (PNCTI) e da necessidade de fortalecer as redes de interações, entre instituições e demais sujeitos envolvidos no processo de ciência, tecnologia, inovação em saúde (CTI/S), foi articulada a ideia de realização de Feiras Tecnológicas Regionais para identificação e disseminação de inovações tecnológicas ligadas às arboviroses, em 2016. A feira, realizada em Salvador em 2017, teve o objetivo de oportunizar espaços de diálogos, de socialização e de disseminação do conhecimento, assim como, ativar redes de cooperação para a apresentação e difusão de inovações tecnológicas, sociais e de serviços de saúde aplicadas à Zika, Dengue e Chikungunya e à Síndrome Congênita do Zika (SCZ).

A feira foi uma estratégia, um projeto de caráter científico e inovador, pois permitiu o intercâmbio de experiências, conhecimentos e a formação de parcerias e redes cooperativas entre públicos de três pilares da ciência e tecnologia em saúde (social, tecnológico e de serviços), bases para a ciência cidadã. A Figura 1 mostra um dos momentos de trocas de experiência na rede cooperativa entre membros de sociedade e técnicos dos serviços de saúde local.



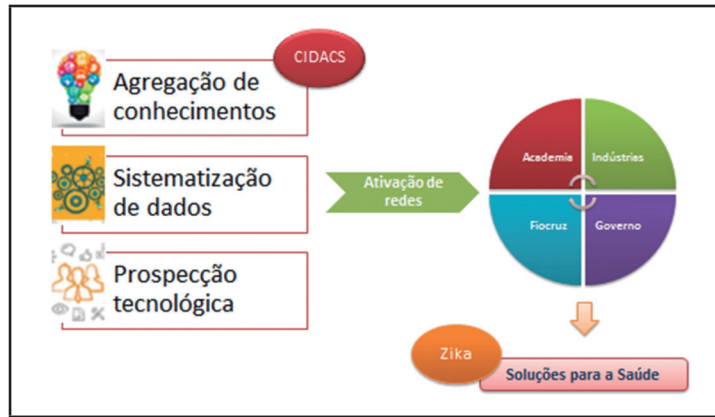
Figura 1: Oficina que reuniu fisioterapeutas, fonoaudiólogos, psicólogos e família. Feira de Soluções para a Saúde. Salvador, 2017



Fonte: Arquivo pessoal

Entretanto, elas não se caracterizaram pela realização de negócios, nem por encontros acadêmicos. Foram desenvolvidas para proporcionar um ambiente de integração de forças da sociedade, capazes de encontrar soluções para um problema de saúde pública imediato, ou a longo prazo, no território de forma saudável e sustentável. Propiciaram a criação de espaços de diálogos, socialização, disseminação do conhecimento e oportunidades entre diferentes atores, com intercâmbio de experiências; ativaram redes de cooperação para a apresentação, difusão e sustentação de inovações tecnológicas, sociais e de serviços denominadas como Soluções para a Saúde. A Figura 2 apresenta o esquema do arranjo concebido para a disseminação do conhecimento durante as Feiras de Soluções para a Saúde.

Figura 2: Arranjo da disseminação do conhecimento nas Feiras de Soluções



Fonte: Elaboração dos autores

### Organização das Feiras de Soluções para Saúde

As Feiras de Soluções para a Saúde objetivaram as discussões de questões de saúde pública relevantes, como a emergência de saúde pública causada pelo vírus Zika e da pandemia do novo coronavírus, Agenda 2030, *One Health*, Territórios Saudáveis e Sustentáveis, dentre outros temas.

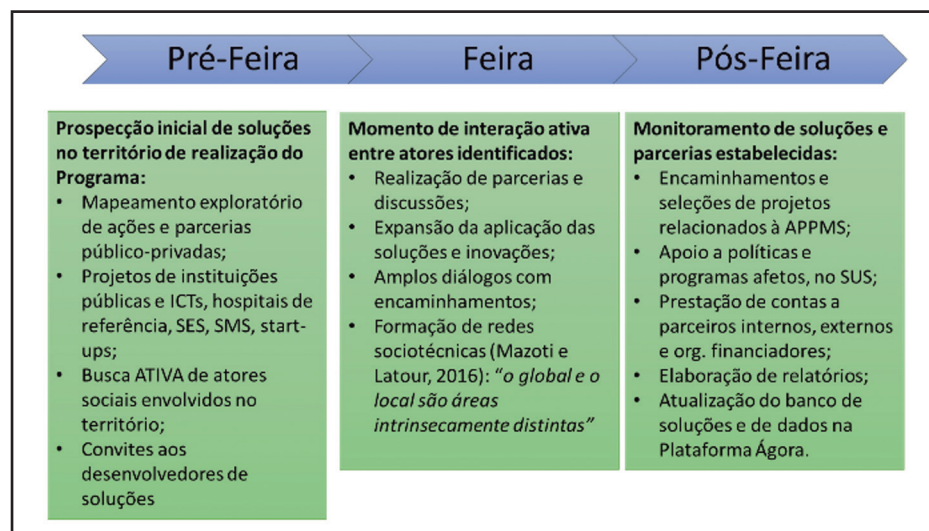
Foram elaboradas em três etapas principais de organização: pré-Feira, Feira e pós-Feira. Na etapa pré-Feira, ocorreu a prospecção inicial de soluções para a saúde no território de realização da feira *per se*, assim: inicialmente, foram mapeadas as ações e parcerias, a partir da busca em bases de dados e na internet, de caráter exploratório, e a partir da combinação de palavras-chave com operadores booleanos. Esta etapa visou a busca de parcerias público-privadas em andamento, projetos de instituições públicas de referência, laboratórios de vigilância em saúde, universidades, grupos de pesquisa, hospitais de referência, Secretarias Estaduais e Municipais de Saúde, empresas e *start-ups*. Em seguida, foi realizada uma busca ativa no território de atores sociais de relevância local e parceiros estratégicos, bem como, convite aos responsáveis pelo desenvolvimento de iniciativas e soluções identificados para cadastrar as soluções no *website* da Feira.

A etapa Feira, realizado pela Fiocruz de forma oportuna e integrada a outros eventos pré-programados no território, promoveu o entrosamento ativo dos atores identificados na etapa anterior, permitindo o conhecimento e o diálogo no contexto da aplicação da inovação, solução ou do produto prospectado.

O processo de ativação de redes sociotécnicas foi induzido na programação e no desenrolar do evento, refletindo a tônica do amplo diálogo entre demandantes e desenvolvedores das soluções em saúde, nos espaços reservados a atividades de rodas de conversas, seminários, apresentação de soluções, *hackathons* e oficinas interativas.

Ao final de cada Feira foi elaborado um relatório com base nos objetivos previamente determinados. Entretanto, somente ao final de 2022, foi encaminhado à SVSA/MS, o relatório final contendo os resultados das quatro Feiras de Soluções para a Saúde realizadas entre os anos de 2017 e 2022 e as atualizações realizadas na Plataforma Ágora.

Figura 3: Descrição Metodológica das Feiras



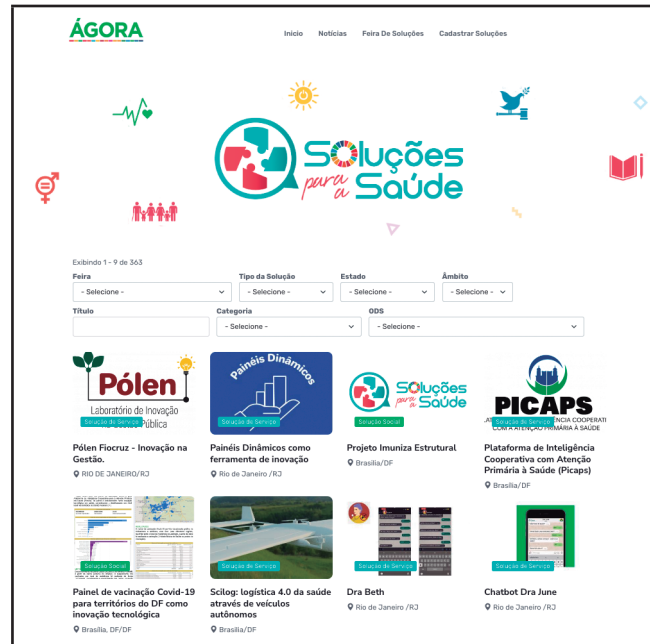
Fonte: Elaboração dos autores

Foram realizadas quatro Feiras de Soluções para a Saúde e cada uma delas teve uma temática central diferente. Na primeira versão do evento, realizada em 2017, na cidade de Salvador - BA, o tema foi o “Zika vírus, microcefalia e síndromes congênitas”.

Em 2019, foram realizadas duas edições: a primeira em Bento Gonçalves – RS, com o tema “Saúde Única para Territórios Saudáveis e Sustentáveis”, e a segunda em Fortaleza – CE, com o tema “Saúde Digital para Territórios Saudáveis e Sustentáveis”. Em 2020, na sua quarta edição, as condições de saúde pública vigentes impuseram a readequação para a modalidade totalmente digital, em decorrência da pandemia causada pelo novo coronavírus, e teve como tema o “Enfrentando as crises sanitárias e epidemias: panoramas e perspectivas”.

A inclusão de uma solução para a saúde como participante de uma Feira estava vinculada ao seu cadastramento na Plataforma Ágora (Fig. 4), instalada em 2012 por iniciativa da Fundação Oswaldo Cruz - Brasília, tendo como objetivo constituir uma rede virtual na área da saúde, um espaço agregador e sistematizador das informações que fluem no ciberespaço de forma articulada à produção cotidiana do Sistema Único de Saúde (SUS).

Figura 4: Plataforma Ágora - Banco de soluções



Fonte: Site [agora.fiocruz.br](http://agora.fiocruz.br)

O uso da Plataforma Ágora (FUNDAÇÃO OSWALDO CRUZ, 2021) permitiu a geração e consolidação de um banco de dados (hoje com 349 soluções), que poderá servir de base para estudos continuados de inovação em saúde pública, e cuja divulgação ampla promoverá a maior integração entre ciência e o universo social e empresarial.

O cadastramento das soluções ocorreu dentro de três grandes eixos: industriais, serviços e sociais. Cada eixo foi organizado em categorias: Atenção, Comunicação e Educação, Controle Social e Participação Popular, Diagnóstico, Gestão e Prevenção.

## Promoção da Saúde, Vigilância em Saúde e Tratamento.

Devido às diferenças na organização de cada feira e da temática abordada, foi utilizada metodologia própria para análise dos resultados, de modo a: propiciar a construção de uma matriz de aspectos positivos e de melhorias; apoiar a sistematização e avaliação do banco de soluções quanto ao nível de maturidade tecnológica para as soluções industriais; e desenvolver ou adaptar metodologia de avaliação do nível de consistência, ou adequabilidade de soluções sociais e de serviços.

O desenvolvimento da metodologia selecionou uma amostra de análise do tipo estratificada proporcional, calculada da seguinte forma:

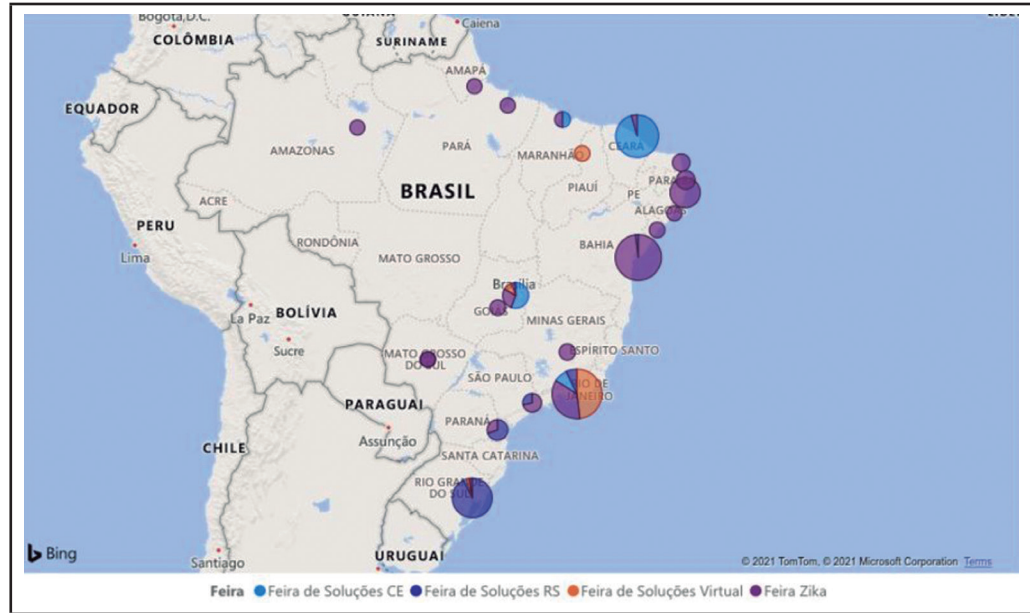
1º) A amostra geral de cada feira foi composta por 20% das soluções cadastradas em cada evento;

2º) A partir deste dado, a amostra foi distribuída em estratos de forma proporcional aos tipos de soluções (serviços, social e industrial) de cada Feira;

3º) Seleção das soluções com base em critérios:

- Preenchimento completo da ficha de inscrição na Plataforma Ágora (detalhes da solução, informações para contato, etc.);
- Leitura do título e descrição da solução, a qual abrangeu a característica inovadora em saúde e o potencial de replicabilidade;
- Classificação de acordo com os oito eixos estratégicos para orientação das inscrições nas feiras tecnológicas (vetor, prevenção, diagnóstico, tratamento, cuidado, vigilância em saúde, gestão e outros);
- Soluções que se relacionavam com os Objetivos de Desenvolvimento Sustentável (ODS).

Figura 5: Soluções distribuídas por UFs



Fonte: Elaboração dos autores

## A Feira de Salvador (BA)

A Feira de Soluções em sua primeira edição, foi realizada em Salvador (08 a 10/08/2017), tendo como objetivo oportunizar espaços de diálogos, de socialização, de disseminação do conhecimento, assim como, ativação de redes de cooperação para a apresentação e difusão de inovações tecnológicas, sociais e de serviços de saúde aplicadas à Zika, dengue, chikungunya e à Síndrome Congênita do Zika (SCZ)<sup>2</sup>.

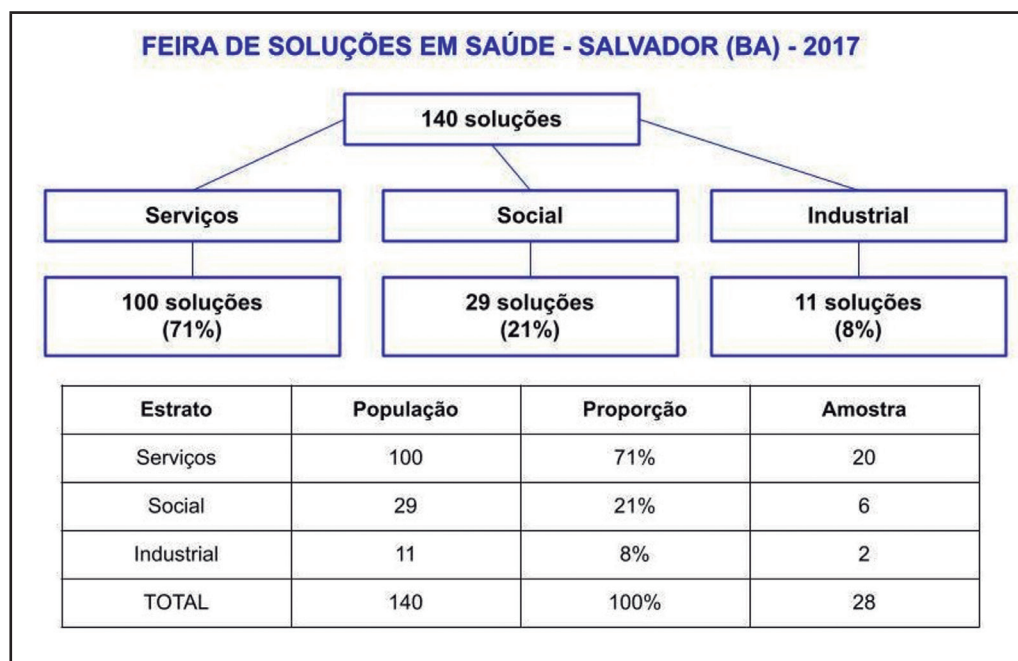
Foram cadastradas no *website* da Feira 140 soluções, abrangendo temas relacionados à reabilitação de crianças com Síndrome Congênita do Zika, meio ambiente, sa-

2. Vídeo sobre a Feira Zika encontra-se disponível em: <https://bit.ly/videofeirazika>



neamento, educação, assistência social, controle do vetor, promoção da saúde e da cidadania, vacinas, kits diagnósticos, desenvolvimento comunitário, direitos sexuais e direitos reprodutivos, equidade de gênero e raça, inclusão e proteção social, aplicativos e plataformas tecnológicas, metodologias de formação e qualificação para profissionais da saúde e equipamentos para estimulação precoce e redução de dores, de baixo custo<sup>3</sup>.

Figura 6: Soluções submetidas na Feira de Soluções para a Saúde, Feira Zika, realizada em Salvador - BA, 2017



Fonte: Correia & Martins, 2021. No prelo

3. A análise completa da Feira de Soluções para a Saúde de Salvador está disponível no Relatório Final da Feira de Soluções para Saúde Aplicadas às Arboviroses e à Síndrome Congênita do Zika. Disponível em: <https://bit.ly/relatoriofeiradesaude>



## A Feira de Bento Gonçalves (RS)

Realizada entre 22 e 25 de abril (2019) e com a temática principal da “Saúde Única para Territórios Saudáveis e Sustentáveis”, abriu espaço para discussão, proposições e apresentação de soluções inovadoras relacionadas aos Objetivos de Desenvolvimento Sustentável (ODS) e para o conceito de *One-Health* (Saúde Única – que articula as saúdes humana, animal e meio ambiente). Contou com a novidade da realização de um *Datathon* - maratona de dados de caráter competitivo, onde profissionais da saúde, desenvolvedores e profissionais de análise de dados buscaram desenvolver soluções com infográficos que facilitaram a comunicação das informações de base de dados com o público interessado<sup>4</sup>.

Em consonância com a concepção do projeto, a feira realizada em Bento Gonçalves ocorreu conjuntamente com o Congresso Anual do Conselho de Secretarias Municipais de Saúde do Estado do Rio Grande do Sul (Cosems/RS). Foi considerado um grande evento de saúde pública no estado, proporcionando a ativação de redes de cooperação e comunicação com diversos atores para além do estado do Rio Grande do Sul<sup>5</sup>.

A Feira de Bento Gonçalves em números:

- Total de participantes: 1.589
- Instituições parceiras: 07 (Cosems, Fiocruz, SES-RS, Governo do RS, Conasems, Ministério da Saúde)
- Número de municípios participantes: 417
- Número de secretários participantes: 275
- Número de representantes de secretários: 28

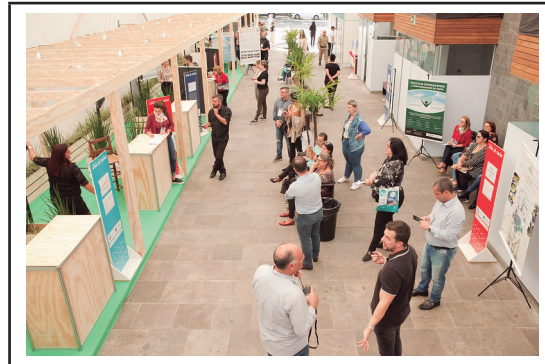
---

4. Um Boletim Informativo desta segunda edição da Feira de Soluções para a Saúde (Disponível em: <https://bit.ly.com/TzuXbdQ>) foi elaborado com relatos, links de vídeos, entrevistas, descrições de soluções tecnológicas e relato de parcerias que refletem seu impacto na promoção da saúde e no fortalecimento dos ODS no território.

5. Resumo detalhado do evento encontra-se disponível em: <https://bit.ly.com/nzcjaQp>.

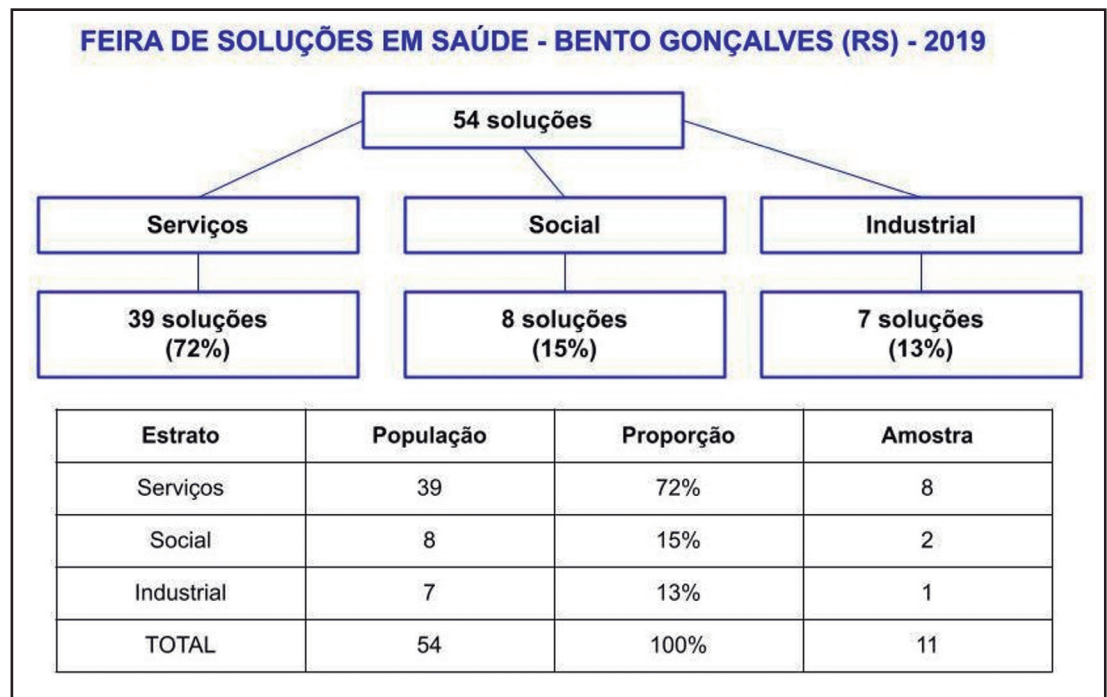
- Mesas redondas: 31
- Apresentações/exposições: 109
- Palestras: 6
- Reuniões com parceiros estratégicos: 14
- Total de Trabalhos da Mostra: 328 inscritos / 115 apresentados
- Soluções para a Saúde: 54

Figura 7: Feira de Soluções para a Saúde, realizada em Bento Gonçalves - RS, 2019



Fonte: Arquivo pessoal

Figura 8: Soluções submetidas na Feira de Soluções para a Saúde, realizada em Bento Gonçalves - RS, 2019



Fonte: Correia & Martins, 2021. No prelo

### A Feira de Fortaleza (CE)

A Feira de Soluções para Saúde, em sua terceira edição, realizada em Fortaleza, Ceará, entre os dias 16 e 19 de outubro (2019), teve como principal objetivo consolidar a metodologia para prospectar, fomentar e fortalecer ações de cooperação e geração de conhecimento em rede, para o fortalecimento do Complexo Econômico e Industrial da Saúde (Ceis) e de parcerias academia-governo-universidades-gestão em saúde.

A feira ocorreu durante a Terceira Feira do Conhecimento do estado do Ceará (CE), simultaneamente com o Primeiro Congresso dos Municípios do Ceará, parceria capital para a ampliação do escopo do evento, oferecendo uma vasta gama de atividades e oportunidades de conhecimento e de colaboração no território.

A temática escolhida foi “Saúde Digital na Interface de Territórios Saudáveis e Sustentáveis”. A era digital trouxe para a saúde grandes avanços, mas também incertezas. Assim, a discussão desta temática não poderia prescindir da discussão em um evento onde o *core* é o diálogo e a formação de redes sociotécnicas.

A atenção e o cuidado da saúde devem ser vistos em uma percepção holística, focada no indivíduo, com suas subjetividades e interações com o ambiente, com o território. Assim, os Objetivos de Desenvolvimento Sustentáveis (ODS) e a Agenda 2030 compõem o roteiro ideal para que os diferentes atores (academia, gestão, sociedade e serviços) em seus territórios, possam trabalhar de forma integrada os eixos da Saúde, Educação, Meio Ambiente, Desenvolvimento Sustentável.

Dentre as atividades da feira destacou-se a Oficina “O SUS e os Desafios da Saúde Digital no Brasil”, cuja tônica foi abordar os desafios para a incorporação da saúde digital no SUS, abordando os temas:

- Diagnósticos e perspectivas da Saúde Digital no Brasil.
- Fomento à inovação em Saúde Digital no Brasil.
- O papel da Fiocruz frente aos desafios da Saúde Digital no Brasil.

Figura 9: Terceira edição da Feira de Soluções para a Saúde, Fortaleza/CE, 2019



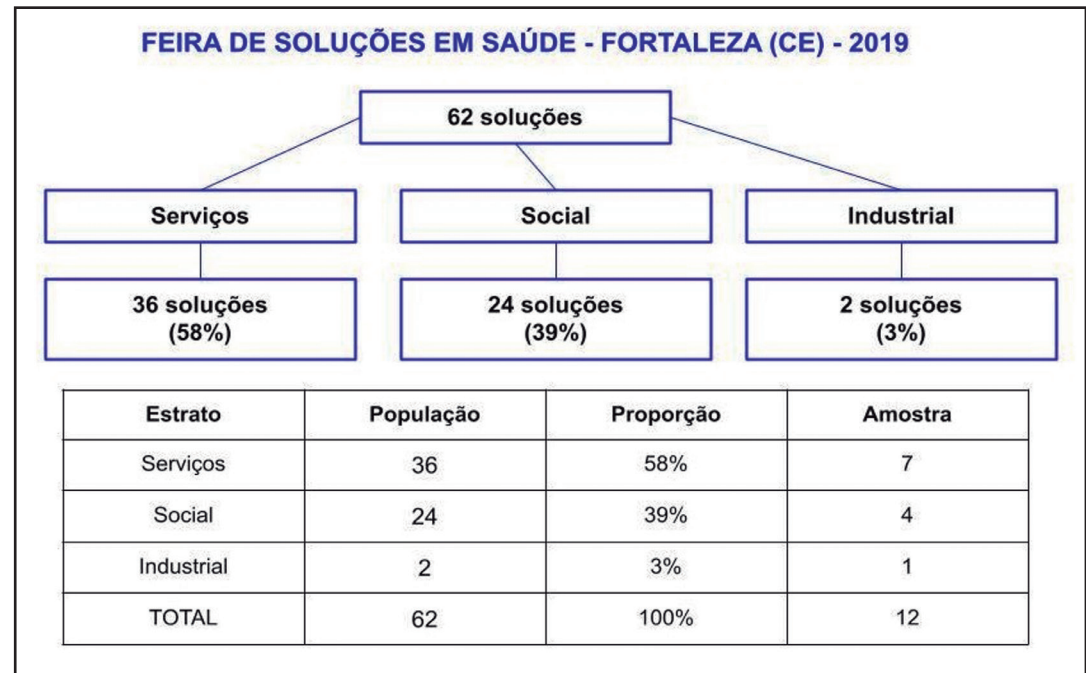
Fonte: Arquivo pessoal

A Estratégia Global para a Saúde Digital 2020-2025<sup>6</sup>, publicada em 2021 pela Organização Mundial da Saúde, estabeleceu em seus objetivos estratégicos, a orientação sobre a transformação global da saúde, o fortalecimento de iniciativas e partes interessadas para melhoria da saúde, bem como, a mitigação dos riscos associados a esta transformação. Assim, essa oficina, oportunamente discutiu a situação da saúde pública de precisão; o investimento na instrumentalização, na capacitação de pessoas da área e na qualidade dos dados produzidos; e a integração de polos de cooperação tecnológica. A Saúde Digital no Brasil tem potencial para tornar-se modelo para outros países, devido às características únicas do nosso sistema de saúde.

Foram cadastradas na Plataforma Ágora para esta edição da Feira, 62 soluções inovadoras para a saúde, que além de separadas por tipo de tecnologia, foram classificadas preliminarmente de acordo com seu enquadramento nos ODS.

6. Global Strategy on Digital Health 2020-2025, disponível em <https://bit.ly/estrategiaglobal>

Figura 10: Soluções submetidas na Feira de Soluções para a Saúde, realizada em Fortaleza - CE, 2019



Fonte: Correia & Martins, 2021. No prelo

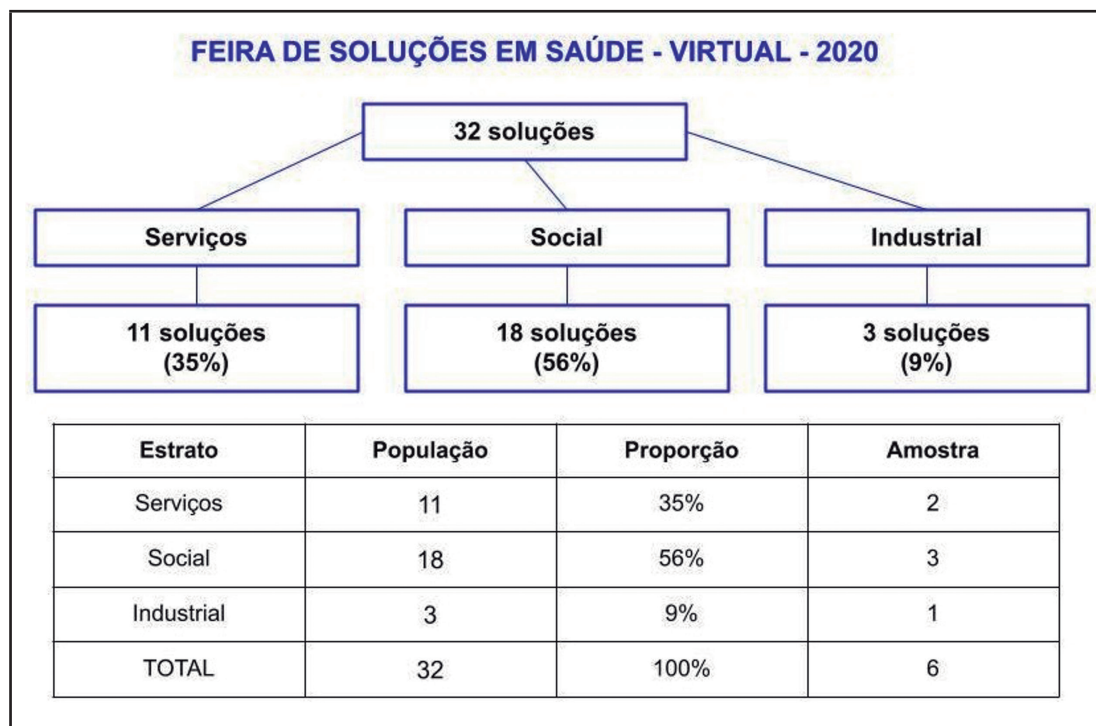
## A Feira Virtual

A quarta edição da Feira de Soluções para a Saúde foi realizada de 9 a 11 de dezembro (2020), em formato totalmente virtual, com o tema “Enfrentando as crises sanitárias e epidemias: panoramas e perspectivas”.

A pandemia pelo SARS-COV-2 impôs forte demanda de desenvolvimento de soluções industriais, organizacionais e virtuais para fazer frente à contenção do vírus e mitigação das consequências provocadas por ele, e dar seguimento à vida de milhares de

peças. A Feira de Soluções para a Saúde foi produzida com o objetivo de intensificar o uso do conhecimento originado em institutos de pesquisa, nas universidades e na própria comunidade, e incentivar a rápida transformação em tecnologias e inovações úteis para a população.

Figura 11: Soluções submetidas na Feira de Soluções para a Saúde, realizada na modalidade virtual em 2020



Fonte: Correia & Martins, 2021. No prelo

A Plataforma Ágora foi o canal utilizado para o cadastro das soluções e favoreceu a criação de redes de produção e difusão de conhecimento, uma vez que promoveu a criação de um banco de dados, de acesso amplo e de interação de conhecimentos de diferentes níveis, desde o nível acadêmico, passando pela assistência e gestão das políticas públicas de saúde.

Nesse contexto, foram cadastradas 32 soluções para a saúde, categorizadas em: serviço, social e industrial e posteriormente, submetidas à avaliação quanto à maturidade tecnológica.

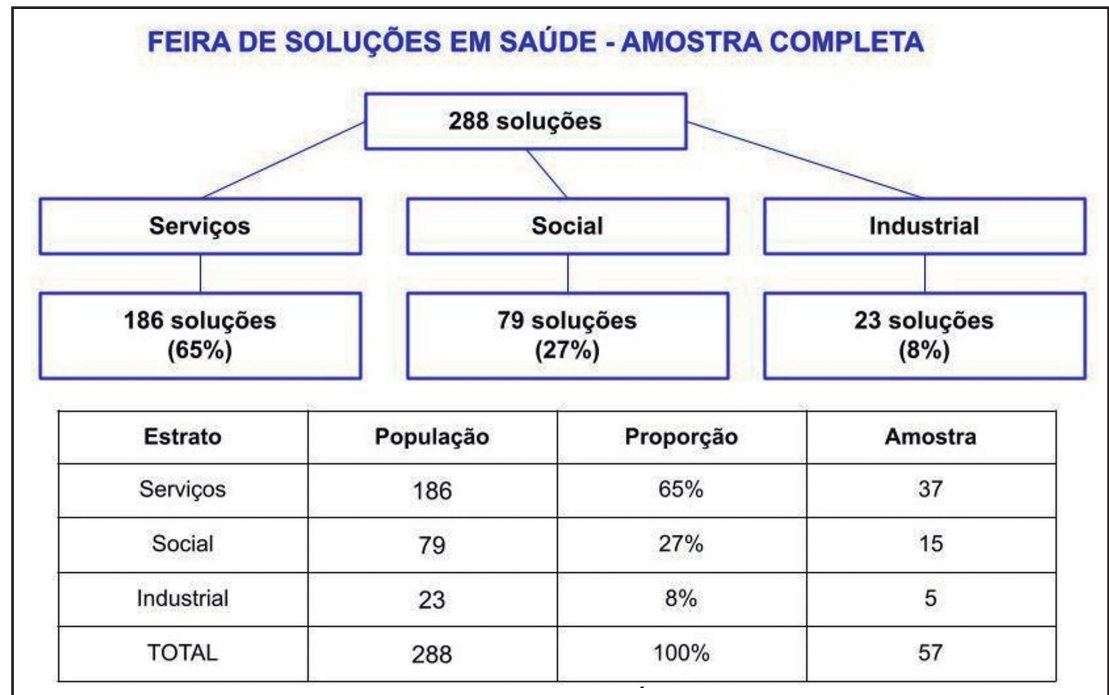
## RESULTADOS

Em suas quatro edições, as Feiras de Soluções para a Saúde cadastraram na Plataforma Ágora 288 soluções inovadoras para a saúde, sendo em maioria as soluções sociais e de serviço.

Notadamente, as soluções industriais, que demandam maior investimento tecnológico para o desenvolvimento e escalonamento, foram as de menor número. Espaços de diálogo e apresentação, como os do tipo feiras de soluções, propiciam a aproximação entre demandantes, academia, desenvolvedores dessas soluções e o Complexo Econômico e Industrial da Saúde (Ceis), como também, a formação de redes de sociotécnicas que podem mitigar esta lacuna.



Figura 12: Total de Soluções submetidas nas quatro Feiras de Soluções para a Saúde, 2021



Fonte: Correia & Martins, 2021. No prelo

### Resultados Quantitativos das Feiras de Soluções para a Saúde

Devido às diferenças na organização de cada feira e da temática abordada, foi desenvolvida metodologia própria para a análise dos resultados das Feiras de Soluções para a Saúde. Uma amostra de análise do tipo estratificada proporcional, de 20% das soluções cadastradas em cada evento, foi utilizada para a sistematização e avaliação do banco de soluções quanto ao nível de maturidade tecnológica para as soluções industriais, como também, desenvolver ou adaptar metodologia de avaliação do nível de consistência, ou adequabilidade de soluções sociais e de serviços.

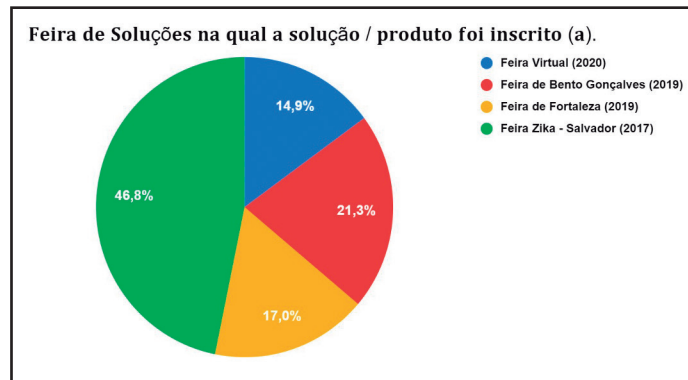
Aplicados os critérios metodológicos, 57 soluções foram selecionadas, representando 20% de todas as soluções cadastradas na Plataforma Ágora. Entretanto, somente 47 aceitaram participar da pesquisa de avaliação, sendo 16 soluções sociais, 3 soluções industriais e 28 soluções de serviços (Figura 13).

Figura 13: Estruturação dos resultados da pesquisa, 2021



Fonte: Correia & Martins, 2021. No prelo

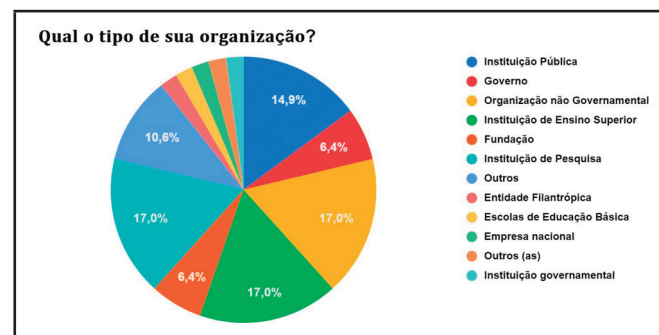
Figura 14: Distribuição das iniciativas selecionadas por Feira de Soluções para a Saúde, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

A Feira Zika (Salvador, 2017) foi a que teve maior número de soluções selecionadas, com 46,8% (n= 22), seguida por 21,3% de Bento Gonçalves - RS (n=10), 17% de Fortaleza - CE (n= 8) e 14,9% da edição virtual (n = 7).

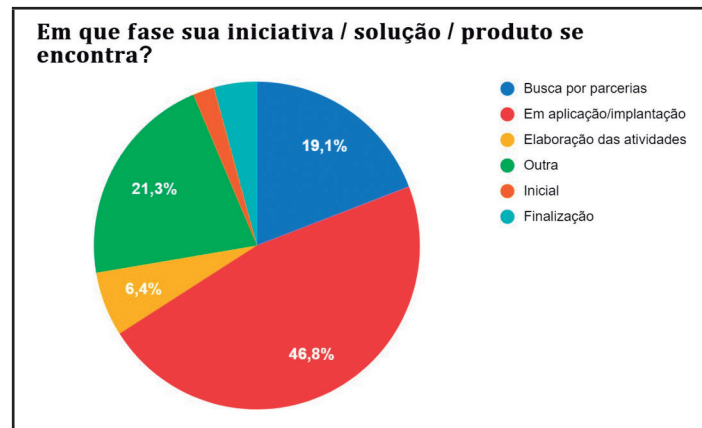
Figura 15: Distribuição das iniciativas selecionadas por tipo de organização, Brasil – 2021



Fonte: Correia & Martins, 2021. No prelo

Em relação à distribuição por tipo de organização que cadastrou soluções para a saúde, nota-se que houve participação de diversas instituições: 17% (ONGs, Instituições de Ensino Superior e Instituições de Pesquisa,  $n = 8$ ); 14,9% (Instituições Públicas,  $n = 7$ ); dentre outras. Contudo, ressalta-se o equilíbrio de participação entre as instituições.

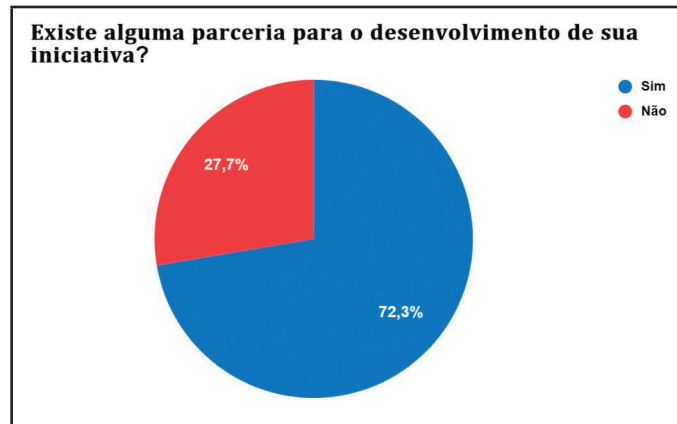
Figura 16: Distribuição das iniciativas selecionadas por fases de desenvolvimento, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

A maioria das soluções selecionadas encontra-se no estágio de implantação/aplicação com 46,8% ( $n = 22$ ), 21,3% em outras ( $n = 10$ ) e 19,1% em busca de parcerias ( $n = 9$ ). Das 47 participantes da pesquisa, 72,3% referiram possuir parceiros ( $n = 34$ ), que ocorre por meio de apoio técnico, divulgação, conhecimento técnico/científico, materiais e equipamentos, insumos ou recursos financeiros.

Figura 17: Distribuição das iniciativas selecionadas com parceria, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

Na Figura 18, são exibidas as nove categorias nas quais as soluções foram enquadradas.

Figura 18: Distribuição das iniciativas selecionadas de acordo com as categorias de ação, Brasil – 2021



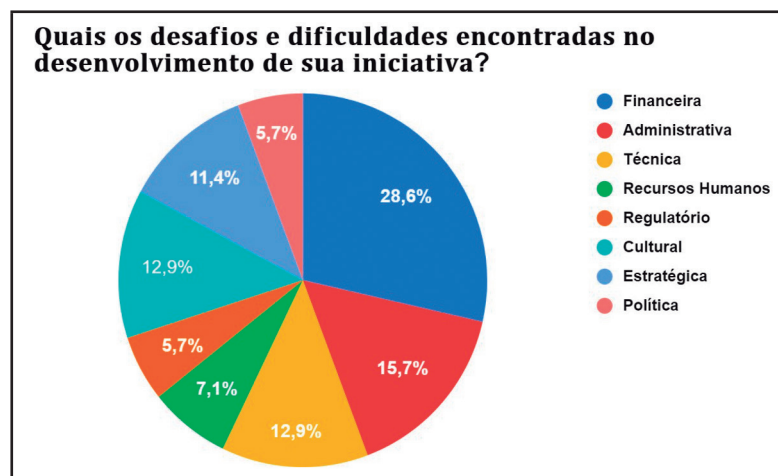
a. Controle do vetor (forma de barreiras, controle, combate, tecnologia social); b. Prevenção (individual, coletiva, vacina, tecnologia social); c. Diagnóstico (reagentes, kits laboratoriais, imagem, tecnologia social); d. Tratamento (fármaco, medicamento, práticas integrativas e complementares, tecnologia social); e. Assistência / Cuidado (tecnologia social, tecnologia assistiva, práticas integrativas e complementares); f. Vigilância em saúde (monitoramento, sistemas de informação, notificação...); g. Gestão (ações e serviços; regulação); h. Comunicação em Saúde; i. Outros.

Fonte: Correia & Martins, 2021. No prelo

A discrepância entre os números de soluções deve-se ao fato de algumas delas poderem estar incluídas em mais de uma categoria, sendo a maioria como propostas voltadas a: melhoria na assistência/cuidado (n = 25), comunicação em saúde e prevenção (n = 17 cada).

Quando foram avaliadas quanto às principais dificuldades para o desenvolvimento e/ou ampliação das iniciativas, 28,6% (n = 20) destacaram a ausência de recursos financeiros como principal desafio, seguidos dos aspectos administrativos, com 15,7% (n = 11).

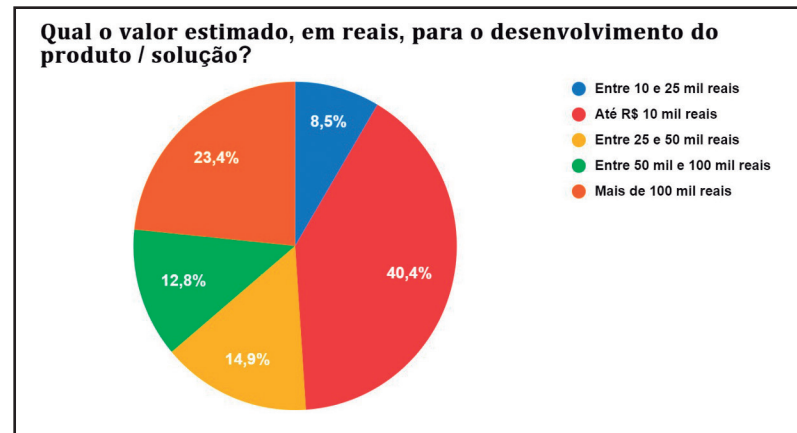
Figura 19: Distribuição das iniciativas selecionadas de acordo com os seus desafios e dificuldades, Brasil – 2021



Fonte: Correia & Martins, 2021. No prelo

Quanto ao custo de desenvolvimento, as soluções industriais foram as que mais se destacaram (acima de 100 mil reais). Contudo, grande parte das soluções selecionadas possuem custo estimado de até 10 mil reais (n = 19; 40,4%).

Figura 20: Distribuição das iniciativas selecionadas de acordo com o valor estimado para desenvolvimento das soluções - 2021

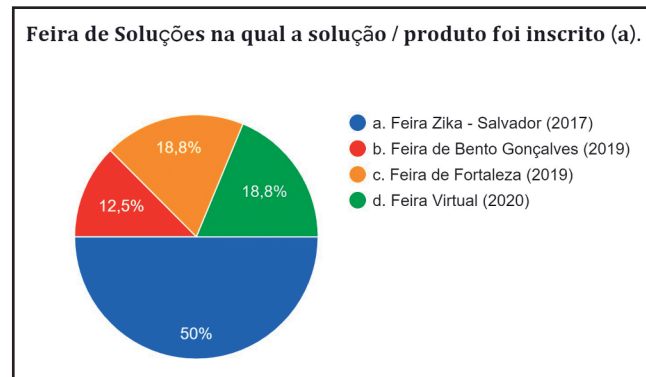


Fonte: Correia & Martins, 2021. No prelo

## Soluções Sociais

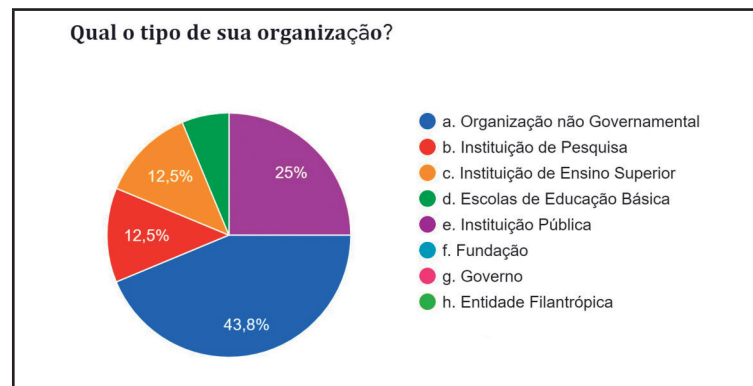
Metade das Soluções Sociais selecionadas foram apresentadas na Feira Zika (2017) (n=8), seguida das Feiras de Fortaleza (2019) e virtual (2020), com três soluções cada.

Figura 21: Distribuição das soluções sociais selecionadas por Feira de Soluções para a Saúde, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

Figura 22: Distribuição das soluções sociais selecionadas por tipo de organização, Brasil - 2021

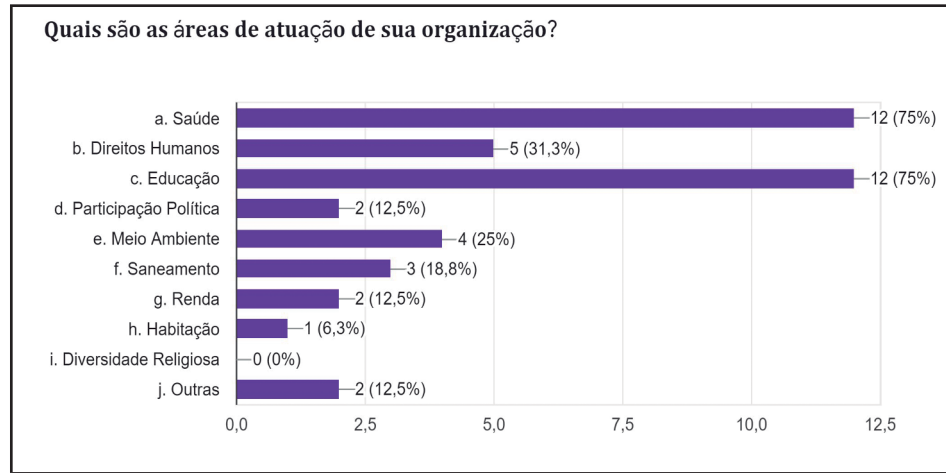


Fonte: Correia & Martins, 2021. No prelo

Destaca-se que as Organizações não Governamentais ( $n = 7$ ) foram as que mais cadastraram soluções sociais nas Feiras de Soluções para a Saúde.

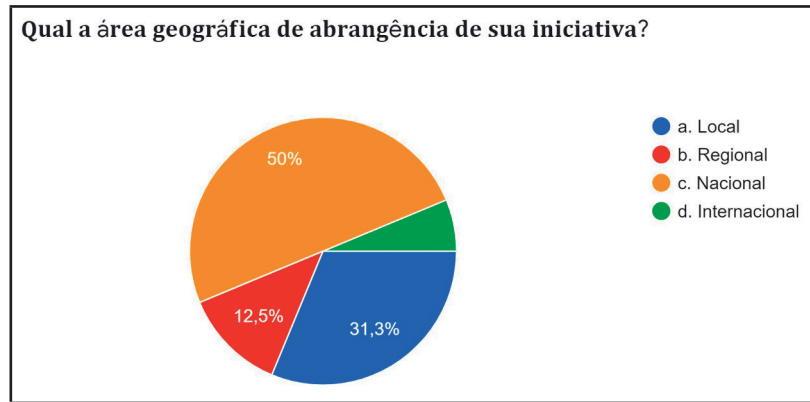


Figura 23: Distribuição das soluções sociais selecionadas por área de atuação das organizações, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

Figura 24: Distribuição das soluções sociais selecionadas por área geográfica de abrangência, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

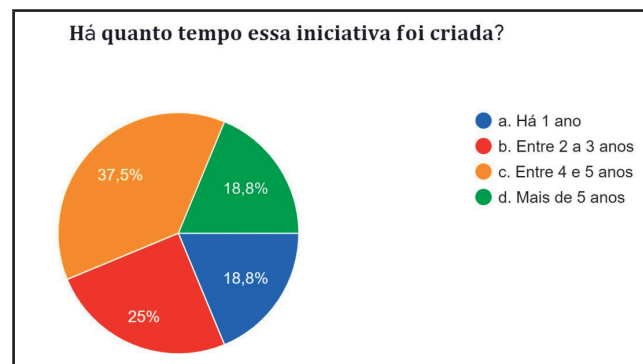
Dentre as soluções sociais, 50% tinham alcance nacional (n = 8), 31,3% local (n = 5), 12,5% regional (n = 2) e 6,3% internacional (n = 1). Têm como público-alvo especialmente a sociedade (56,3%) e foram criadas há mais de quatro anos (56,3%).

Figura 25: Distribuição das soluções sociais selecionadas segundo público-alvo, Brasil – 2021



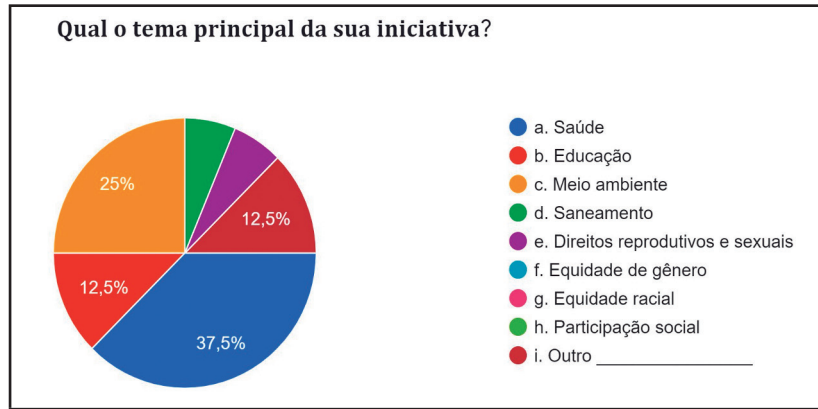
Fonte: Correia & Martins, 2021. No prelo

Figura 26: Distribuição das soluções sociais selecionadas segundo tempo de criação da iniciativa, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

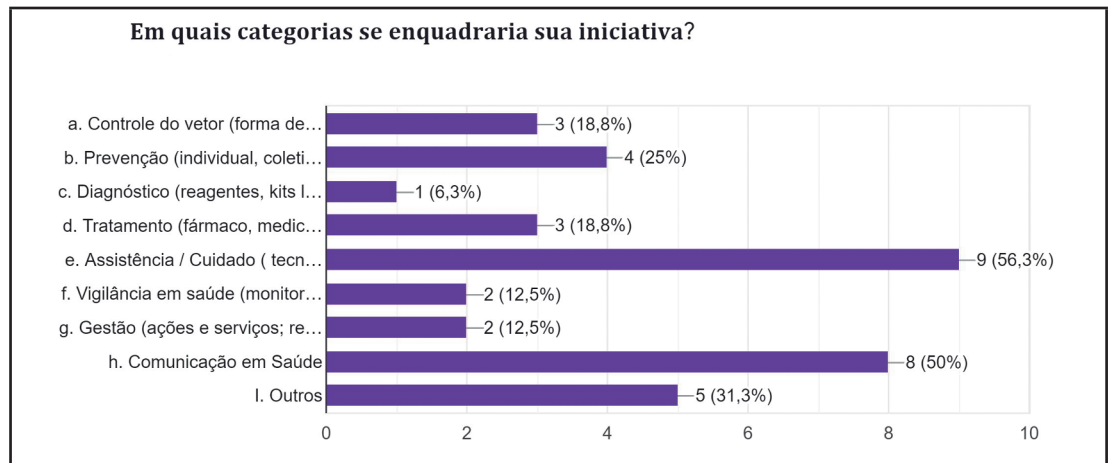
Figura 27: Distribuição das soluções sociais selecionadas segundo tema principal, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

Saúde (n = 6) e meio ambiente (n = 4) foram os temas principais das soluções sociais.

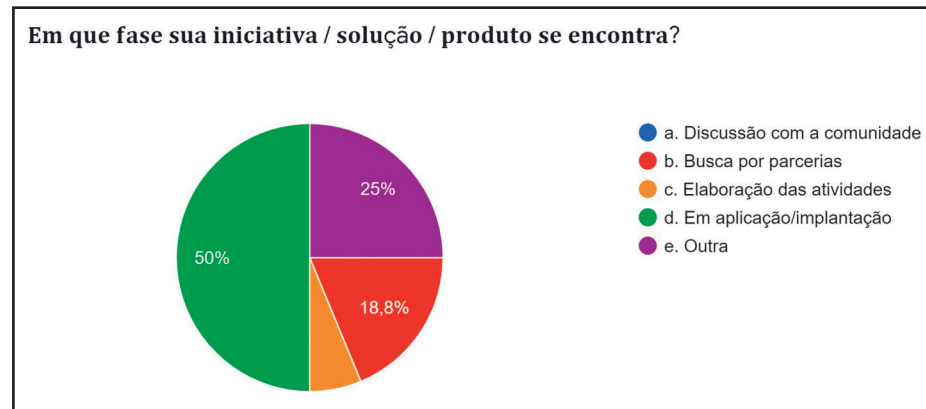
Figura 28: Distribuição das soluções sociais selecionadas segundo categorias, Brasil - 2021



a. Controle do vetor (forma de barreiras, controle, combate, tecnologia social); b. Prevenção (individual, coletiva, vacina, tecnologia social); c. Diagnóstico (reagentes, kits laboratoriais, imagem, tecnologia social); d. Tratamento (fármaco, medicamento, práticas integrativas e complementares, tecnologia social); e. Assistência / Cuidado ( tecnologia social, tecnologia assistiva, práticas integrativas e complementares); f. Vigilância em saúde (monitoramento, sistemas de informação, notificação...); g. Gestão (ações e serviços; regulação); h. Comunicação em Saúde; i. Outros Observação: era possível selecionar mais de uma categoria.

Fonte: Correia & Martins, 2021. No prelo

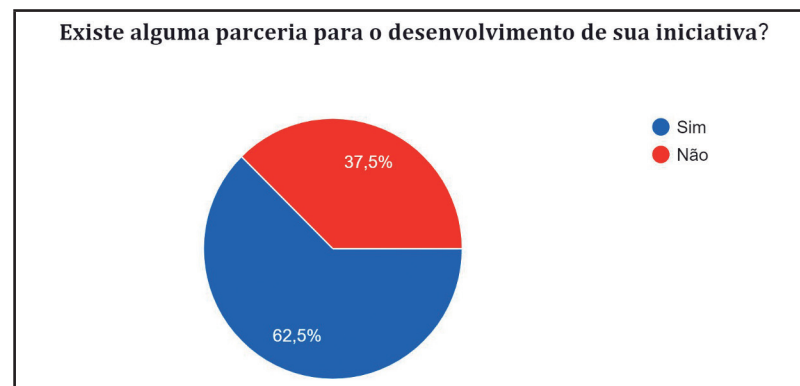
Figura 29: Distribuição das soluções sociais selecionadas segundo fases de desenvolvimento, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

Metade das soluções selecionadas mencionou que estavam em fase de implantação, ou aplicação (n = 8) e 25% referiram estar em outras fases (n = 4). Em relação a parcerias, 62,5% (n= 10) mencionaram ter parceiros para o desenvolvimento destas iniciativas que contribuem por outras formas, conhecimentos técnicos e divulgação.

Figura 30: Distribuição das soluções sociais selecionadas segundo parcerias, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

A mobilização com as comunidades e parceiros era realizada pela realização de eventos (62,5%), rodas de conversas (50%) e uso de mídias sociais locais (50%), e dentre as principais dificuldades, a falta de recursos financeiros foi apontada como a principal, com 75% (n = 12). Dentre as potencialidades mencionadas, o apoio institucional e o baixo custo foram os mais relevantes.

Figura 31: Feira de Soluções de Salvador/BA (Feira Zika, 2017)



Fonte: Fiocruz Brasília

Figura 32: Feira de Soluções de Fortaleza/CE, 2019



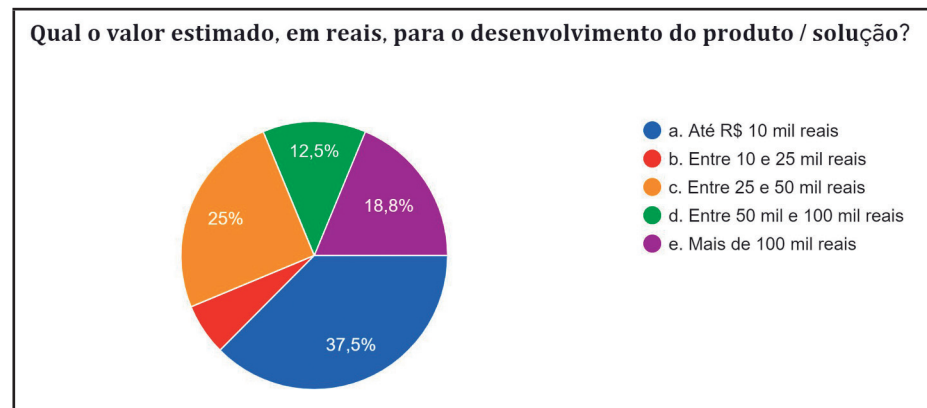
Fonte: Fiocruz Brasília

Figura 33: Distribuição das soluções sociais selecionadas segundo desafios e dificuldades no desenvolvimento dessas iniciativas, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

Figura 34: Distribuição das soluções sociais selecionadas segundo valor estimado para desenvolvimento do produto/solução, Brasil - 2021

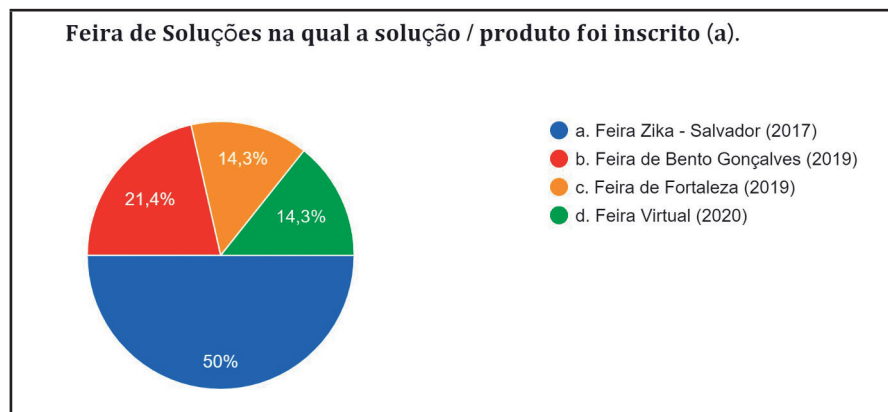


Fonte: Correia & Martins, 2021. No prelo

O custo estimado é inferior a R\$ 10 mil reais (Figura 34) para 37,5% (n = 6).

## Soluções de Serviços

Figura 35: Distribuição das soluções de serviços selecionadas por Feira de Soluções para a Saúde, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

A maioria das soluções de serviços (n = 14) que participou deste estudo estavam inscritas na Feira de Soluções para a Saúde de Salvador - Feira Zika.

Figura 36: Distribuição das soluções de serviços selecionadas por tipo de organização, Brasil - 2021

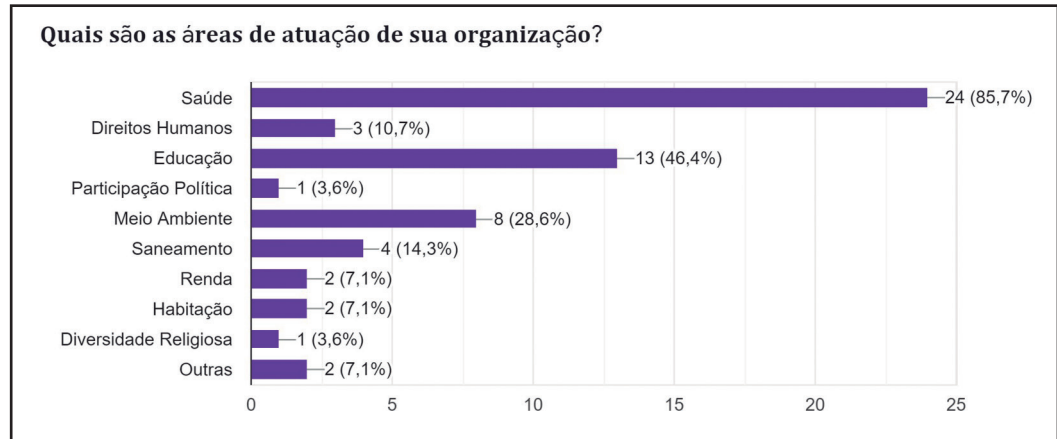


Fonte: Correia & Martins, 2021. No prelo

As instituições públicas ( $n = 7$ ), de ensino superior ( $n = 6$ ) e instituições de pesquisa ( $n = 6$ ) foram as que mais cadastraram soluções de serviços nas quatro feiras realizadas. Dentre os temas abordados pelas soluções de serviços, a maior parte abordou os temas de saúde ( $n = 24$ ) e educação ( $n = 13$ ).

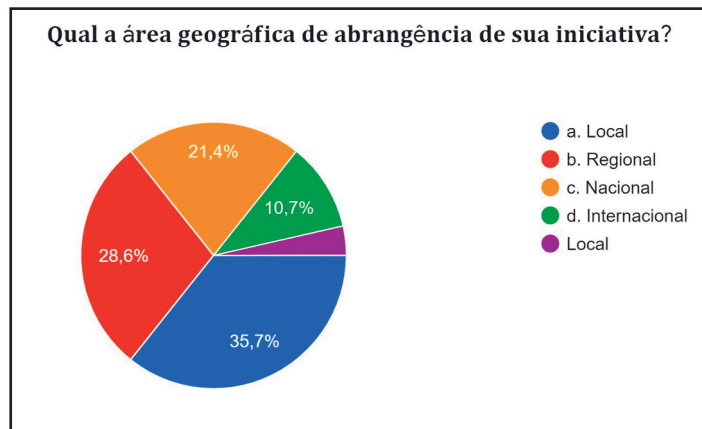


Figura 37: Distribuição das soluções de serviços selecionadas por área de atuação das organizações, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

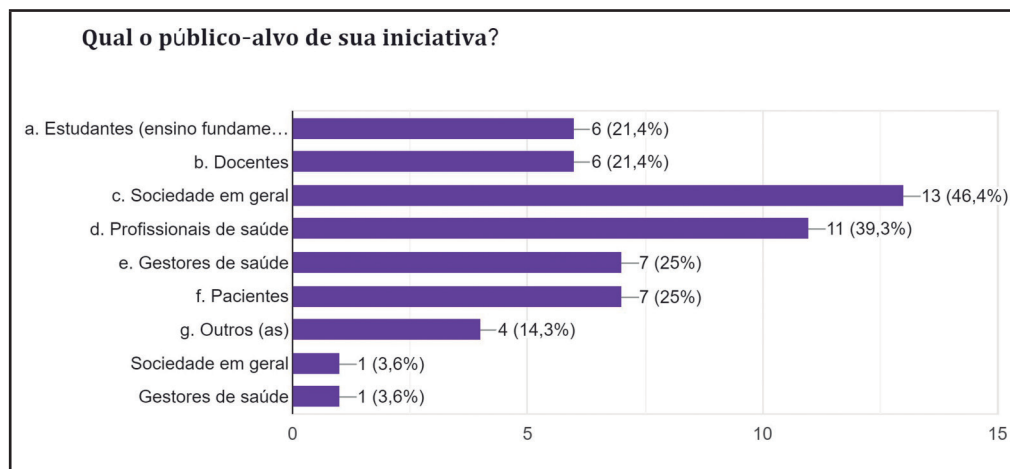
Figura 38: Distribuição das soluções de serviços selecionadas por área geográfica de abrangência, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

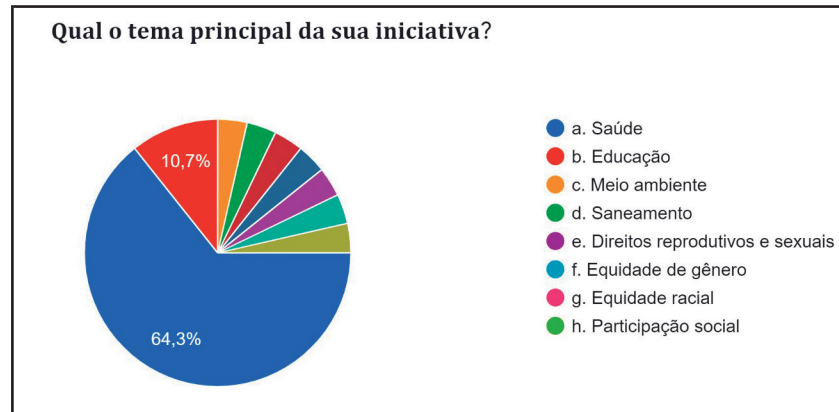
Dois terços das iniciativas de serviços tinham abrangência local (n = 10) e regional (n = 8). As soluções apresentadas tinham como público-alvo principal a sociedade em geral (n = 13) e profissionais da saúde (n = 11).

Figura 39: Distribuição das soluções de serviços selecionadas segundo público-alvo, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

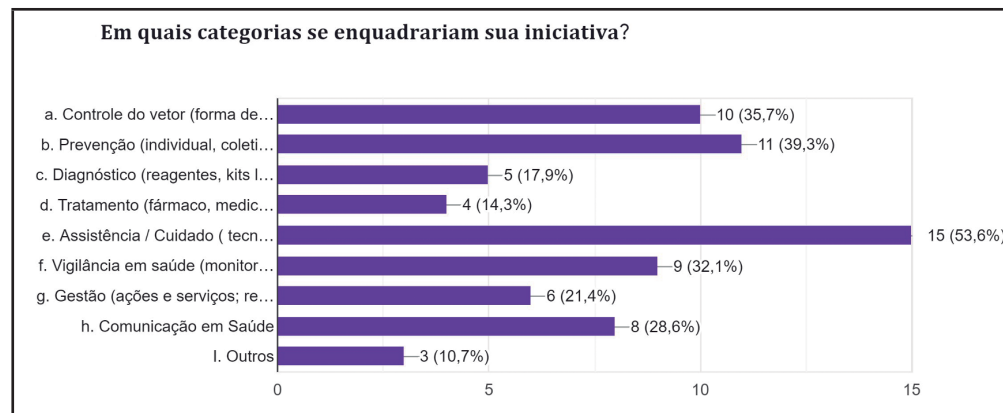
Figura 40: Distribuição das soluções de serviços selecionadas segundo tema principal, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

Um total de 64,3% (n = 18) das soluções de serviços estavam relacionadas de forma direta à temática da saúde.

Figura 41: Distribuição das soluções de serviços selecionadas segundo categorias de ação, Brasil - 2021



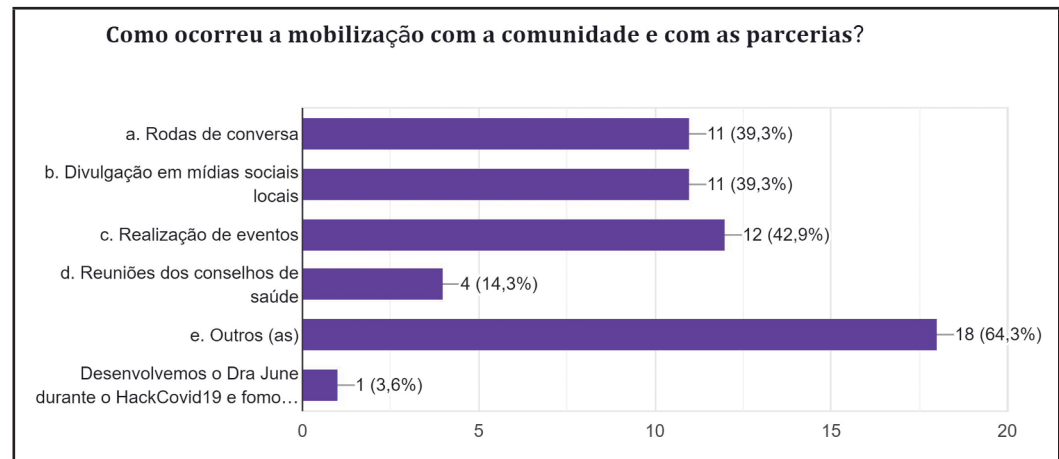
a. Controle do vetor (forma de barreiras, controle, combate, tecnologia social); b. Prevenção (individual, coletiva, vacina, tecnologia social); c. Diagnóstico (reagentes, kits laboratoriais, imagem, tecnologia social); d. Tratamento (fármaco, medicamento, práticas integrativas e complementares); e. Assistência / Cuidado (tecnologia social, tecnologia assistiva, práticas integrativas e complementares); f. Vigilância em saúde (monitoramento, sistemas de informação, notificação...); g. Gestão (ações e serviços; regulação); h. Comunicação em Saúde;

i. Outros. Observação: era possível selecionar mais de uma categoria

Fonte: Correia & Martins, 2021. No prelo

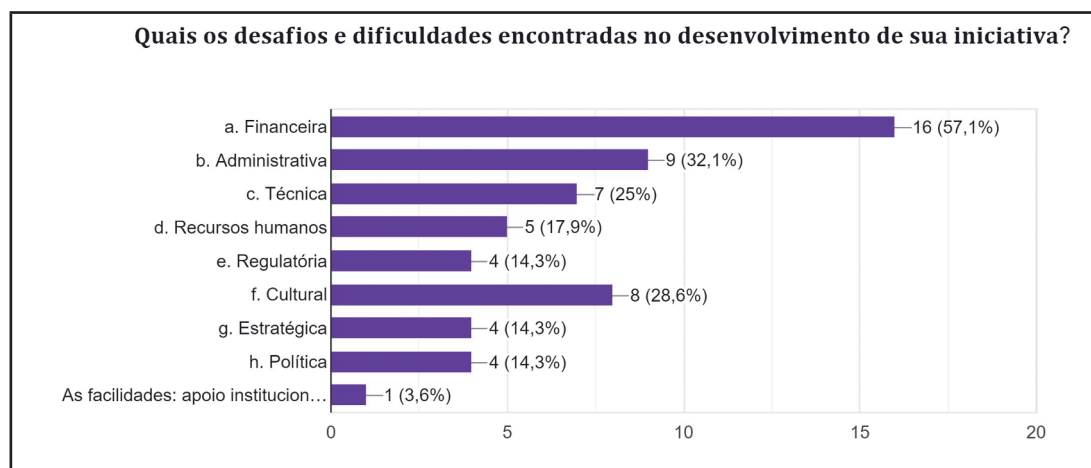
As iniciativas apresentadas em relação aos serviços distribuíram-se em várias temáticas: assistência / cuidado (n = 15), prevenção individual (n = 11) e controle do vetor (n = 10). Em relação às fases em que estas iniciativas se encontram, a maioria encontra-se em aplicação/implantação (n = 14).

Figura 42: Distribuição das soluções de serviços selecionadas segundo forma de mobilização com a comunidade e parcerias, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

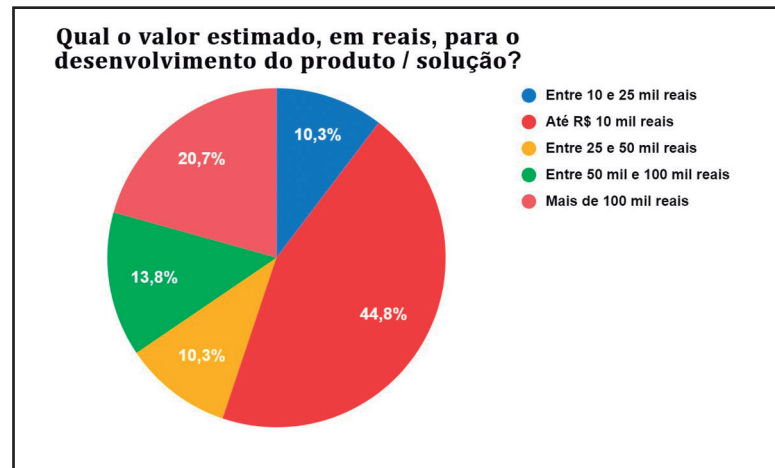
Figura 43: Distribuição das soluções de serviços selecionadas segundo desafios e dificuldades no desenvolvimento dessas iniciativas, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

Em relação aos desafios e dificuldades no desenvolvimento destas soluções/iniciativas, as mais presentes foram: financeira (57,1%); administrativa (32,1%) e culturais (28,6%). As principais facilidades e potencialidades foram: baixo custo (70,4%), apoio institucional (59,3%) e recursos humanos qualificados (51,9%).

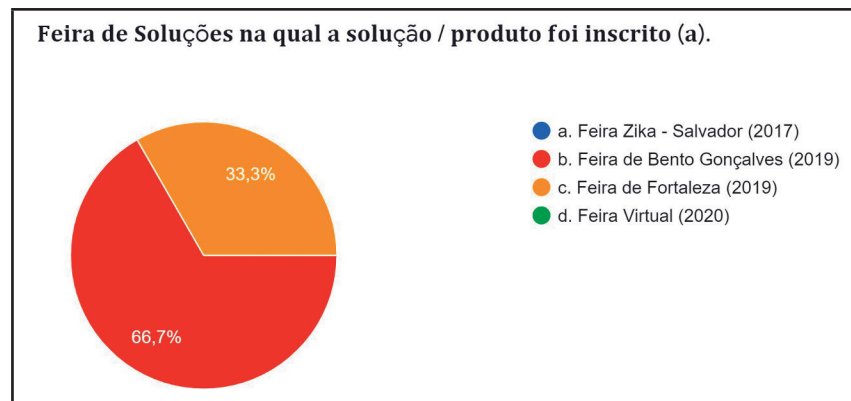
Figura 44: Distribuição das soluções de serviços selecionadas segundo valor estimado para desenvolvimento do produto / solução, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

O custo médio estimado para as Soluções de Serviço é de a R\$ 10 mil reais para 44,8% (n = 13).

Figura 45: Distribuição das soluções industriais selecionadas por Feira de Soluções para a Saúde, Brasil - 2021



Fonte: Correia & Martins, 2021. No prelo

As soluções industriais foram pouco representadas nesta análise porque somente três representantes aceitaram participar da pesquisa. Assim, 66,7% (n = 2) foram apresentadas na Feira de Bento Gonçalves (2019) e 33,3% (n = 1) na Feira de Fortaleza (2019, sendo vinculadas a instituição governamental (n = 1), empresa nacional (n = 1) e outros (n=1), situadas principalmente na região Sul (n = 2) e Nordeste (n = 1).

De acordo com os dados apresentados, as linhas de desenvolvimentos destas indústrias distribuem-se entre: diagnóstico, gestão, prevenção, tratamento e vigilância em saúde. De acordo com o estágio de desenvolvimento, duas estão em fase de finalização e uma está em fase inicial.

Até o momento de desenvolvimento deste estudo, nenhuma das soluções selecionadas possui registro em outro país, e as tecnologias utilizadas foram adquiridas por intermédio de parcerias com outras empresas/institutos, empresas públicas ou acordos comerciais. A origem do financiamento foi proveniente de recursos públicos federais, privados nacionais e do terceiro setor e os custos para ampliação das soluções industriais foram todos acima de R\$ 100 mil.

As plataformas tecnológicas envolvidas nas Soluções Industriais foram classificadas como biológico, biotecnologia ou equipamento, e segundo o valor agregado, foram consideradas soluções inovadoras para o mundo, com duas soluções já patenteadas e uma patenteável.

Como desafios para a execução e ampliação destas soluções, os desenvolvedores pontuaram: questões administrativas; estratégicas; financeiras e técnicas.

## Nível de Maturidade Tecnológica das Soluções Industriais

Para a análise das 23 Soluções Industriais apresentadas nas quatro feiras foi realizado um estudo adicional de Nível de Maturidade Tecnológica (NMT).

De acordo com critérios do Quadro 1, as soluções foram enquadradas e a maioria alcançou o maior nível de maturidade, NMT 9 (52%), tendo em vista que os produtos já estão sendo comercializados, ou possuem registro na Agência Nacional de Vigilância Sanitária (Anvisa). Uma solução (4%) foi classificada com NMT 8, por ter sido implantada apenas em um setor específico. Outra solução (4%) foi categorizada com NMT 7 por ter somente o protótipo lançado. O restante das soluções (56%) apresenta um NMT abaixo de 6, sendo cinco classificadas com NMT 4 (encontram-se em estágio de pesquisa e desenvolvimento), uma com NMT 2 e três com NMT 1 (não possuem informações adicionais que permitem enquadrá-las em níveis superiores).

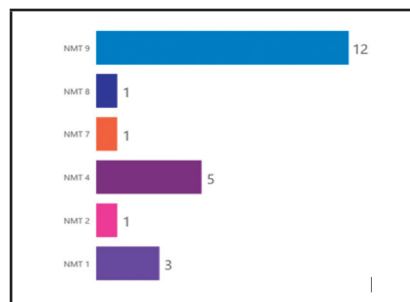
Quadro 1: Níveis de Maturidade Tecnológica (NMT)  
das soluções industriais para a saúde

NMT 1	NMT 2	NMT 3	NMT 4	NMT 5	NMT 6	NMT 7	NMT 8	NMT 9
Pesquisa básica		Pesquisa e desenvolvimento			Desenvolvimento do sistema/subsistema (estudos clínicos)			
Pesquisa básica		Pesquisa pré-clínica	Pesquisa pré-clínica tardia		Fase 1	Fase 2	Fase 3	Fase 4, Registro
Pesquisa de viabilidade		Demonstração da tecnologia			Teste, lançamento e operação do sistema			
Observação e registro da ideia básica da solução	Formulação dos conceitos tecnológicos ou da aplicação da solução	Prova de conceito da solução	Validação laboratorial dos componentes da solução	Validação das funções críticas dos componentes da solução em ambiente de produção (validação de processo)	Demonstração das funções críticas do modelo de sistema/subsistema ou protótipo da solução em ambiente de produção (validação da capacidade de processo)	Demonstração do protótipo do sistema da solução em ambiente operacional (validação da capacidade de processo em lote econômico)	Finalização e qualificação do sistema atual da solução por testes e demonstração (validação da capacidade sobre uma parcela das partes)	Aprovação e implantação da solução no ambiente de aplicação proposto (validação da capacidade sobre todas as partes e por longos períodos). Ex.: Produto com registro nos órgãos sanitários ou comercializado.

Fonte: Rezende & Martins, 2021, adaptado de VELHO *et. al.* (2017)



Figura 46: Soluções industriais por Nível de Maturidade Tecnológica (NMT)



Fonte: Silva & Martins, 2021. No Prelo

## PRODUTOS

### Plataforma Ágora/Fiocruz

As Feiras de Soluções para a Saúde em suas quatro edições, propiciaram a construção de um banco de dados com 281 Soluções para a Saúde, onde a cooperação entre diferentes atores possibilita a formação de redes virtuais, integrando diferentes níveis de desenvolvimento da ciência e maturidade tecnológica para a solução de problemas de saúde<sup>7</sup>.

### Soluções para a saúde em fase de implementação

Das 281 Soluções para a saúde cadastradas na Plataforma Ágora, 37 delas encontravam-se em fase de implantação no momento de suas inscrição nas Feiras de Soluções.

7. Banco de Soluções disponível em <https://solucoes.agora.fiocruz.br>

## Soluções Sociais

**Programa Mães Produtivas AMAR** (Feira de Soluções, Salvador - 2017). É um programa que visa gerar renda e empoderamento social a mulheres cuidadoras de pessoas com doenças raras no país. Atua em módulos, em parcerias com instituições afins à causa, inclusive microcefalia. As beneficiárias são selecionadas para preencher uma modalidade à sua escolha, desde graduação na modalidade do Ensino à Distância (EAD), até cursos como artesanato e culinária, atividades que elas podem conciliar com o cuidado das crianças e, em paralelo, atuar como promotoras de qualidade de vida das famílias.

**AMAR Cuidando de quem cuida** (Feira de Soluções, Salvador - 2017). O Programa é o desenvolvimento da missão institucional da AMAR. Com a percepção de marginalização das cuidadoras de pessoas com doenças raras, em especial, da microcefalia oriunda da Síndrome Congênita do Zika Vírus, percebeu-se que as políticas, atenção e cuidados estavam voltados unicamente para a pessoa doente, como se a mãe não fosse também sujeito de direitos. A AMAR desenvolve as ações de empoderamento das cuidadoras, acolhimento e orientação jurídica, psicológica e encaminhamento terapêutico para parceiros. “Cuidando de Quem Cuida” é o programa de entrada das famílias assistidas pela AMAR na Organização Não Governamental. O objetivo principal é a assistência social e psicológica e a identificação de necessidades destas famílias, a partir do que se realiza um plano de ação, e cada família é assistida conforme sua especificidade. Em comum a todas as famílias, há o módulo de Participação Cidadã, onde as cuidadoras com laços afetivos e suas famílias são estimuladas a interagir entre si, a conhecer seus direitos e a participar de movimentos sociais aderentes à AMAR.

**Exposição Fotográfica aBRAÇO através do olhar** (Feira de Soluções, Salvador - 2017). O nome da exposição foi inspirado no seu próprio objetivo: propor um novo olhar sobre a infância das crianças com microcefalia. A exposição é composta de aproximadamente 25 fotos, tamanho 0,60 m x 0,90 m e de autoria da fotógrafa Natália Borges. As fotos são impressas em preto e branco, ou coloridas, e expostas em painéis. A

exposição funciona em conjunto com a Loja aBRAÇO itinerante para potencializar a divulgação da associação.

**Abraço:** Centro de Referência em Acolhimento, Reabilitação e Fortalecimento da Rede Social (Feira de Soluções, Salvador - 2017). A Associação Abraço à Microcefalia apoia as famílias de crianças com alterações neurológicas e microcefalia em uma rede de apoio e solidariedade, proporcionando oficinas, com diversos profissionais, no intuito de orientar famílias e criar um ambiente de convivência, onde haja troca de experiências e momentos de relaxamento. Além dos atendimentos individualizados com fonoaudiólogos, terapeutas ocupacionais, psicólogos e fisioterapeutas, as famílias são acolhidas, dando-lhes suporte e orientação.

**Livro infantil Convivendo com as Diferenças** (Feira de Soluções, Salvador - 2017). A solução visa disseminar informações sobre as características das crianças com a Síndrome Congênita do Zika Vírus (SCZV), por meio de um livro infantil que aborda o tema da microcefalia de maneira lúdica e sensível. A solução destina-se a crianças de 2 a 9 anos, suas famílias, além de educadores e profissionais de escolas públicas e privadas. A expectativa é distribuir 1000 exemplares e que o livro contribua para inclusão no ambiente escolar, com destaque para a criança com microcefalia.

**Aplicativo AMMA:** AMor MAior (Feira de Soluções, Salvador - 2017). O aplicativo assim denominado, será de distribuição gratuita e conterá informações quanto à estimulação precoce dos bebês, através de atividades do cotidiano e diversão, no formato de vídeo e imagens. No aplicativo, as famílias poderão aprender a estimular seus bebês de forma simples e fácil, visando sempre ajudá-los o quanto antes, para quem sabe, incluí-los no mercado de trabalho, ou até mesmo, alcançar independência das atividades domésticas.

**Materiais de baixo custo para apoio à estimulação precoce de crianças com a Síndrome do Zika Vírus** (Feira de Soluções, Salvador - 2017). Desenvolvimento de metodologia de construção de materiais de estimulação sensorial, utilizados por famílias e profissionais da área de saúde, com o objetivo de realizar atividades de estimulação

precoce. A produção dos materiais é de simples execução, com qualidade estética, segurança e impactos testados junto a parceiros locais. Os materiais (brinquedos) foram desenvolvidos por uma equipe multidisciplinar (fonoaudióloga, especialista em brinquedos e *designers*) e criados a partir de caixas, bacias e tecidos. Podem ser utilizadas por famílias, cuidadores e profissionais de atendimento com adaptações e intervenções de simples execução, com qualidade estética, segurança e validados no ambulatório de estimulação precoce do Instituto Fernandes Figueira (IFF).

**“Infobike - pedalando por “mais direitos, menos Zika”** (Feira de Soluções, Salvador - 2017). O “Infobike” foi um projeto pensado por jovens para jovens e teve como objetivo levar informações e insumos de prevenção (preservativos) a áreas de comunidades periféricas em bairros de Salvador e Recife. A partir de uma bicicleta adaptada com equipamento de som, dialogam com jovens homens e mulheres sobre direitos sexuais e direitos reprodutivos em tempos de Zika (pensando também dimensões como raça, geração, local de moradia, orientação sexual e identidade de gênero). A solução apresentada é uma bicicleta personalizada a partir da identidade visual do Coletivo Mangueiras, que com uma caixa de som acoplada e saída para microfone, atraía a atenção nas comunidades com músicas e *spots* de áudio informativos. Também foi acoplado à bicicleta um baú de madeira, que permitia transportar insumos de prevenção e materiais informativos a serem distribuídos ao longo do trajeto. A bicicleta itinerante percorreu ruas e becos de comunidades de Recife e Salvador, convidando as pessoas para conversar sobre direitos sexuais e direitos reprodutivos e o contexto do Zika. Optou-se pelo uso da bicicleta por tratar-se de uma ferramenta comunicacional comum ao cotidiano das comunidades, o que possibilitou uma boa aproximação com as mulheres e jovens que encontravam. Os percursos desenvolvidos consideraram os pontos de maior vulnerabilidade das regiões, além de pontos de encontro de jovens, como praças e escolas. Anteriormente ao momento de incidência com a Infobike, as/os jovens articuladoras/es envolvidas/os nas ações participaram de espaços de formação teórico-política desenvolvidos pelo Coletivo Mangueiras e também por organizações parceiras. Vale a pena conferir o *hyperlink* do coletivo Mangueiras: <https://bit.ly/3B7EukD>.

**Implantação de uma comunidade intencional sustentável em uma área rural, visando o envelhecimento com autonomia** (Feira de Soluções, Bento Gonçalves - 2019).

A configuração física das comunidades intencionais se apresenta como alternativa viável para o envelhecimento visando a recuperação e responsabilidade dos cuidados primários, estimulando a interação e o desenvolvimento de um senso de comunidade sustentável. Objetivo: Implantar uma comunidade intencional sustentável em uma área rural, visando o envelhecimento com autonomia. A Implantação usa o trabalho colaborativo na produção, comercialização e consumo de produtos como mel, hortifrutigranjeiros, plantas nativas, frutíferas, ornamentais, espaço coletivo para práticas integrativas complementares para a comunidade interna e externa, impactando positivamente na vida dos beneficiários. O produto é uma hospedagem rural Holztel, que propicia a sustentabilidade financeira de alguns projetos que estão sendo implementados na comunidade intencional estabelecida no entorno do empreendimento. Todas as instalações são de livre acesso aos hóspedes, bem como, a liberdade de poder usufruir do pomar e da horta, sem custo. São produzidos apenas produtos orgânicos, sem uso de fertilizantes e defensivos agrícolas (Site: <https://www.holztel.com.br/>).

**Formação para catadores de resíduos sólidos na área de saúde e ambiente** (Feira de Soluções, Ceará - 2019).

Os trabalhadores que fazem a triagem desse material em dezenas de cooperativas e associações e participam da coleta seletiva de seus municípios, não recebem apoio técnico, de formação e acompanhamento. A Apoena Socioambiental realiza formação para estes trabalhadores com temas essenciais para sua atuação e relação ambiente-saúde. Os temas são escolhidos pelos catadores e catadoras, por exemplo: uso de equipamentos de proteção individual (EPI), vacinação, saúde do trabalhador, alimentação, violência doméstica, gestão financeira (pessoal e da cooperativa), protagonismo feminino, gestão de resíduos e educação ambiental (para os trabalhadores e comunidade no entorno). O objetivo desta solução é trazer conhecimento de uma maneira acessível e didática, levando em consideração o ambiente em que eles trabalham e a sua atuação. Cada oficina é realizada em um espaço de apoio (universidades, escolas, centros de educação), ou em seus próprios espaços de trabalho.

**Viveiro Educador** (Feira de Soluções, Ceará - 2019). Tecnologia social desenvolvida a partir do plantio de propágulos de mangue vermelho em espaço próximo a árvores adultas da mesma espécie, possibilitando o aviamento pelo próprio, quando da chegada da maré e a proteção das novas sementes pelas árvores já existentes.

**Ação Sustentável: Papel Semente** (Feira de Soluções, Ceará - 2019). Por meio da observação da rotina da escola, constatou-se um alto índice de descarte de papel nas salas de aula e nas demais dependências. Constatou-se que, em um período de vinte dias letivos, o descarte de papel chega a aproximadamente 50 Kg, e que esse volume de papel produzido constituiria um grande desperdício de recursos financeiros e, sem a reciclagem, um grande impacto ambiental, pois a produção de papel consome grande quantidade de recursos naturais. Da constatação do problema, discutiu-se uma forma de resolvê-lo. Encontrou-se uma maneira de reciclar o papel e transformá-lo em suporte de sementes, que se plantado germinaria novas plantas. Objetivo geral: Reciclar papel transformando-o em suporte de sementes, para dar destino sustentável ao papel descartado na Escola Eduardo Alves Ramos, Eusébio-CE. Objetivos específicos: estimular o uso do papel reciclado; projetar o uso de papel reciclado para cartões feitos pela escola; evitar a poluição do meio ambiente; estimular a comunidade escolar a produzir suas hortas caseiras a partir do plantio do papel semente. Relevância do projeto: Dentro do conceito de sustentabilidade, a reciclagem do papel também é essencial para reduzir sua quantidade, além de favorecer o consumo consciente de recursos naturais, já que se gasta mais água e energia elétrica para a produção do papel comum do que para o reciclado. Quanto ao papel semente, sua importância se justifica pelo fato de que após o seu uso, ele poderá ser plantado, evitando a poluição e gerando novas plantas, tornando-se dessa forma, um meio eficaz disponível para contribuir com a diminuição dos impactos da ação humana no meio ambiente.

**Saneamento Ecológico:** implantação de tecnologias sociais para reúso das águas cinzas (Feira de Soluções, Ceará - 2019). Implantação de tecnologias sociais acessíveis e de baixo custo voltadas ao saneamento ecológico no Assentamento Palestina / Oiticuinha, localizado no município de Independência-CE, considerando o contexto social, econômico, ambiental e cultural, pois a falta e dificuldades do acesso às polí-

ticas públicas, principalmente referentes ao saneamento básico no campo, causam agravos à saúde e ao ambiente. Nesse contexto, o projeto trabalhou ações de formação e aplicação de tecnologias de tratamento de água cinza, sendo elas o Círculo de Bananeira e Espinhel de Peixe para reutilização em atividades que as famílias desenvolvem nos quintais de suas residências, como o cultivo de frutíferas e hortaliças com base na agroecologia. O projeto coloca-se como uma alternativa a soluções de uma problemática encontrada no assentamento, e também é uma referência para outras comunidades, para ser fomentada junto às pautas de luta por políticas públicas que melhorem a vida no campo.

**Gincana da Saúde: Desafio Melhor de 3** (Feira de Soluções Virtual - 2020). Uma forma lúdica de conversar com crianças e adolescentes dos ensinos fundamental e médio sobre o conceito ampliado de saúde e sobre *fake news*, especialmente sobre a Covid-19 e vacinas. São três jogos que utilizam ferramentas digitais *on-line* gratuitas para divulgar conhecimento científico.

### Soluções de Serviço

**Desenvolvimento de um sistema de baixo custo para contagem automática de ovos de *Aedes aegypti*** (Feira de Soluções, Salvador - 2017). O monitoramento entomológico é uma maneira de melhorar a capacidade de predição e na detecção precoce de epidemias das doenças transmitidas pelo mosquito *Aedes aegypti*. Este monitoramento é majoritariamente baseado no índice larvário, o qual lista a quantidade de casas infectadas, ou a quantidade de ovos de *Aedes* coletados em palhetas inseridas em ovitrampas. A solução consiste na criação de um aplicativo de celular e um protocolo para aquisição de imagens para contagem dos ovos de *Aedes aegypti* com as seguintes características: facilidade de uso, alta acurácia e custo baixo. Destina-se a pesquisadores e laboratórios com alta demanda de contagem dos ovos em palhetas de algodão, ou Eucatex.

**Kit Pedagógico: Missão Zika Zero** (Feira de Soluções, Salvador - 2017). O kit foi produzido pela Pró-Reitoria de Extensão, proposto pelo Grupo de Trabalho Zika Zero da Universidade Federal do Recôncavo da Bahia (UFRB). É destinado ao público infantil, com o objetivo de informar de maneira lúdica sobre a proliferação do mosquito *Aedes Aegypti* e suas consequências. Trata-se de um conjunto composto por revista em quadrinhos, atividades lúdicas, cartazes, folders, adesivos, dentre outros, disponíveis virtualmente no site da UFRB e impressos, distribuídos em atividades educativas.

**Curso on-line “Zika: abordagem clínica na Atenção Básica”** (Feira de Soluções, Salvador - 2017). O curso *on-line* “Zika: abordagem clínica na Atenção Básica” com 45 horas, objetiva capacitar os profissionais de saúde para minimizar os impactos da Zika no país. Para tanto, as quatro unidades do curso tratam dos aspectos epidemiológicos, fisiopatológicos, promoção à saúde e prevenção de infecção pelo vírus Zika. Há uma unidade dedicada aos cuidados às mulheres (em especial às gestantes e parturientes), que trata do planejamento reprodutivo e do diagnóstico de microcefalia no recém-nascido. A última unidade trata de temas como notificação compulsória; instrumentos específicos de notificação de suspeita e confirmação; vigilância de complicações decorrentes da infecção pelo vírus Zika e atuação do profissional-sentinela na identificação de complicações.

**Armadilha para mosquitos feita com garrafas PET** (Feira de Soluções, Salvador - 2017). “O objetivo é fazer uma armadilha simples e barata, que pode ser confeccionada em casa com materiais muito comuns: garrafas PET e tela de mosquiteiro. O mosquito vê a água parada dentro da garrafa, deposita seus ovos que, posteriormente, não vão poder sair devido à tela de impedimento, pois ficam presos. É aconselhável espalhar alguns pela casa para uma maior eficiência.

**Aedes Play: Jogo de tabuleiro como proposta educativa em prevenção e controle ao Zika vírus** (Feira de Soluções, Salvador - 2017). O *Aedes Play* é um jogo de tabuleiro desenvolvido como ferramenta e estratégia educacional nas escolas para difundir aspectos relacionados ao vetor (*Aedes aegypti*) e às doenças transmitidas, com ênfase no Zika vírus e as maneiras de prevenção. A sua execução obteve um resultado



satisfatório na ampliação do conhecimento dos alunos participantes, e constituiu uma importante forma de divulgação de medidas de combate a esta arbovirose e outras.

**Selo Sérgio Arouca de Qualidade da Informação: Certificação de *Sites* de Saúde** (Feira de Soluções, Salvador - 2017). Esta solução tem o objetivo de avaliar e certificar *sites* de Zika, Dengue e Chikungunya e Síndromes Congênitas com o Selo Sérgio Arouca da Escola Nacional de Saúde Pública da Fiocruz. O público é formado por gestores públicos e organizações não governamentais (ONGs) responsáveis por *sites* de saúde. Resultados: A informação de saúde certificada pode empoderar o cidadão, melhorar sua qualidade de vida, facilitar o autocuidado e reduzir os gastos com serviços de saúde.

**Kit Multissensorial: Multiplicando e disseminando esta iniciativa** (Feira de Soluções Salvador - 2017). Trata-se da apresentação de um *Kit* Multissensorial que foi reproduzido tomando como base uma das ações do Projeto Redes de Inclusão, iniciativa do Fundo das Nações Unidas para a Infância (Unicef) com outros parceiros, entre eles a Fundação Altino Ventura. O objetivo deste trabalho é multiplicar e disseminar tal ideia, ressaltando sua utilização não apenas na Síndrome Congênita do Zika Vírus, mas também em qualquer outra condição que a criança apresente transtorno do seu desenvolvimento. 1ª Atividade - Reprodução de *kit* multissensorial (0-3 anos) para crianças com microcefalias ou outras deficiências; 2ª Atividade - Distribuição de calças de posicionamento; 3ª Atividade - Confeção de mini *kits* para estimulação de crianças com microcefalia, ou outras deficiências (0 a 3 meses); 4ª Atividade - Aplicação de protocolo de acompanhamento do desenvolvimento para crianças com transtornos do desenvolvimento (0 a 3 anos).

**Eliminar a Dengue - Desafio Brasil** (Feira de Soluções, Salvador - 2017). Arboviroses transmitidas pelo *Aedes aegypti* geram efeitos para a saúde e economia de países tropicais e subtropicais ao redor do globo. A descoberta do efeito protetor da *Wolbachia* em *Drosophila* sp. (mosca da fruta) para DCV (vírus C de *Drosophila* sp.) suscitou a ideia de transferir mecanicamente a *Wolbachia* desta espécie de mosca para o *Aedes aegypti*. Os efeitos foram semelhantes, garantindo uma redução da replicação viral no

organismo do vetor. O antes denominado “Projeto Eliminar a Dengue - Desafio Brasil”, hoje tornou-se um programa internacional sem fins lucrativos, desenvolvido em 11 países. Tem como base a utilização da bactéria *Wolbachia* no controle das arboviroses transmitidas pelo mosquito *Aedes aegypti*. A presença da *Wolbachia* nas células do vetor reduz sua capacidade em transmitir vírus como Dengue, Zika e Chikungunya.

**Estação Disseminadora de Larvicida** (Feira de Soluções, Salvador - 2017). O princípio é o uso dos próprios mosquitos para disseminar larvicida. Utilizam-se baldes plásticos com água para atrair as fêmeas dos mosquitos. As paredes internas dos baldes são recobertas com um pano preto sobre o qual se aplica larvicida em forma de pó. Quando um mosquito adulto pousa nessa superfície, partículas do inseticida se aderem às pernas e ao corpo do inseto. Como as fêmeas de *Aedes* (e outros mosquitos) visitam muitos criadouros para colocar uns poucos ovos em cada um, elas disseminam o larvicida.

**Curso Prático para o acompanhamento escolar de crianças acometidas pela Síndrome Congênita da Zika** (Feira de Soluções, Salvador - 2017). Formação prática, de natureza gratuita, ministrada por profissionais da atenção especializada nas temáticas de reabilitação física, intelectual, auditiva e visual para profissionais do contexto escolar. Ofertadas duas turmas com carga horária total de 28 horas, distribuídas em encontros mensais de 4 horas/aula por módulo, totalizando sete módulos que abordaram temas relacionados ao cuidado interdisciplinar da criança com deficiência múltipla, tais como AVD (Atividades da Vida Diária), alimentação, postura, mobiliário e recursos adaptados, higiene e recursos pedagógicos adaptados, tendo como base o conceito e os recursos da Tecnologia Assistiva para a melhor qualidade no trabalho com estas crianças. Público-alvo prioritário: 40 profissionais (sendo 20 professores e 20 cuidadores) que atuam em creches e/ou escolas no município de Salvador. Foram convidadas 21 escolas e/ou creches da Rede Municipal de Ensino de Salvador-BA, pelo critério de estarem localizadas em Gerências Regionais que contemplavam os bairros de residência das crianças diagnosticadas com a Síndrome Congênita da Zika, atendidas no Centro Especializado em Reabilitação das Obras Sociais Irmã Dulce (CER IV/OSID). Algumas vagas também foram disponibilizadas para profissionais de educação de outros muni-

cípios da Bahia. Assim foram inscritas 42 pessoas nas duas turmas; com média de 32 participantes por módulo, e especificamente na cidade de Salvador/BA, contemplou 13 escolas e/ou creches da Rede Municipal de Ensino. O curso foi desenvolvido nas dependências do CER IV/OSID, pelos profissionais do centro, das seguintes especialidades: Educação Física (2), Fisioterapia (1), Fonoaudiologia (4), Odontologia (1), Pedagogia (4), Psicologia (2), Serviço Social (2) e Terapia Ocupacional (5).

**Programa Cada Jovem Conta** (Feira de Soluções, Bento Gonçalves - 2019). O Programa Cada Jovem Conta organiza comitês intersetoriais nos territórios para identificação e acompanhamento de crianças e adolescentes em situação de risco para a violência, construindo um plano único de ação preventiva. Tem como objetivos: estreitar vínculos das famílias com a rede de atenção básica, educação e assistência social; reduzir fatores de risco e ampliar fatores de proteção, impactando positivamente na frequência e rendimento escolar.

**Programa Mais Saúde: Exercite essa ideia** (Feira de Soluções, Bento Gonçalves - 2019). Em nossa sociedade, ser velho, na maioria das vezes, significa ser relegado a uma vida solitária, excluído de atividades sociais e sem prazeres. Objetiva-se neste projeto transformar esta realidade no município de Lavras do Sul, introduzindo através da atividade física e do estímulo ao fortalecimento de vínculos, a promoção da saúde e melhora na qualidade de vida da pessoa idosa. Os idosos participam de atividades semanais como: dança, caminhada, alongamento, técnicas psicológicas, aferição de pressão arterial, palestras informativas com profissionais especializados, trabalho voltado para prevenção de acidentes e promoção de saúde. Os encontros são organizados pela fisioterapeuta e pela psicóloga do município, além de trabalhar na prevenção para redução das incapacidades e potencializar o envelhecimento ativo. Visam estimular a convivência e troca de experiências, para uma maior valorização pessoal e melhoria da autoestima, reforçando a visão da pessoa idosa enquanto cidadão ativo na comunidade.

**Farmácia Verde no Sistema Prisional** (Feira de Soluções, Bento Gonçalves - 2019). No sistema prisional são oferecidas oficinas com a finalidade de reintegrar e recuperar os

apenados promovendo a reinserção social. Objetivou-se construir a farmácia verde no presídio estadual local, desde o plantio, manejo e multiplicação de mudas, até a distribuição nas Unidades Básicas de Saúde (UBS), após as capacitações com os grupos de regime fechado e o regime semiaberto do presídio, abordando o manejo correto de plantas medicinais, secagem e empacotamento.

**Instrumento *on-line* de monitoramento e avaliação intersetorial do programa Bolsa Família** (Feira de Soluções, Ceará - 2019). Implantação de um instrumento *on-line* de monitoramento e avaliação intersetorial entre a saúde e assistência social para fortalecer as estratégias, promover vínculos e criar fluxo de acompanhamento, qualificando o serviço ofertado aos beneficiários do programa e ampliando o acesso às variáveis vinculadas às condicionalidades da saúde.

**Utilização de pessário vaginal para tratamento conservador de prolapso de órgãos pélvicos no âmbito do Sistema Único de Saúde** (Feira de Soluções, Ceará - 2019). Os Prolapsos de Órgãos Pélvicos (POP) são condições ginecológicas comuns que acometem mulheres em idades variadas. Caracterizam-se pela descida das paredes vaginais, cúpula vaginal e/ou útero em direção às carúnculas himenais, causando uma sensação de peso vaginal ou “bola na vagina”. Essa condição impacta negativamente na qualidade de vida. Os pessários são instrumentos de silicone inseridos na vagina para dar suporte aos órgãos pélvicos, sendo indicados nos casos em que há contraindicação à cirurgia, quando as pacientes não desejam operar, ou ainda como tratamento temporário enquanto se espera pela cirurgia. Além de ser um tratamento não invasivo, são de baixo custo e com poucos efeitos colaterais.

**Fossas Verdes - ou Bacias de Evapotranspiração** (Feira de Soluções, Ceará - 2019). O Módulo de Fossa Verde (MFV) é uma tecnologia social para tratar esgoto doméstico unifamiliar, promovendo espaço para quintal produtivo, evitando assim a contaminação do solo e corpos hídricos. O sistema corresponde a uma vala impermeabilizada com uma câmara na porção interior, onde o esgoto é tratado por biorremediação. A água e os nutrientes presentes no esgoto são absorvidos, por fluxo ascendente, pelas raízes das plantas cultivadas na parte superior do canteiro. Essa tecnologia social tem

se apresentado viável como uma alternativa para o saneamento rural e periurbano na região semiárida brasileira. Mais detalhes podem ser acessados no artigo disponível no link: <https://bityli.com/MpkLRDe>

**UBS *On-line*** (Feira de Soluções virtual - 2020). A UBS *On-line* é um serviço de teleatendimento criado pelo Departamento de Medicina Social da Universidade Federal de Pelotas (UFPel), para que usuários das UBS pudessem resolver suas demandas por teleatendimento sem precisar deslocar-se até a unidade e poder seguir a recomendação de ficar em casa durante a pandemia. O recurso foi um celular por UBS, com linha telefônica e acesso à internet. Os profissionais foram indicados entre aqueles afastados da UBS, por serem parte do grupo de risco para Covid-19. Na experiência vivenciada até o momento, observou-se que praticamente todas as demandas do escopo da Atenção Primária em Saúde acabam sendo apresentadas. Mesmo aquelas que apresentam manifestações clínicas passíveis de exame, tem-se o recurso da imagem e vídeo que conseguem ainda resolver boa parte dos casos. O serviço permite também o acompanhamento dos casos suspeitos domiciliares de coronavírus que, neste momento, em particular, são relevantes. O único recurso material imprescindível para o início do trabalho foi um celular para cada UBS, já adquiridos pela UFPel. Os recursos humanos são os médicos lotados em cada UBS que, no período da Pandemia, cumprirão sua jornada de trabalho em revezamentos quinzenais, de modo que sempre haja algum profissional médico no seu domicílio realizando atendimentos dos pacientes da sua área pelo *Whats app*. 1. Garantindo a longitudinalidade do cuidado na APS durante a pandemia; 2. Fortalecendo o vínculo entre equipe e usuários; 3. Abrindo uma nova forma de acesso a APS durante a pandemia.

### Soluções industriais

**Block Insetos** (Feira de Soluções, Salvador - 2017). *Block Insetos* é a única tela mosquiteira fixada com imã em qualquer tipo de janela. Suas hastes de PVC junto às exclusivas fitas magnéticas possibilitam que a tela seja fixada à superfícies de alumínio,

madeira, PVC ou ferro. Basta fita métrica e tesoura para adaptar a tela ao tamanho perfeito da janela. A instalação é prática, sem furo e necessidade de técnico especializado, dentro/fora de casa. Veda integralmente a janela, bloqueando a entrada de poeira e insetos, permitindo apenas a entrada do ar fresco.

**Teste rápido Zika IgG / IgM Combo** (Feira de Soluções, Salvador - 2017). Primeiro teste sorológico brasileiro para detecção de Zika vírus a ser registrado pela Anvisa, o Zika IgG / IgM Combo da Bahiafarma, fruto de parceria com a sul-coreana Genbody, é composto por dois cassetes portáteis. O primeiro reage ao anticorpo IgM e identifica infecções recentes (até duas semanas), enquanto o segundo, que reage ao IgG, identifica se o paciente foi infectado previamente. Os resultados saem em até 20 minutos. Produto distribuído para todo o país pelo Ministério da Saúde.

**Armadilha para mosquito MOSQUIFIM, 100% ecológica** (Feira de Soluções, Salvador - 2017). As armadilhas para mosquitos MOSQUIFIM usam uma tecnologia inovadora para atrair os insetos pela reação chamada fotocatalise (combinação de luz UV + Dióxido de titânio), imitando o cheiro humano. Funciona por sucção dos mosquitos, que morrem por desidratação. Usando essas características, o Mosquifim ajuda a eliminar do ambiente pernilongos, muriçocas, borrachudos e também o *Aedes Aegypti*, de forma totalmente ecológica.

**Multiteste Sorológico** (Feira de Soluções, Bento Gonçalves - 2019). Para a redução da mortalidade materno-infantil e aumento da efetividade nas ações controle de doenças infecciosas que acometem a mãe e geram risco ao feto durante a gravidez, o IBMP está desenvolvendo um teste sorológico capaz de identificar com uma única gota de sangue doenças como: HIV, Sífilis, Hep. B, Hep. C, Toxoplasmose e Chagas. O teste trará a capacidade de fornecer resultados com a qualidade de um laboratório de análises clínicas, estendendo seus benefícios aos sistemas de saúde.

### ***Hackathon***

Com o objetivo de criar soluções e produtos resolutivos e aplicáveis para um determinado desafio em tempo recorde, durante a Feira Zika foi realizada a 1ª Maratona de Desenvolvimento de Soluções Tecnológicas para enfrentamento da Dengue, Chikungunya, Zika e Síndrome Congênita – *Hackathon-Zika*.

Estratégia de composição de grupos de trabalho para elaboração de soluções tecnológicas: grupos formados com participantes de várias partes do país foram compostos para pensar em soluções tecnológicas para o tratamento e a prevenção da Zika.

Sete grupos, com cinco componentes, foram formados no Hackathon-Zika e apesar de todas as soluções terem sido bem avaliadas pelos juízes do evento, três se destacaram e foram premiadas com bolsa de pesquisa para o desenvolvimento e implantação da solução:

1º lugar – Solução: Zap Zika

2º lugar – Solução: Deu Zika

3º lugar – Solução: Painel de Informação Zika Brasil (PIZB)

### ***Game Jam***

*Game Jam* é uma reunião presencial de desenvolvedores de jogos que tem como objetivo criar um jogo em um curto intervalo de tempo. Nas Feiras de Soluções para a Saúde, foi utilizada para a promoção da saúde de forma criativa, divertida, de modo a alcançar um público diferenciado, crianças, jovens, além dos adultos que se divertem com jogos eletrônicos. O *Game Jam* realizado na Feira de Soluções para a Saúde de Fortaleza/CE (2019), teve como tema o desenvolvimento de jogos para promoção da saúde com a temática Vacinação.



A maratona de desenvolvimento durou 24 horas e teve 28 participantes divididos em 6 grupos.

Figura 47: Maratona de jovens desenvolvedores de jogos digitais



Fonte: Secitece

Cinco equipes terminaram a maratona de *games* e apresentaram seus trabalhos a banca julgadora.

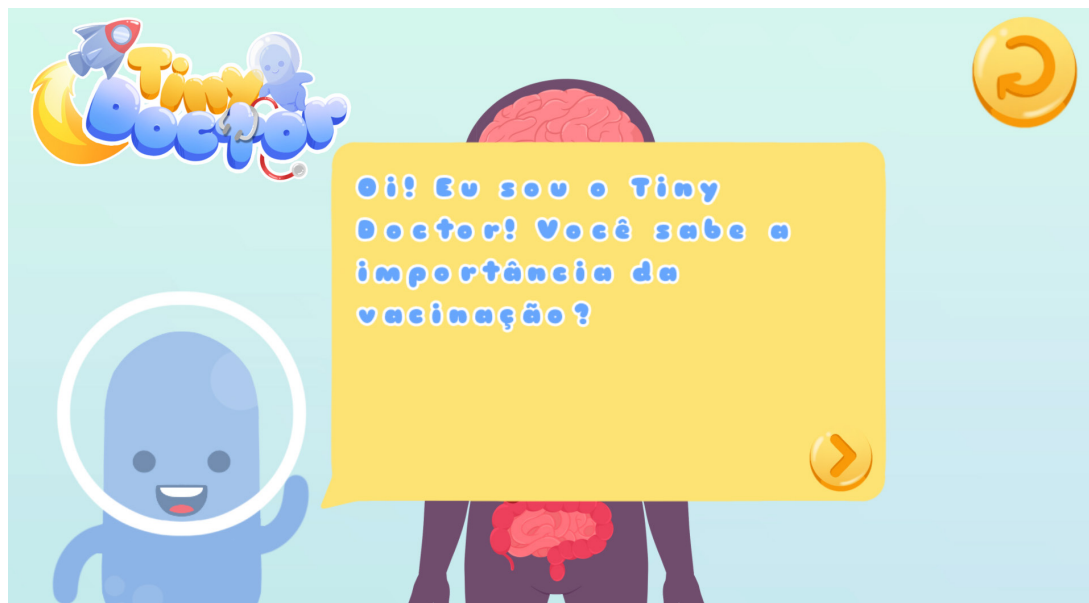
- A 1ª Equipe: apresentou o projeto Vacina, onde o jogador deve vacinar a família de acordo com o tipo de vacina indicada para a idade, segundo o Calendário de Vacinação proposto pelo Programa Nacional de Imunizações (PNI), explorando a ideia de que é melhor e mais fácil prevenir uma doença do que tratá-la.
- A 2ª Equipe: apresentou Vaccine, que abordou a correlação entre vacinas e os agentes etiológicos dos agravos a que elas se destinam a prevenir. Uma brincadeira entre a imunização e a produção de anticorpos para proteção contra vírus e bactérias.
- A 3ª Equipe: apresentou o Imuno Cruz, jogo que buscou demonstrar o funcionamento das vacinas como agentes estimuladores do sistema imunológico a produzir anticorpos que podem combater doenças infecciosas.



- A 4ª Equipe: apresentou o jogo Agente de Saúde, no qual o agente de saúde vai de porta em porta demonstrando a importância das vacinas para a população.
- A 5ª Equipe: apresentou um jogo interativo, *Tiny Doctor*, no qual o personagem principal, a “vacina”, enfrenta os vírus e bactérias combatendo as doenças à medida que caminha pelo corpo humano.

A equipe do jogo *Tiny Doctor* foi a premiada do evento. Recebeu investimento em formato de bolsa no valor de R\$ 1.250,00, durante três meses, para a finalização do desenvolvimento e disponibilização do jogo em plataformas *on-line*.

Figura 48: Jogo *Tiny Doctor*, vencedor do *Game Jam*



Fonte: <https://alisson-alencar.itch.io/tiny-doctor>

## FEIRAS DE SOLUÇÕES PARA SAÚDE COMO INSTRUMENTO DO COMPLEXO INDUSTRIAL DA SAÚDE

Dentre os objetivos do projeto “Plataforma de Vigilância de Longo Prazo para a Zika e Microcefalia no âmbito do SUS” está o fortalecimento das propostas de colaboração e cooperação na rede do Complexo Industrial da Saúde, assim como, o desenvolvimento de bancos de dados que seguissem os preceitos do movimento global de Ciência Aberta (*Open Science*).

O diálogo entre o fortalecimento do Complexo Industrial da Saúde e a Ciência Aberta provocou a construção de novos arranjos de pesquisa, desenvolvimento, inovação e divulgação do conhecimento, de modo a maximizar os recursos financeiros, humanos, técnicos e de bens e serviços em prol da sociedade. Nesse contexto, surgiram as Feiras de Soluções para a Saúde.

As experiências exitosas das quatro Feiras de Soluções evidenciaram a capacidade catalisadora dos eventos para o compartilhamento de experiências bem-sucedidas, divulgação e fortalecimento de ações de saúde, assim como, ativação de redes socio-técnicas nos territórios.

A maior parte das soluções apresentadas tem baixo custo de desenvolvimento e ampliação, o que propicia sua implantação nos diferentes níveis do Sistema Único de Saúde. Entretanto, a falta de recursos financeiros foi relatada como o principal motivo para o não seguimento de projetos, ampliação ou implantação das soluções.

No escopo da política de Ciência Aberta, a formação de bancos de dados abertos de fácil acesso, com 288 soluções para a saúde cadastradas na Plataforma Ágora, permite o mapeamento e a mobilização de ideias no enfrentamento de emergências em saúde.

Muitas das soluções avaliadas possuem potencial de escalonamento, sendo que 37 delas já estão em fase de implementação. Entretanto, a falta de financiamento mos-

trou-se o principal entrave para esse processo de finalização do desenvolvimento, escalonamento e disseminação.

A construção de uma matriz de resultados para as Soluções para a Saúde possibilitou a identificação de melhorias, extensão de abrangência (local, estadual, regional, nacional ou internacional); parcerias de desenvolvimento; custo médio para desenvolvimento e produção, gerando assim, subsídios para o apoio a projetos no âmbito da Agenda Nacional de Prioridades na Pesquisa em Saúde do Ministério da Saúde.

A saúde digital, explorada na forma de eventos como *Hackthons* e *Game Jam* mostrou-se uma ferramenta lúdica potente, atrativa para ações de promoção da saúde, de modo que, novos investimentos, no campo da saúde digital, devem ser direcionados a eventos desta natureza.

O arranjo produtivo na forma de feira mostrou-se, sem dúvida, uma base para a consolidação de uma ciência cidadã.

## CONCLUSÕES

As Feiras de Soluções para a Saúde como estratégia científica inovadora para identificação e disseminar tecnológicas em saúde, mostrou-se um arranjo propício à formação de redes sociotécnicas, disseminação do conhecimento e de parcerias institucionais para promoção da Ciência, Tecnologia & Inovação em Saúde (CT&IS) e translação do conhecimento.

Motivadas pelas dificuldades e necessidades impostas pelas arboviroses, sobretudo pela emergência de saúde causada pelo vírus Zika, as feiras mostraram-se nicho de ciência e cidadania, orientadas pelos Objetivos de Desenvolvimento Sustentável (ODS) e a Agenda 2030 da Organização das Nações Unidas, assim como, pelo preenchimento das lacunas de pesquisa identificadas pelo Ministério da Saúde e estabelecidas em sua Agenda de Prioridades para Pesquisa (2018).

Na primeira versão do evento, realizada em 2017, na cidade de Salvador - BA, o tema foi o “Zika vírus, microcefalia e síndromes congênitas”. Em 2019, foram realizadas duas edições: a primeira realizada em Bento Gonçalves – RS, com o tema “Saúde Única para Territórios Saudáveis e Sustentáveis”, e a segunda realizada em Fortaleza – CE, com o tema “Saúde Digital para Territórios Saudáveis e Sustentáveis”. Em 2020, na sua quarta edição, as condições de saúde pública vigentes impuseram a readequação para a modalidade totalmente digital, em decorrência da pandemia causada pelo novo coronavírus e, teve como tema o “Enfrentando as crises sanitárias e epidemias: panoramas e perspectivas”.

Em suas quatro edições, as Feiras de Soluções para a Saúde cadastraram na Plataforma Ágora 288 soluções inovadoras, sendo predominante as de caráter social e de serviço com parcerias de desenvolvimento e implantação de baixo custo e aplicação em diferentes níveis do Sistema Único de Saúde. A metade, tinha abrangência nacional. Os arranjos locais, ou de cobertura no território, tiveram destaque nas edições locais de cada feira, com resultados que geraram subsídios para o apoio a projetos no âmbito da Agenda Nacional de Prioridades na Pesquisa em Saúde do Ministério da Saúde.

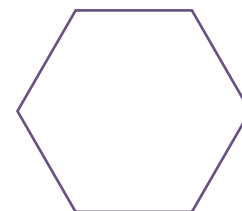
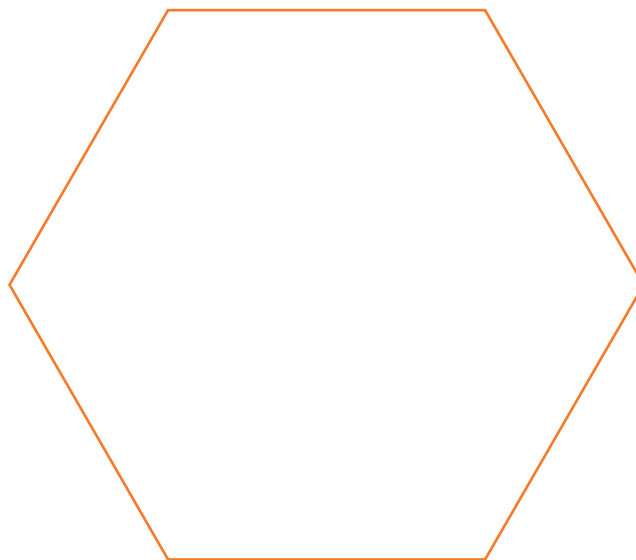
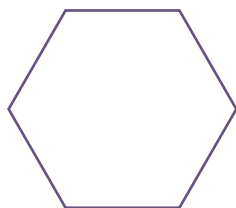
Destaca-se que as soluções industriais, que demandam maior investimento financeiro e tecnológico para o desenvolvimento e escalonamento, foram as de menor número. Entretanto, espaços de diálogo e apresentação propiciaram a aproximação entre demandantes, academia, desenvolvedores dessas soluções e o Complexo Econômico e Industrial da Saúde (CEIS), dando oportunidade à formação de redes sociotécnicas que poderiam mitigar lacunas da pesquisa translacional.

Mais uma vez, as instituições de ensino e pesquisa e as parcerias públicas foram protagonistas do cenário de desenvolvimento e financiamento à inovação científica e tecnológica. A falta de recurso foi relatada como o principal motivo para o não seguimento de projetos, ampliação ou implantação das soluções, o que refletiu o cenário nacional do financiamento à pesquisa e inovação.

Muitas das soluções avaliadas possuíam potencial de escalonamento, sendo que 37 delas já se encontravam em fase de implementação.

A Saúde Digital, explorada na forma de eventos como *Hackthons* e *Game Jam* mostrou-se uma ferramenta lúdica potente, atrativa para ações de promoção da saúde, de modo que novos investimentos devem ser direcionados a eventos dessa natureza.

Sem dúvida, a experiência bem-sucedida alcançada no arranjo produtivo na forma de feira mostrou-se um importante instrumento de interação com a sociedade civil, que nos possibilita a identificação de soluções sociais, criadas para que as comunidades, antecipando-se tanto ao Estado quanto à Academia e de forma solidária, pudessem enfrentar as epidemias e suas consequências. Trabalhando como uma inteligência cooperativa territorial que contribui com o enfrentamento das crises sanitárias e o fortalecimento do Sistema Único de Saúde.



## ARTIGOS CIENTÍFICOS E PRODUTOS RESULTANTES

BRASIL. Fundação Oswaldo Cruz (Fiocruz). **Boletim Feira Soluções para a Saúde: Fiocruz Brasília**. Brasília, DF: Fiocruz, 2019. 37 p. Disponível em: <https://bityli.com/Tzu-XbDQ>. Acesso em: 2 fev. 2023.

CORREIA, C. V. S. R.; MARTIS, W. J. **Relatório técnico** contendo os resultados sistematizados de pesquisas de opinião realizadas com os atores da academia, gestores dos serviços públicos, empresários e cidadãos participantes das Feiras de Soluções para a Saúde: etapas: Salvador, Bento Gonçalves e Ceará. Brasília: Fiocruz Brasília, 2021. [141] f. No prelo.

FUNDAÇÃO OSWALDO CRUZ (Fiocruz). **Feira Soluções para a Saúde – Zika**. 2017. Disponível em: <https://www.youtube.com/watch?v=i26cAqUTdpw>. Acesso em: 02 fev. 2023.

FUNDAÇÃO OSWALDO CRUZ (Fiocruz). **Relatório Final da Feira de Soluções para Saúde Aplicadas às Arboviroses e à Síndrome Congênita do Zika**. 2018. Disponível em: <https://www.fiocruzbrasil.fiocruz.br/wp-content/uploads/2019/11/Livro-fiocruz-ZICA-11-10-2018-WEB-1.pdf>. Acesso em: 02 fev. 2023.

FUNDAÇÃO OSWALDO CRUZ (Fiocruz). **Soluções para a Saúde**. 2021. Disponível em: <https://solucoes.agora.fiocruz.br/>. Acesso em: 02 fev. 2023.

REZENDE, K. S.; MARTIS, W. J. **Relatório técnico** contendo avaliação dos resultados da interação entre academia, gestores, empresários e cidadãos das Feiras de Soluções para a Saúde em construções de soluções industriais, sociais, e de serviço para atendimento de demandas de saúde em Emergências de Saúde Pública e melhoria de Determinantes Sociais de Saúde. Brasília: Fiocruz Brasília, 2021. 31 f. No prelo.

SCHRAMM, J. M. A. *et al.* Feira de Soluções para a Saúde: saúde única para territórios saudáveis e sustentáveis: Fundação Oswaldo Cruz Fiocruz. In: CONGRESSO ANUAL DO CONSELHO DE SECRETARIAS MUNICIPAIS DE SAÚDE DO ESTADO DO RIO GRANDE DO SUL (COSEMS/RS), [31.]. 22-25 abr. 2019. Bento Gonçalves, RS. **Relatório [...]**. Brasília: Fiocruz, 2019.

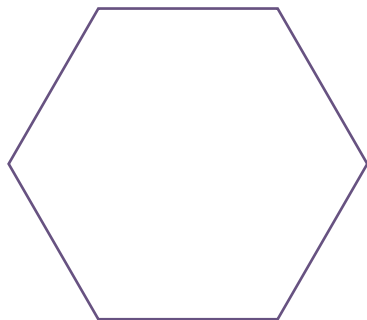
SILVA, G. O.; MARTIS, W. J. **Relatório técnico** contendo o resultado da sistematização do banco de soluções das Feiras de Soluções para a Saúde, das etapas Salvador, Bento Gonçalves e Ceará com as respectivas avaliações prospectivas em relação à maturidade de soluções cadastradas. Fiocruz Brasília, 2021. [40] f. No prelo.

## REFERÊNCIAS

BRASIL. **Agenda de Prioridades de Pesquisa do Ministério da Saúde - APPMS. Ministério da Saúde**. Secretaria de Ciência, Tecnologia e Insumos Estratégicos. Departamento de Ciência e Tecnologia. [recurso eletrônico] . Brasília: Ministério da Saúde, 2018.

VELHO, S. R. K.; SIMONETTI, M. L.; SOUZA, C. R. P.; IKEGAMI, M. Y. Nível de maturidade tecnológica: uma sistemática para ordenar tecnologias. **Parc Estrat.**, Brasília, DF, v. 22, n. 45, p. 119–140, jul./dez. 2017. Disponível em: [https://seer.cgee.org.br/parcerias\\_estrategicas/article/view/867](https://seer.cgee.org.br/parcerias_estrategicas/article/view/867). Acesso em: 2 fev. 2023.

WORLD HEALTH ORGANIZATION (WHO). **Global strategy on digital health 2020-2025**. Geneva: World Health Organization, 2021. Disponível em: <https://apps.who.int/iris/bitstream/handle/10665/344249/9789240020924-eng.pdf>. Acesso em: 2 fev. 2023.





# Ciência Aberta e Dados para Pesquisa: contexto, desafios e oportunidades

---

*Bethânia de Araújo Almeida  
Alexandro Rodrigues Pinto  
Márcia Luz da Motta  
Paula Xavier dos Santos  
Vanessa Arruda Jorge*



A ciência sempre foi aberta na medida em que teorias, métodos e resultados são avaliados por pares e divulgados amplamente. Contudo, nas últimas décadas, com a globalização e a ampla utilização de tecnologias da informação e comunicação, o chamado movimento da Ciência Aberta vem ganhando força e visibilidade, principalmente a partir de 2002 com a *Budapest Open Access Initiative*, que defende a disponibilização livre e gratuita de publicações científicas (BOAI, 2002).

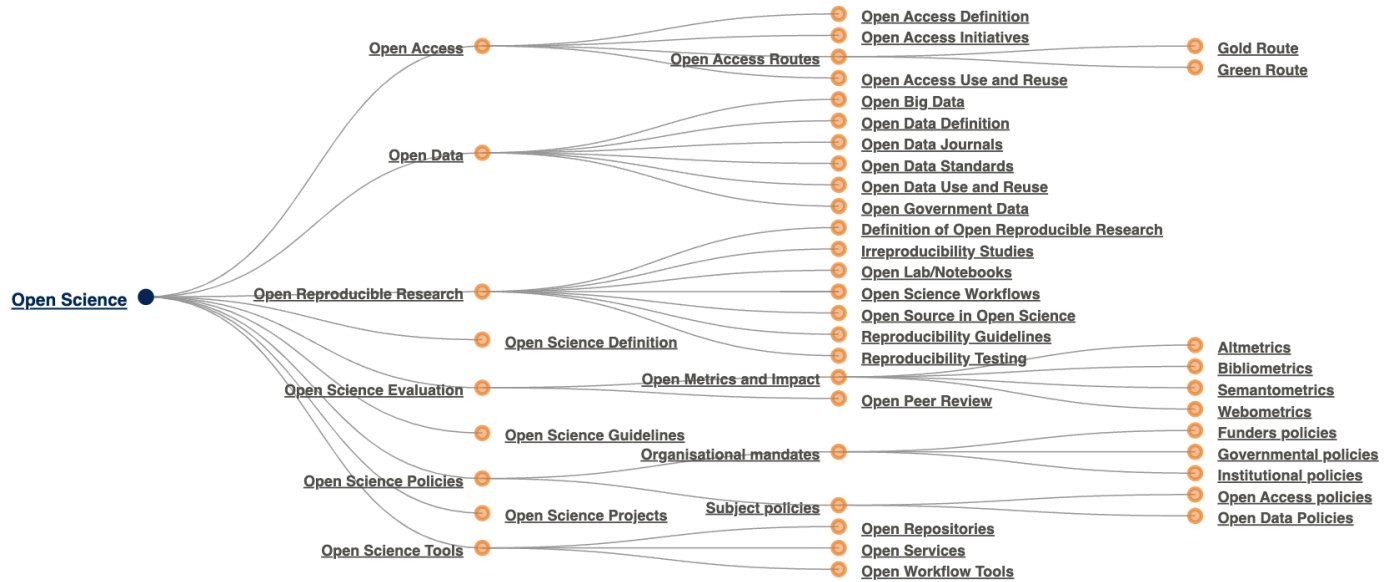
Em 2007, com a publicação de diretrizes para orientar o compartilhamento de dados de pesquisas financiadas com recursos públicos pela Organização para Cooperação e Desenvolvimento Econômico (OECD, 2007), o compartilhamento e a abertura de dados científicos tornaram-se uma das principais propostas do movimento da Ciência Aberta. Proposta que suscita um intenso debate sobre os princípios e os padrões a serem seguidos para o acesso aos dados de pesquisas.

A abertura de dados ganhou relevância pelo potencial de otimizar tempo e recursos dedicados à pesquisa, promovendo, assim, o aumento na circulação de informações e conhecimento, que devem ser orientados por valores como liberdade acadêmica e integridade científica. O objetivo é obter uma ciência mais transparente e colaborativa com progressos mais rápidos e alinhada às necessidades e desafios das sociedades contemporâneas.

A Ciência Aberta vai além do acesso a publicações e do compartilhamento de dados originados em projetos de pesquisa, na medida em que promove a abertura de todo o processo científico e a translação do conhecimento. Esse movimento, reforça o conceito de responsabilidade social científica e é potencialmente capaz de ampliar os impactos sociais e econômicos da ciência.

Nessa perspectiva, novas práticas voltadas à disponibilização aberta de conhecimento científico começam a surgir para viabilizar o seu acesso e reutilização. Exemplos dessas práticas estão presentes na Taxonomia da Ciência Aberta, elaborada pela *Facilitate Open Science Training for European Research* (FOSTER, 2016).

Figura 1: Taxonomia da Ciência Aberta



Fonte: <https://www.fosteropenscience.eu/foster-taxonomy/open-data?page=1>

Os principais benefícios de uma ciência aberta, que constituem a base de suas motivações, são: reprodutibilidade, transparência científica, velocidade de circulação da informação e reúso de dados, resultando numa ciência de maior qualidade e progressos mais rápidos.

A despeito do novo paradigma científico ser pautado no compartilhamento e reutilização de dados, de métodos e do conhecimento científico, pesquisadores e suas instituições estão imersos em contextos nacionais que podem potencializar, ou dificultar a abertura e o acesso a eles.

Em 2021, a Unesco lançou recomendações sobre a Ciência Aberta aos Estados-membros, que deverão ser aplicadas a partir de medidas apropriadas, aderentes às estruturas jurídicas e governamentais de cada Estado. Para tal, as recomendações deverão ser de conhecimento das autoridades e órgãos responsáveis pela ciência, tecnologia e inovação e discutidas com atores envolvidos com a Ciência Aberta a nível nacional, internacional e multilateral (UNESCO, 2021)<sup>1</sup>.

Essas recomendações ganham maior relevância e complexidade no setor de pesquisa em saúde, por ser uma área sensível e estratégica para a sociedade, particularmente em países que possuem sistemas universais de saúde, como o Brasil, e em cenários de enfrentamento de emergências sanitárias, como no caso da Zika e mais recentemente da pandemia de Covid-19.

No âmbito do enfrentamento de emergências de saúde pública, destaca-se a declaração das academias de ciências das nações do Grupo dos Sete (G7), que engloba as sete maiores economias do mundo. O documento recomenda uma série de ações a serem tomadas para viabilizar o compartilhamento de dados entre países para enfrentar as próximas emergências de saúde pública (ROYAL SOCIETY, 2021).

---

1. Recomendação da UNESCO sobre Ciência Aberta disponível em português. Ver UNESCO, 2021

De forma pioneira no País, em 2017, a Fiocruz constituiu o Grupo de Trabalho em Ciência Aberta (GTCA) no âmbito do projeto “Plataforma de Vigilância de Longo Prazo para Zika e Microcefalia no âmbito do SUS” – Plataforma Zika – com a participação inicial das seguintes instâncias do órgão: Centro de Integração de Dados e Conhecimentos para Saúde (Cidacs/Fiocruz Bahia), Gerência Regional de Brasília (Gereb/Fiocruz) e da Vice - Presidência de Educação, Informação e Comunicação (VPEIC)<sup>2</sup> da Fiocruz. Em seguida, passou a contar com membros de outras instâncias da Fiocruz, e começou a ser chamado de GTCA ampliado.

A criação do GTCA objetivou promover o desenvolvimento de princípios e diretrizes para orientar o acesso, compartilhamento e abertura de dados para pesquisa na perspectiva da Ciência Aberta, inicialmente no âmbito do projeto, para posterior adoção institucional<sup>3</sup>, com potencial de replicação em outras instituições públicas.

Os resultados obtidos por meio dos estudos realizados pelo GTCA até o segundo semestre de 2018 foram apresentados aos dirigentes e técnicos dos Ministérios da Saúde, da Ciência, Tecnologia, Inovação e das Comunicações por ocasião do Seminário sobre Abertura, Gestão e Compartilhamento de Dados para Pesquisa, realizado em 25 de outubro de 2018. O seminário resultou da parceria entre a Fiocruz Brasília, o Cidacs/Fiocruz Bahia, o Departamento de Ciência e Tecnologia do Ministério da Saúde (DECIT/MS) e a Vice-Presidência de Educação, Informação e Comunicação (VPEIC) da Fiocruz. Além desses atores, participaram também dirigentes do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), do Instituto Brasileiro de Informação em Ciência e Tecnologia (Ibict) e da Comissão Nacional de Ética em Pesquisa (Conep), bem como, trabalhadores da Fiocruz.

A iniciativa promoveu o debate sobre o estado da arte da Ciência Aberta no mundo e os impasses e perspectivas para implementação da abertura de dados pelas instituições brasileiras, em especial, o Ministério da Saúde (MS), uma vez que este é, na

---

2. Membros do GTCA: Alexandre Pinto, Anne Clio, Bethânia Almeida, Flávia Elias, Gabriela Oliveira, Márcia Motta, Patrícia Henning, Paula Xavier, Paulo Guanaes e Vanessa Jorge.

3. Para obter informações adicionais sobre o trabalho do GTCA, acesse: <https://cidacs.bahia.fiocruz.br/plataformazika/ciencia-aberta/>

atualidade, uma das instituições do Governo Federal mais acionadas pelos brasileiros na procura de dados, tanto administrativos como aqueles gerados pelas pesquisas financiadas pelo órgão.

Além dos debates promovidos durante o Seminário, foram entregues ao Ministério da Saúde todo o material produzido pelo grupo até aquele momento, a saber: o “Livro Verde - Ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional”; o livro intitulado “Marcos regulatórios e a abertura de dados para pesquisa em saúde: dados pessoais, sensíveis ou sigilosos e propriedade intelectual”; e o Termo de Referência sobre Gestão, Compartilhamento e Abertura de Dados para Pesquisa.

Na sequência, apresenta-se, em ordem cronológica e de forma concisa, os produtos listados acima, bem como, as demais ações desenvolvidas pelo GTCA. Informações adicionais e mais detalhadas poderão ser encontradas nos *links* das publicações e nas notas de rodapé disponibilizados.

## 2017

### **LIVRO VERDE - CIÊNCIA ABERTA E DADOS ABERTOS: MAPEAMENTO E ANÁLISE DE POLÍTICAS, INFRAESTRUTURAS E ESTRATÉGIAS EM PERSPECTIVA NACIONAL E INTERNACIONAL**

No mesmo ano de constituição do GTCA, entre março e agosto de 2017, foi realizado um estudo exploratório sobre abertura de dados científicos e governamentais, baseado na análise de documentos oficiais e sítios eletrônicos de instituições governamentais, multilaterais e de fomento, além de outros documentos contendo relatos de iniciativas nessa mesma direção.

Figura 2: Livro Verde



Fonte: <https://www.arca.fiocruz.br/handle/icict/24117>

O estudo recebeu o nome de Livro Verde por representar um mapeamento inicial sobre o tema, uma vez que este é dinâmico e encontra-se em constante evolução, visto que proliferam iniciativas em prol da ciência aberta.

A iniciativa objetivou obter um panorama internacional sobre a abertura de dados para pesquisa, por meio do mapeamento e análise de políticas, infraestruturas e estratégias voltadas a esse fim.

A escolha dos países participantes do estudo teve por critério a identificação daqueles considerados mais destacados na discussão acerca do tema e do país de interesse de aplicação, no caso o Brasil. O estudo analisou iniciativas em sete países (Alemanha, Austrália, Brasil, Canadá, Estados Unidos da América, Holanda, Portugal, Reino Unido) e em um bloco econômico e político, a União Europeia<sup>4</sup>.

---

4. Os resultados do estudo, por país, poderão ser consultados diretamente na publicação.

## O que foi encontrado?

Especificamente no que tange ao estímulo para abertura de dados científicos, a maioria das iniciativas foi liderada por agências de fomento à pesquisa. A apresentação de planos de gestão de dados e o depósito de dados em repositórios passaram a figurar como critérios para concessão de financiamento por essas agências. Passaram a ser permitidos nos orçamentos dos projetos, a inclusão dos custos relacionados à operacionalização dos planos de gestão de dados, como os que incluem o tratamento, a organização e a documentação dos dados, bem como, investimentos em infraestrutura (*hardware* e *software*) e pessoal especializado<sup>5</sup>.

Neste sentido, cabe chamar atenção para a importância de pessoal especializado voltado à curadoria de dados, que além de auxiliar na organização, descrição e documentação dos dados, e no provimento de metadados para sua adequada localização e utilização, também auxilia na proteção e gerenciamento de riscos relacionados aos acessos, usos e reúsos dos dados<sup>6</sup>.

A gestão de dados na perspectiva da Ciência Aberta reconhece aspectos éticos e legais, dentre outros, como proteção de direitos humanos e segurança nacional para não compartilhar dados, desde que as restrições sejam proporcionais e justificadas. Portanto, os dados poderão ser fechados, acessados de forma controlada (sob determinados termos e condições), ou abertos (acessados livremente, sem qualquer restrição). Mesmo que os dados não sejam compartilhados, metadados deverão ser disponibilizados de maneira aberta, para que pesquisadores e demais partes interessadas saibam da sua existência, conteúdo e período de criação.

Os principais aspectos considerados no planejamento da gestão e compartilhamento de dados são: ética (particularmente consentimento e privacidade); estruturas e padrões de qualidade dos dados, viabilidade de armazenamento e acesso a longo prazo.

5. Algumas iniciativas que auxiliam na estimativa de custos de gestão de dados de projetos de pesquisa: OpenAire (<https://www.openaire.eu/how-to-comply-to-h2020-mandates-rdm-costs>); UK Data Service (<https://ukdataservice.ac.uk/learning-hub/research-data-management/plan-to-share/costing/>); Horizon 2020 ([https://twiki.pp.rl.ac.uk/twiki/pub/Main/PpdBids/EC\\_Horizon\\_2020\\_General\\_costing\\_guide.pdf](https://twiki.pp.rl.ac.uk/twiki/pub/Main/PpdBids/EC_Horizon_2020_General_costing_guide.pdf))

6. Para saber mais sobre curadoria de dados de pesquisa, acesse o site do Digital Curation Centre: <https://www.dcc.ac.uk/>

A busca por convergência internacional no tocante à gestão de dados de pesquisa para otimizar a sua reutilização na perspectiva da Ciência Aberta, tornou os denominados “princípios FAIR” o padrão adotado. O acrônimo FAIR significa que o conjunto de dados deve ser encontrável (*Findable*), acessível (*Accessible*), interoperável (*Interoperable*) e reutilizável (*Reusable*) no ecossistema global de dados digitais<sup>7</sup>.

Os princípios FAIR buscam orientar a aplicação de uma série de requisitos relacionados à gestão e disponibilização de dados, para que estes possam ser localizados, acessíveis, compreensíveis, rastreáveis, interoperáveis e reutilizados. Destaca-se que fazem distinção entre dados e metadados para apoiar uma ampla gama de circunstâncias, dentre as quais aderência aos marcos éticos e legais.

Quanto à infraestrutura para depósito, acesso, compartilhamento e abertura de dados originados de projetos de pesquisa, existem diversos repositórios e plataformas vinculados a universidades e instituições de pesquisa, financiados por governos e/ou agências de fomento à pesquisa. Há ainda repositórios de áreas especializadas do conhecimento, em geral promovidos por sociedades ou associações científicas<sup>8</sup>.

No campo da gestão pública, iniciativas voltadas à abertura de dados visam a transparência e o acesso à informação pelos cidadãos e organizações da sociedade civil<sup>9</sup>. Notadamente no que tange aos dados de pesquisa, destacamos iniciativas nacionais voltadas ao provimento de infraestrutura, ferramentas e serviços integrados para depósito, busca, reutilização e preservação de dados de projetos de pesquisa na perspectiva da Ciência Aberta<sup>10</sup>.

Em relação à formação de recursos humanos, ainda de forma tímida, algumas iniciativas mencionam a necessidade de formação em áreas estruturantes como informática, estatística, ciência de dados, ética, governança e infraestrutura, relacionadas ao

---

7. Para saber mais sobre os princípios FAIR, acesse: <https://www.go-fair.org/fair-principles/>

8. Uma boa fonte de consulta sobre repositórios de dados de pesquisa é o Re3data.org, financiado pela Fundação Alemã de Pesquisa - German Research Foundation (<https://www.re3data.org/about>).

9. <https://www.opengovpartnership.org/about/>

10. European Open Science Cloud ([https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc\\_en](https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/open-science/european-open-science-cloud-eosc_en)); Australian Research Data Commons ([https://ardc.edu.au/about\\_us/](https://ardc.edu.au/about_us/)).



campo que emerge como *E-Science*. A maior ênfase volta-se à capacitação de pesquisadores para elaboração de planos de gestão de dados gerados por projetos de pesquisa financiados com recursos públicos.

A ênfase em planos de gestão de dados é compreensível pela necessidade de provimento de informações e documentação sobre a organização dos dados ao longo do ciclo de vida do projeto, principalmente sobre os dados que serão compartilhados e preservados para reutilização. Nesse sentido, agências de fomento e instituições de pesquisa passaram a disponibilizar seus respectivos modelos de plano de gestão de dados com tutoriais de preenchimento, alguns dos quais encontram-se disponíveis em plataformas dedicadas ao tema<sup>11</sup>.

A despeito de especificidades, em termos gerais, a maioria dos países busca viabilizar a abertura de dados de pesquisa financiadas com recursos públicos, por considerar que são bens produzidos em nome do interesse público. Como tal, devem ser disponibilizados de forma aberta, com as menores restrições possíveis, de maneira oportuna e responsável.

## Considerações

O mapeamento e análise de políticas, infraestrutura e outras estratégias para abertura de dados na perspectiva da Ciência Aberta aponta desafios e oportunidades relacionados à necessidade de mudanças culturais, políticas e técnicas. A indução através de requisitos obrigatórios nos editais de financiamento e disponibilização de cursos e tutoriais voltados à gestão de dados de pesquisa, foram identificados como os principais meios para sensibilizar pesquisadores acerca da necessidade de gerir e compartilhar dados de suas pesquisas. Constata-se ainda que países passam a investir de

---

11. Alguns exemplos de plataformas que disponibilizam modelos de planos de gestão de dados de pesquisa: DMPTool Building your Data Management Plan (<https://dmptool.org/>); DMPONLINE (<https://dmponline.dcc.ac.uk/>)/(<https://dmponline.dcc.ac.uk/>). Exemplo de curso sobre o tema Research Data Management Training: <https://mantra.ed.ac.uk/>

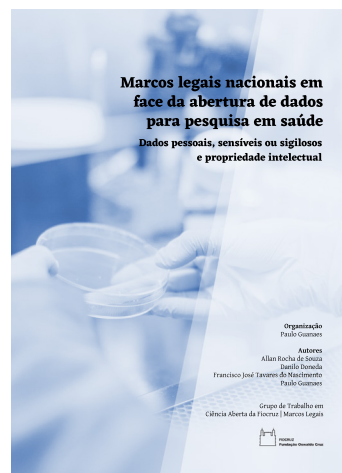
forma estratégica em infraestrutura e pessoal especializado, para adequada gestão, armazenamento, preservação e acesso aos dados para pesquisa científica por diferentes áreas disciplinares, e em alguns casos, os investimentos são declaradamente relacionados às suas políticas de desenvolvimento científico, econômico e social.

## 2018 E 2019

### **MARCOS REGULATÓRIOS E A ABERTURA DE DADOS PARA PESQUISA EM SAÚDE: DADOS PESSOAIS, SENSÍVEIS OU SIGILOSOS E PROPRIEDADE INTELECTUAL**

No ano seguinte, buscou-se entender o ordenamento jurídico brasileiro no que concerne à identificação dos marcos legais que incidem sobre temas vinculados à abertura de dados para pesquisa em saúde, com a finalidade de fornecer subsídios e diretrizes aderentes às estruturas jurídicas do país, para orientar o acesso, compartilhamento e abertura de dados na perspectiva da Ciência Aberta no projeto e na Fiocruz.

Figura 3: Marcos legais nacionais em face da abertura de dados para pesquisa em saúde: dados pessoais, sensíveis ou sigilosos e propriedade intelectual



Fonte: <https://www.arca.fiocruz.br/handle/icict/28838>

O estudo foi pautado em pesquisa documental, com base em fontes legislativas primárias do período de 1988 a 2018. As normas identificadas incidem na abertura de dados para pesquisa em saúde – seja de forma direta, ou indireta. A publicação traz um conjunto de normas em uma planilha específica, onde são apresentados os dispositivos pertinentes, seu conteúdo (objeto), as razões de sua seleção (justificativa) e comentários explicativos.

### **O que foi encontrado?**

O estudo destaca que existem normas relacionadas à abertura de dados para pesquisa, mas que ainda não havia regramento que regulasse de forma específica o acesso, o tratamento e o uso de dados pessoais para pesquisa acadêmica e estudos em saúde pública, ressalvada neste caso, a Lei Geral de Proteção de Dados Pessoais - LGPD, Lei n. 13.709 - (BRASIL, 2018), que conta com dispositivos a respeito. À época da elaboração do estudo, a LGPD havia sido sancionada, mas ainda não estava em vigência.

Especificamente no que concerne aos direitos de propriedade intelectual sobre bancos de dados, desde a seleção, organização ou disposição de seu conteúdo, estes são protegidos por direitos autorais, sempre se considerando as especificações da Lei do *Software*, quando digitais. Entretanto, dados e informações científicas *per se* não são protegidos por direitos autorais, embora os dados clínicos e demais relacionados aos inventos sujeitos a patentes e mesmo segredos de negócios, possam ser protegidos pela Lei de Propriedade Industrial.

O objeto da proteção por direitos autorais é o banco de dados em si, ou seja, o conjunto ordenado e organizado de dados. Uma vez protegido, sua reprodução, distribuição e comunicação, bem como, sua adaptação, reordenação e modificação bem como, a comunicação e reprodução dos resultados das operações, são exclusivas dos titulares, que podem autorizar ou não, além de determinar as condições de seu uso por via contratual.

Por fim, indica que cabe às instituições científicas e às instituições de fomento regular, o compartilhamento e a abertura de dados para pesquisa em saúde, com base nas diretrizes e limites normativos para garantir o acesso aos dados da maneira mais ampla possível, sem comprometer a integridade dos dados e as normativas éticas e legais.

### **Considerações**

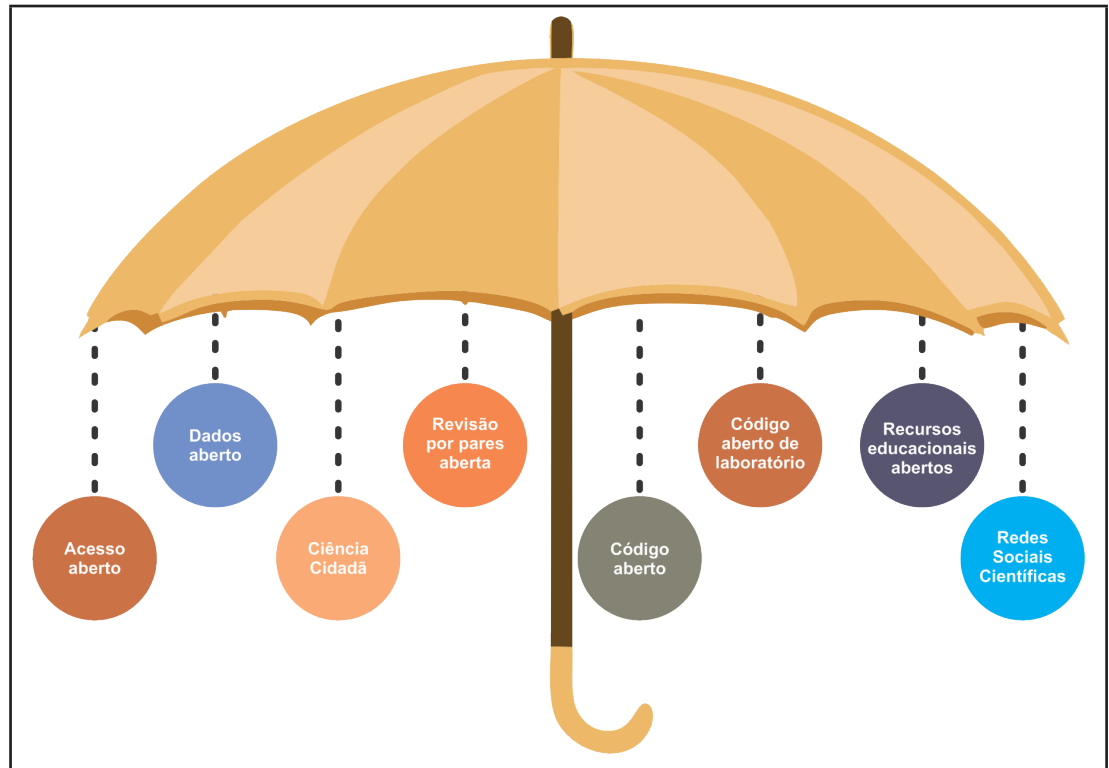
A principal constatação do estudo volta-se à recente aprovação de normas substanciais específicas, que dispõem sobre tratamento e proteção de dados pessoais e sobre dados para pesquisa em saúde. Também foi identificada uma variedade de normas jurídicas pertinentes ao tema, como a Lei de Direitos Autorais, a Lei de Acesso à Informação, o Código Civil e a Constituição Federal. Apesar dos múltiplos interesses (científicos, comerciais, políticos, entre outros) em torno da proteção e utilização dos dados pessoais – elemento crucial na pesquisa em saúde –, a efetividade normativa de sua proteção ainda está em construção.

As instituições de pesquisa, universidades públicas e agências de fomento públicas brasileiras necessitam criar normas internas na forma de políticas institucionais que regulamentem as relações que decorrem desse processo. O estudo buscou subsidiar um debate no âmbito do projeto e da Fiocruz, na medida em que indica as diretrizes a serem seguidas e as fronteiras que não devem ser transpostas diante da necessidade de se autorizar o acesso, o compartilhamento e a abertura de dados custodiados pela instituição.

## SENSIBILIZAÇÃO E CAPACITAÇÃO SOBRE CIÊNCIA ABERTA

Em 2018, a Fiocruz lançou uma Formação Modular sobre Ciência Aberta, que contou com a participação de integrantes do GTCA como conteudistas. O curso<sup>12</sup> foi organizado em séries temáticas que reúnem cursos gratuitos, na modalidade ensino à distância, e abertas a qualquer pessoa interessada.

Figura 4: Formação Modular sobre Ciência Aberta



Fonte: Site do Campus Virtual Fiocruz

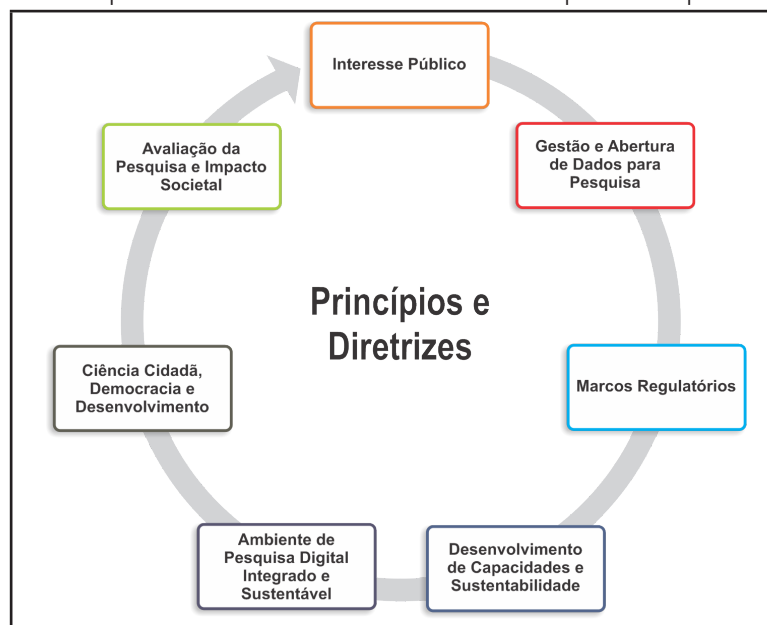
12. Disponível em: <https://bit.ly/formacaocienciaaberta>

## Avanços institucionais

Com base nas discussões iniciadas e nos estudos conduzidos pelo GTCA, foi elaborado um Termo de Referência sobre Gestão, Compartilhamento e Abertura de Dados para Pesquisa<sup>13</sup>. O documento sistematiza um conjunto de princípios e diretrizes e foi colocado em consulta pública para a construção de uma visão coletiva para subsidiar as diretrizes institucionais.

O Termo de Referência foi a base de discussão da comunidade Fiocruz e em 2019 foi formalmente disponibilizado para contribuições em sistema de consulta pública interna. Após análise e incorporação das contribuições recebidas, em 2020, uma nova versão do documento foi aprovada pelo Conselho Deliberativo da Fiocruz.

Figura 5: Princípios e Diretrizes no Termo de Referência sobre Gestão, Compartilhamento e Abertura de Dados para Pesquisa



Fonte: Figura elaborada pelo GTCA, 2019

13. Disponível em: <https://bit.ly/termodereferenciaca>

## 2020

A versão aprovada do documento tornou-se a “Política de gestão, compartilhamento e abertura de dados para pesquisa: princípios e diretrizes da Fiocruz”, instituída por meio da Portaria nº 5.760, de 7 de dezembro de 2020<sup>14</sup>.

A construção dessa Política envolveu um amplo e extenso processo de discussões e pactuações internas, com diferentes atores que realizam pesquisas e com diversas instâncias de gestão e consultivas, vinculadas às atividades de investigação científica na Instituição (câmaras técnicas de pesquisa e de informação e comunicação, unidades da Fiocruz e grupos que se dedicam a estudos em temas específicos).

A criação de uma política institucional que define novos parâmetros e contornos de produção, organização e comunicação da ciência, de forma democrática e inclusiva, em uma instituição complexa, com distintos campos de atuação, exigiu a adoção de um método que abarcasse as diferentes visões sobre as tensões e benefícios gerados pela abertura de dados. Esse processo amplo e participativo permite que a metodologia utilizada e os conhecimentos gerados possam ser extrapolados a outros órgãos públicos, a exemplo do próprio Ministério da Saúde, com as devidas adaptações de contexto, perfis e missões institucionais.

---

14. Disponível em: <https://www.arca.fiocruz.br/handle/icict/46408>

Figura 6: Política de gestão, compartilhamento e abertura de dados para pesquisa, Fiocruz



Fonte: Acervo Fiocruz

Cabe ainda registro que a implementação dessa Política encontra-se em andamento, a partir da criação do Fórum de Ciência Aberta em 2021, instância de governança das políticas de Ciência Aberta na Fiocruz, estabelecida pela Portaria 157, de 12 de maio de 2021<sup>15</sup>.

Ao considerar que os objetivos iniciais do GTCA foram alcançados - tendo em vista o estabelecimento dos princípios e diretrizes que atualmente direcionam a gestão, acesso, compartilhamento e abertura de dados para pesquisa -, tanto no âmbito da Plataforma Zika como em toda a Fiocruz<sup>16</sup>, iniciamos uma nova fase de trabalho com estudos voltados à utilização de dados administrativos para pesquisa científica e à realização de estudos em saúde pública que abordam, se relacionam ou geram conhe-

15. Disponível em: <https://bit.ly/portaria157>

16. As ações de implementação da Política de gestão, compartilhamento e abertura de dados na Fiocruz é de responsabilidade da Vice-Presidência de Educação, Informação e Comunicação (VPEIC).



cimentos para subsidiar a adoção de boas práticas para gestão, compartilhamento e abertura de dados para pesquisa.

## 2020/2021

### **CENÁRIOS SOBRE ZIKA - CONTRIBUIÇÕES DA PESQUISA CIENTÍFICA COM USO DE DADOS ADMINISTRATIVOS VINCULADOS**

O uso de dados para pesquisa não se limita aos dados coletados ou gerados no âmbito de projetos de pesquisa, caracterizados como dados científicos. Para entender o potencial uso de dados administrativos para pesquisa, buscamos analisar as contribuições de dois estudos conduzidos no âmbito da Plataforma Zika, desenvolvidos a partir do acesso e integração de bases de dados administrativos<sup>17</sup>.

A primeira pesquisa analisou a taxa de letalidade e as características materno-infantis dos casos e óbitos relacionados à microcefalia/Síndrome Congênita do Zika (SCZ) no Brasil, entre os anos de 2015 a 2017, a partir da vinculação de três bases de dados: Registro de Eventos de Saúde Pública (Resp), Sistema de Informações sobre Nascidos Vivos (Sinasc) e o Sistema de Informações sobre Mortalidade (SIM). O resultado do estudo alterou o conhecimento anterior sobre a taxa de letalidade em crianças com microcefalia, que antes era considerada em torno de 4-8% e passou a ser de 10%.

Já o segundo estudo selecionado teve como objetivo avaliar o impacto da epidemia de Zika no registro de crianças com anomalias congênitas no Brasil, que considerou o período de janeiro de 2010 a setembro de 2015 como um “evento pré-Zika” e o período logo após setembro de 2015 a dezembro de 2017 como um “evento pós-Zika”. A pesquisa observou o aumento na taxa de notificação de anomalias congênitas, durante e depois da epidemia de Zika, sugerindo uma melhoria geral no registro de malforma-

17. O termo “dados científicos” refere-se aos dados produzidos e/ou coletados em atividades científicas, enquanto o termo “dados administrativos” se refere àqueles coletados por diferentes departamentos governamentais para subsidiar atividades administrativas e prestação de serviços à população.

ções congênitas, incluindo aquelas não atribuídas ao Zika.

O estudo demonstra que o uso de dados administrativos vinculados foi fator determinante para o alcance dos resultados e contribuições das pesquisas acerca dos potenciais cenários e impactos da Zika no Brasil. Adicionalmente, apontamos desafios éticos, legais e metodológicos em torno da integração de dados administrativos enfrentados pela Plataforma Zika, a partir do estabelecimento de arranjos de segurança da informação adequados e pessoal especializado para receber, processar, integrar e prover acesso aos dados administrativos integrados com qualidade e acurácia para viabilizar a condução dos estudos.

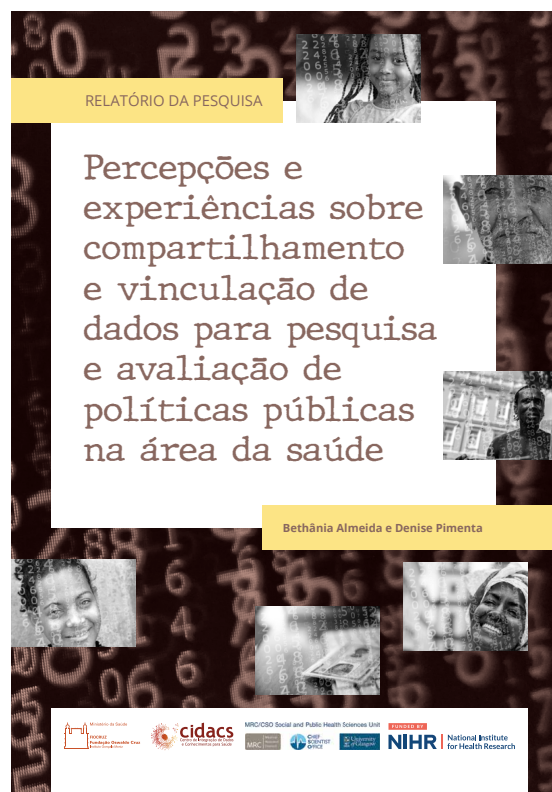
## **PERCEPÇÕES E EXPERIÊNCIAS SOBRE COMPARTILHAMENTO E VINCULAÇÃO DE DADOS PARA PESQUISA E AVALIAÇÃO DE POLÍTICAS PÚBLICAS NA ÁREA DA SAÚDE**

No contexto da entrada em vigor da LGPD, marco na regulamentação e utilização de dados pessoais no Brasil, entre 2020 e 2021, foi conduzido um estudo exploratório para compreender as experiências e percepções de indivíduos de determinados segmentos da sociedade acerca do compartilhamento e vinculação de dados pessoais tanto na esfera pública (pesquisas científicas e políticas públicas), quanto na esfera privada. A investigação deu especial atenção aos dados administrativos, entendendo-os como bens públicos de grande valor para a sociedade.

Por ser um tema novo, com muitas lacunas para a sua compreensão, particularmente no Brasil, os participantes foram escolhidos a partir de características diversas, objetivando garantir uma variedade de impressões. Foram entrevistados representantes de grupos de pacientes, beneficiários de programas sociais, pesquisadores usuários de dados administrativos das áreas de saúde, ciências sociais, história e patrimônio e ciência da computação e gestores públicos com experiência nas áreas de saúde e proteção social.

Para condução das entrevistas semiestruturadas foram construídos roteiros específicos para cada grupo de entrevistados, três ao todo. Foram mais de 24 horas de gravação de áudio e vídeo e mais de 540 páginas transcritas, analisadas e comparadas. Através deste rico material, foi possível encontrar padrões sobre temas que subsidiaram o mapeamento de categorias que emergiram a partir das análises comparadas das entrevistas.

Figura 7: Relatório da pesquisa de percepções e experiências sobre compartilhamento de dados



Fonte: Almeida e Pimenta, 2021

## O que foi encontrado?

A maioria das entrevistas aponta que o processamento de dados pessoais por parte do governo é algo que gera maior receio do que ações de grandes empresas de tecnologia (*Big Techs*) que oferecem serviços de busca, *e-mail*, aplicativos e redes sociais, por exemplo. Receio justificável, se pensarmos que é a relação com o Estado que confere cidadania ao indivíduo, garantindo a liberdade e o acesso a direitos, com seus contornos jurídicos amparados pela Constituição e contornos sociológicos.

Figura 8: Depoimento do Relatório da pesquisa de percepções e experiências sobre compartilhamento de dados

“

*“Até porque quando eu fui dar entrada no Minha Casa Minha Vida, eles provavelmente juntaram os dados pra ver que eu morava com a minha mãe... então assim, foram dados que pra casa eu não tinha fornecido. Na verdade, quando me pediram os dados, pediram os meus dados com os de meu filho. E aí, quando veio a negativa... na verdade não veio uma negativa, né, na verdade saiu no diário oficial com o nome de outras mães e o meu não estava.*

*E aí, foi toda uma burocracia pra saber o que foi que havia acontecido, mas pra eles verem que minha mãe fazia parte do meu círculo de residência, no caso vamos dizer assim, eles sabiam, entendeu?! Porque assim, eu recebo um benefício que é um salário mínimo, minha mãe recebe um salário mínimo também. Somos nós duas. Aqui em casa, somos em quatro: eu, minha mãe, meu pai e meu filho. Eu tô com o pai do meu filho ainda, mas assim, ele mora com os pais dele. Então, aqui em casa somos nós quatro. Então, o que acontece, eles ficaram sabendo... Mas eu acho que eles têm acesso... a gente assim fica meio preocupada... E a gente começa a pensar várias coisas porque, querendo ou não, são todos os nossos dados. Então é isso, a gente fica preocupado sim.”*

*- Beneficiária de Pensão Vitalícia (antigo Benefício de Prestação Continuada BPC), mãe de criança com microcefalia decorrente do Zika vírus (entrevistada para esta pesquisa)*

Fonte: Almeida e Pimenta, 2021

Entretanto, os titulares de dados tendem a aprovar o uso de seus dados custodiados pelo Estado para pesquisas científicas e para políticas públicas, desde que tenham salvaguardas (proteção e garantias de direitos) e que isso gere algum benefício na vida delas. O acesso à informação e o retorno dos resultados de pesquisas para pessoas e comunidades às quais os dados se referem, também foram considerados importantes, pois poderão fazer uso destes para saber mais sobre suas condições de vida e saúde e ter subsídios para acessar direitos respaldados e garantidos por lei.

Figura 9: Depoimento do Relatório da pesquisa de percepções e experiências sobre compartilhamento de dados

“

*“Porque gente, hoje tem um grande número de casas [terreiros de religião de matriz africana] que não sabe nem que pode fazer isso. Ou que precisam fazer isso. Que têm isenção de IPTU ou não têm. A cidade de Salvador já deu isenção de IPTU, mas a maior parte das casas continua pagando IPTU porque nem sabe que pode deixar de pagar. Então assim, eu acho que tem uma série de informações que não chegou para o público-alvo e tem uma série de informações que a gente adoraria saber e que a gente também não tem”.*

*- Pesquisadora da área de História e Patrimônio (entrevistada para esta pesquisa)*

Fonte: Almeida e Pimenta, 2021

Especificamente no que concerne à pesquisa científica e estudos em saúde pública, os achados apontam para a necessidade de delineamento e implementação de formas de envolvimento e engajamento da sociedade para abarcar, dentre outras coisas, sugestões de temas de pesquisa, discussão de projetos de pesquisa e retorno dos seus resultados para indivíduos, grupos e populações aos quais as pesquisas se referem.

Figura 10: Depoimento do Relatório da pesquisa de percepções e experiências sobre compartilhamento de dados

“

*“Dialogar com a sociedade civil é uma das coisas que eu aprendi trabalhando, principalmente no ano passado, no enfrentamento à pandemia, com populações em situação de rua. Nunca falar ‘por’ eles e sim falar ‘com’ eles. Então, é sentar, é botar na mesa assim ‘ó, a gente, tem essa lei aqui de proteção de dados e a gente tá pensando isso, o que é que vocês pensam em relação às comunidades quilombolas?’. Eu acho que é a primeira coisa, entendeu?! Porque a gente fica muito por dentro de nós mesmos assim e elucubrando aqui uma coisa, e se você conversa com uma liderança dessas, ela já vai te dar uma noção. Então, eu acho que mesmo que a gente trabalhe com dados secundários, com dados administrativos, a gente tem que abrir o canal de diálogo com essas lideranças. Tanto que é uma coisa que eu fico falando toda hora... ‘tem que falar com as lideranças, tem que conversar com as lideranças’. Entendeu, abrir esse canal. Eu acho que o CIDADCS tem a possibilidade de abrir esses canais e não ficar só no meio científico e acadêmico. Porque a gente faz esse movimento muitas vezes quando a gente vai a campo, né?! Coleta dados primários. Quando a gente vai trabalhar com dados secundários, dados de domínio público, a gente não tem esse cuidado e que é extremamente relevante. Extremamente!”.*

*- Pesquisadora da área de Saúde (entrevistada para esta pesquisa)*

Fonte: Almeida e Pimenta, 2021

Pesquisadores também apontaram a falta de transparência e impessoalidade no processo de cessão de bases de dados administrativos, que devem ser pautados por critérios bem estabelecidos e transparentes, pautados pela isonomia.

Figura 11: Depoimento do Relatório da pesquisa de percepções e experiências sobre compartilhamento de dados

“

*“Olha, eu não acho que tenha muita transparência. Têm dados, o Brasil é um país ... não é super transparente... Acho que eu já tive esse problema mais de uma vez, mas eu tô lembrando agora de uma pesquisa de microcrédito que eu fiz e era uma avaliação de impacto de microcrédito. E o pessoal resolveu que não ia dar os dados, porque eram dados que você tinha que acessar quem recebeu o crédito, então você acabava tendo a... Naquela época, isso já faz uns... quase 20 anos, a gente conseguiu resolver conversando mesmo tudo com uma pessoa que tinha muito prestígio e sentando na mesa e explicando o quê que era e acabaram liberando. Mas, é o mesmo problema de prontuário né, a gente tem algumas coisas que a gente quer estudar, aí bate nesse problema de acesso a prontuário”.*

*- Pesquisadora da área de Ciências Sociais (entrevistada para esta pesquisa)*

Fonte: Almeida e Pimenta, 2021

Os gestores apresentaram a mesma preocupação que os pesquisadores entrevistados quanto à indispensabilidade de um sistema de governança de dados, capaz de propiciar transparência e acesso a boas informações para políticas públicas e para geração de novos conhecimentos e evidências em prol do interesse público, principalmente, dos mais vulneráveis.



Figura 12: Depoimento do Relatório da pesquisa de percepções e experiências sobre compartilhamento de dados

“

*“Na minha experiência é regulação... A Política Pública tem uma finalidade e a finalidade institucional vai ser melhorar a vida do cidadão e destes grupos e, pra isso, ela precisa do uso dos dados. Dificultar o acesso para políticas ou para pesquisa de instituições sérias de pesquisa é dificultar a própria capacidade de êxito da política. Fraude, a gente trata como fraude, erro a gente trata como erro, mas isso não tem que eivar todo um outro constructo que não pode nascer do receio do mau uso do dado, o que a gente tem que fazer é impedir o mau uso do dado, não impedir o bom uso do dado com receio do mau uso. Se você for numa família vulnerável, numa família indígena e perguntar : ‘você prefere ceder seu dado, ou, não ceder e não chegar, aqui, a água?’ Acho que quando vocês forem entrevistar as pessoas mais pobres, elas vão dizer: ‘não... então toma meu dado aí, então, traz o serviço, traz o benefício, faz aqui o que você puder fazer por mim.’ É isso, vamos regrar e impedir o mau uso, é o que pode ser feito. Punir o que há de errado e deixar as coisas seguirem, né?! Se não, a gente não faz mais nada. Se cada um, cada pesquisador e cada instituição pública for transformado em um potencial fraudador ou num potencial Serasa...”*

*- Gestora de Políticas Públicas (entrevistada para esta pesquisa)*

Fonte: Almeida e Pimenta, 2021

Destacamos ainda que dados administrativos são um tipo de *Big Data*, cujo valor não está pautado apenas no grande volume de dados e capacidade computacional, mas principalmente na capacidade de extrair conhecimento para que se tornem fonte de evidências.

Figura 13: Depoimento do Relatório da pesquisa de percepções e experiências sobre compartilhamento de dados

“

*“(...) vamos pensar numa etapa... por isso que é bom que tenha alguém, que não precisa ser de ciência da computação mas faça ciência de dados, sabe? Que use essas ferramentas e esses conhecimentos dessa grande área pra gerar e encontrar relações que ainda não foram encontradas. E veja; isso são relações que ainda não foram encontradas, não foram pensadas e que podem ser ainda mais eficientes pro tipo de processo de gestão de dado administrativo que a gente tem. Porque como é dado secundário, talvez a gente encontre coisas que o nosso processo de geração... nossa forma de pensar na população, na saúde populacional, talvez seja muito aquém das possibilidades que eu tenho, sabe? Só estou tentando dizer que... (...) Agora imagina a potência que é você lidar com grandes bases de dados pensando dessa forma. São relações que você consegue olhar de uma forma muito mais macro e relações que ao serem encontradas você vai pensar: ‘pô, como assim?’ Olha a grande sacada! Os epidemiologistas conseguiram descobrir que o fato de você ter poucos cômodos em casa ou dos seus cômodos serem muito pequenos, isso está muito relacionado, fortemente relacionado à incidência de tuberculose, sabe? Maravilha! Olha essa relação! Isso partiu de uma hipótese. Ninguém viu isso no dado. Alguém pensou nisso a priori. Olha como o ser humano é criativo a ponto de conseguir compreender essa relação a partir de observação de fenômeno. Agora imagina a possibilidade que é fazer com que a máquina teste todas as possibilidades possíveis dentro do dado e que você navegue entre as hipóteses que fazem mais sentido pro espírito do tempo, sabe?”*

*Pesquisador da área de Ciência de Dados (entrevistado para esta pesquisa)*

Fonte: Almeida e Pimenta, 2021

A governança de dados inclui a gestão de dados, com informações sobre a coleta, processamento, seleção de informações e qualidade dos dados, por exemplo, imprescindíveis para proporcionar contexto para as análises. Nesse sentido, os dados também precisam ser pensados em termos epistemológicos, pois são utilizados para representar diferentes aspectos da realidade, cujas interpretações dependerão das circunstâncias específicas dos processos de análise e dos significados que lhes são atribuídos.



Figuras 14 e 15: Depoimentos do Relatório da pesquisa de percepções e experiências sobre compartilhamento de dados

“

*“Então eu sempre... eu digo que a genética não é nada sem epidemiologia e epidemiologia não é nada se ela não discutir a antropologia, as culturas”.*

*- Pesquisadora Indígena da Área da Saúde-Antropologia (Geneticista) (entrevistada para esta pesquisa)*

“

*“E se a gente não entende esse processo, se a gente não traz pra discussão todas as possibilidades, tipo; a pessoa de epidemiologia vai lá e explica, ok, massa, mas a pessoa de gestão explica também. Traz a pessoa de computação, olha pro processo e também explica. E eu acho que essas três explicações, elas não são, nem contraditórias e também não são antagônicas, eu acho que elas são complementares. E é daí que nasce um estudo muito mais, eu diria, muito mais, como é que eu posso dizer? Multidisciplinar”.*

*- Pesquisador da área de Ciência da Computação (entrevistado para esta pesquisa)*

Fonte: Almeida e Pimenta, 2021

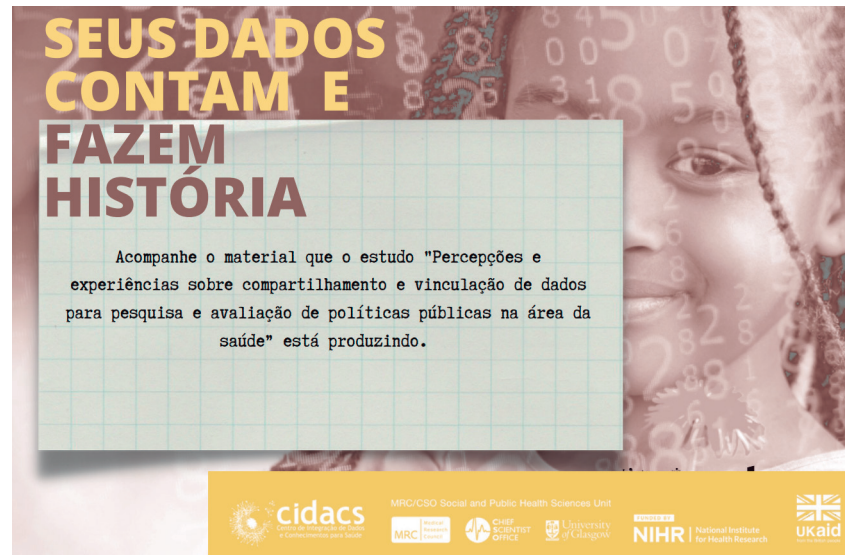
## Considerações

A despeito do limitado potencial de generalização do estudo, em virtude do tamanho da amostra e de sua representatividade, os achados da pesquisa são indicativos importantes que apontam para a necessidade de ações voltadas à literacia sobre dados em um mundo digitalizado:

- criação de mecanismos de envolvimento e engajamento da sociedade civil nas decisões sobre os usos e reúsos de seus dados, particularmente de dados custodiados pelo Estado;
- estabelecimento de um ecossistema nacional de gestão e governança de dados administrativos para subsidiar a geração de conhecimento científico e de evidências para políticas públicas, que inclua a vinculação de dados, com mecanismos de transparência, controle e responsabilização orientados pelo legítimo interesse público.

## SEUS DADOS CONTAM E FAZEM HISTÓRIA

Figura 16: Material do projeto “Seus dados contam e fazem história”



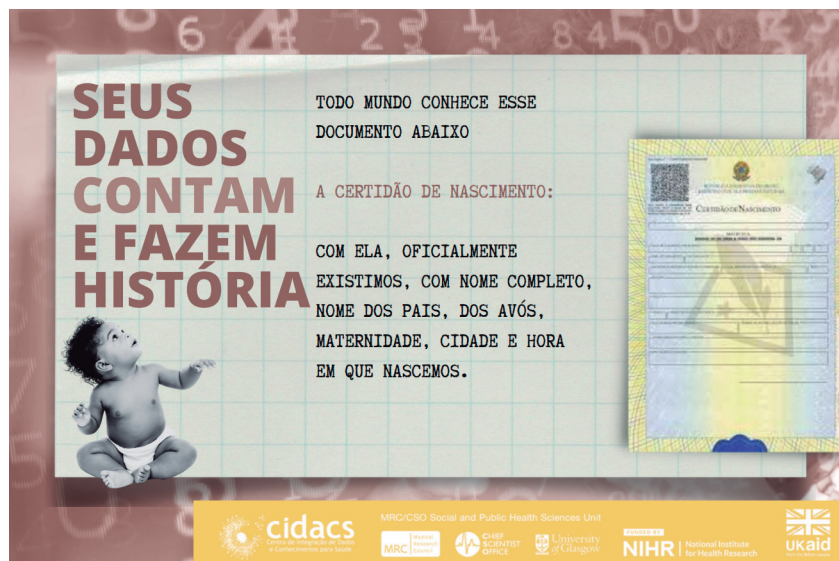
Fonte: <https://cidacs.bahia.fiocruz.br/>

A partir dos achados da pesquisa intitulada “Percepções e experiências sobre compartilhamento e vinculação de dados para pesquisa e avaliação de políticas públicas na área da saúde”, elaboramos conteúdos com informações sucintas, de fácil acesso, mas sem perder a profundidade para alcançar diversos públicos. O material produzido foi disseminado semanalmente por dois meses nas redes sociais do Cidacs e está disponível em seu *site*, de forma aberta, para ampla utilização<sup>18</sup>.

As postagens, além de informar sobre definições e conceitos, apresentaram exemplos e cenários de usos e reúsos de dados para pesquisa científica e para políticas públicas, destacando o fato de que produzimos dados administrativos desde o nosso nascimento.

18. O material do projeto está disponível no link <https://cidacs.bahia.fiocruz.br/2021/11/24/material-do-projeto-seus-dados-contam-e-fazem-historia/>

Figura 17: Material do projeto “Seus dados contam e fazem história”



Fonte: <https://cidacs.bahia.fiocruz.br/>

Dentre os conteúdos, buscou-se responder a perguntas como: O que são dados primários e dados secundários? O que são dados pessoais e dados sensíveis? Quais são os aspectos éticos do uso de dados pessoais para pesquisa científica? Qual é a importância dos dados administrativos para a ciência e para gestão pública? O que é LGPD? O que é *linkage*?

Figuras 18 e 19: Material do projeto “Seus dados contam e fazem história”

**SEUS DADOS CONTAM E FAZEM HISTÓRIA**

E OS DADOS OBTIDOS PELO GOVERNO QUANDO A GENTE PREENCHE UM CADASTRO, FICHA OU FORMULÁRIO SÃO CHAMADOS DE DADOS ADMINISTRATIVOS, POR EXEMPLO PARA MATRÍCULA NA ESCOLA, ATENDIMENTO MÉDICO, IMPOSTO DE RENDA OU BENEFÍCIO SOCIAL.

**cidacs** MRC/CSO Social and Public Health Sciences Unit  
 MRC CHEF SCIENTIST OFFICE University of Glasgow NIHR National Institute for Health Research ukaid

**SEUS DADOS CONTAM E FAZEM HISTÓRIA**

VOCÊ JÁ PAROU PARA PENSAR SOBRE A IMPORTÂNCIA DOS DADOS ADMINISTRATIVOS PARA O GOVERNO E PARA POPULAÇÃO? E QUE ESTES DADOS ALIMENTAM DIFERENTES SISTEMAS DE INFORMAÇÃO DO MUNICÍPIO, DO ESTADO E DO PAÍS?

QUAL O TAMANHO DA POPULAÇÃO? TEM MAIS MULHERES?  
 HOMENS? QUANTOS NASCIMENTOS? QUANTAS MORTES?  
 QUANTAS HOSPITALIZAÇÕES?

TUDO ISSO, OS DADOS ADMINISTRATIVOS PODEM NOS AJUDAR A RESPONDER.

**cidacs** MRC/CSO Social and Public Health Sciences Unit  
 MRC CHEF SCIENTIST OFFICE University of Glasgow NIHR National Institute for Health Research ukaid

Fonte: <https://cidacs.bahia.fiocruz.br/>

O propósito foi auxiliar no entendimento sobre direitos dos titulares de dados pessoais, como também acerca da importância dos dados administrativos para promoção de políticas públicas e realização de pesquisas científicas.

## **COMPARTILHAMENTO E ABERTURA DE DADOS NA CIÊNCIA ABERTA E A PESQUISA EM SAÚDE PÚBLICA NA LEI GERAL DE PROTEÇÃO DE DADOS**

Dados podem ser utilizados para diferentes propósitos, e a depender da situação e dos interesses envolvidos, serão utilizados como mercadorias (*commodities*), ou como bens públicos. Assim, o aumento da quantidade de dados produzidos por distintas fontes, dentro e fora do âmbito científico e em um mundo cada vez mais digitalizado, suscita inúmeras reflexões sobre as condições e finalidades de acesso para o (re) uso de dados.

Dados são imprescindíveis para geração de conhecimento científico e de evidências para políticas e programas dos governos, que incluem o Sistema Único de Saúde (SUS). Adicionalmente, a pesquisa em saúde tradicionalmente utiliza dados pessoais e sensíveis a partir de normas específicas de proteção aos direitos dos titulares dos dados. Nessa perspectiva, analisamos as principais orientações internacionais voltadas ao compartilhamento e abertura de dados para pesquisa na Ciência Aberta e discutimos as especificidades da pesquisa científica e de estudos em saúde pública na Lei Geral de Proteção de Dados (LGPD)<sup>19</sup>.

---

19. Consulte o artigo publicado sobre o tema em ALMEIDA; PIMENTA; BARRETO, 2021.



## O que achamos?

Como dados de diferentes tipos e origens passaram a ser cada vez mais utilizados e reutilizados com distintos propósitos e interesses, discussões éticas e legais em torno da coleta e processamento de dados e as finalidades de uso, principalmente de dados pessoais e sensíveis, tornaram-se frequentes.

Ao considerar que o compartilhamento e abertura de dados com finalidade de pesquisa manifesta-se de maneira diferente nas diversas áreas disciplinares e países, devendo seguir princípios éticos e marcos regulatórios, abordamos aspectos relacionados ao processamento de dados para pesquisa científica no Brasil, particularmente dados pessoais e sensíveis, e as especificidades da pesquisa em saúde pública no âmbito da Lei Geral de Proteção de Dados – Lei nº 13.709.

No Brasil, toda pesquisa envolvendo seres humanos requer aprovação do Sistema CEP/Conep, com o objetivo de proteger o participante da pesquisa e assegurar que o estudo será realizado de acordo com princípios éticos a partir de resoluções e normativas deliberadas pelo Conselho Nacional de Saúde.

A LGPD, sancionada em agosto de 2018 e com vigência a partir de agosto de 2020, volta-se ao estabelecimento de princípios e conceitos norteadores para preservar o equilíbrio entre a necessidade de proteger efetivamente os direitos dos titulares dos dados, ao mesmo tempo em que permite o processamento de dados pessoais e sensíveis para determinadas finalidades.

Verifica-se que a pesquisa científica e a pesquisa em saúde pública são consideradas contextos específicos de processamento de dados na Lei, desde que existam salvaguardas para proteger os direitos dos titulares dos dados, que deverão estar em consonância com a regulamentação do setor.

No que tange à gestão de dados, considerada uma condição indispensável para a reutilização de dados de projetos de pesquisa na perspectiva da Ciência Aberta, cabe sa-

liantar que esta também é requerida na LGPD, uma vez que a lei exige que os responsáveis pelo tratamento dos dados pessoais organizem e mantenham registros acerca de qualquer atividade relacionada ao processamento desses dados pessoais sob sua responsabilidade, pois os titulares devem ter acesso facilitado às informações sobre qualquer tratamento aos quais seus dados sejam submetidos.

Novas formas de produzir conhecimento científico com os adventos do *Big Data* e da abertura de dados na perspectiva da Ciência Aberta colocam desafios que precisam ser enfrentados, dentre eles a harmonização do marco regulatório nacional para utilização de dados pessoais e sensíveis para produção de conhecimento científico com segurança jurídica, transparência e responsabilidade social. Requerendo, portanto, infraestruturas computacionais robustas e sustentáveis, ferramentas, serviços e formação de pessoal para gestão e análise de dados.

### **Considerações**

Verifica-se que ainda não existe consenso acerca do acesso e gestão de dados secundários por instituições de pesquisa, que incluem dados administrativos e outros tipos de dados oriundos de distintas fontes para pesquisa científica e para estudos em saúde pública. As principais diretrizes encontradas na literatura e nas experiências internacionais sobre o tema, que incluem integração de dados, são voltadas ao estabelecimento claro de acesso aos dados, que envolve medidas de segurança apropriadas, uso dos dados para finalidades específicas, credenciais do solicitante e aprovação ética do estudo proposto.

No Brasil, temos o exemplo do Centro de Integração de Dados e Conhecimentos para a Saúde (Cidacs), vinculado à Fiocruz Bahia, que estruturou um complexo sistema de gerenciamento de dados com métodos, técnicas e protocolos para receber, tratar, processar, integrar e preservar grande volume de dados. O provimento de acesso aos dados vinculados e anonimizados para responder perguntas de pesquisas relaciona-

das aos determinantes sociais e aos impactos de políticas públicas sobre a saúde da população brasileira, é fornecido sob determinados termos e condições em ambiente seguro e controlado.

## CONCLUSÕES

O trabalho teve início com um mapeamento e análise documental de artigos científicos, documentos oficiais, sítios eletrônicos de instituições governamentais e de fomento à pesquisa e iniciativas sobre a abertura de dados praticadas em outros países. Em seguida, buscamos conhecer os marcos legais nacionais que incidem na pesquisa científica e mais especificamente na pesquisa em saúde, para orientar o estabelecimento de princípios e diretrizes voltados à gestão, compartilhamento e abertura de dados para pesquisa. Esses estudos subsidiaram ações no âmbito da Plataforma Zika e o estabelecimento da política institucional da Fiocruz, que poderão ser utilizados como orientações pela Secretaria de Vigilância em Saúde e Ambiente (SVSA/MS) e outras instituições que custodiam dados utilizados para pesquisa científica.

Ao longo do processo, constatou-se que as discussões sobre a reutilização de dados na perspectiva da Ciência Aberta tinham por foco dados gerados por projetos de pesquisa. Para suprir lacunas sobre o acesso e integração de dados administrativos para pesquisa, tipos de dados que subsidiaram os estudos da Plataforma de Vigilância de Longo Prazo para Zika e Microcefalia no âmbito do SUS, iniciamos uma nova fase de trabalho.

A utilização de dados administrativos para pesquisa científica e estudos em saúde pública apresenta uma série de desafios, pois seu acesso, tratamento, compartilhamento e utilização apresentam questionamentos éticos e legais que necessitam ser respondidos através de um conjunto de medidas, que visam mitigar riscos e ampliar benefícios para indivíduos e sociedade.



Ao reconhecermos que a Ciência Aberta requer transparência e participação social para que seja responsável e inclusiva, buscamos compreender as experiências e percepções de representantes de grupos de pacientes, beneficiários de programas sociais, pesquisadores usuários de dados administrativos e gestores públicos, acerca do compartilhamento e integração de dados com finalidade de pesquisa e de estudos em saúde pública.

Iniciamos assim, processos de literacia voltados a diferentes segmentos interessados (governo, academia, organizações da sociedade civil e público em geral), e elaboramos e disseminamos conteúdos sobre usos e reúsos de dados com finalidades de pesquisa e de políticas públicas, todos disponíveis em formato aberto para ampla utilização.

E como dados são imprescindíveis para geração de conhecimento científico e de evidências para políticas e programas dos governos, que incluem o SUS, buscamos discutir alguns aspectos relacionados à Ciência Aberta, *Big Data* e especificidades da pesquisa científica e de estudos em saúde pública à luz da LGPD.

O trabalho desenvolvido, intitulado Dados e Ciência Aberta, aponta uma série de desafios a serem enfrentados para produção de conhecimento científico confiável e socialmente responsável na era do *Big Data*, principalmente quando são utilizados dados administrativos integrados para responder perguntas complexas.

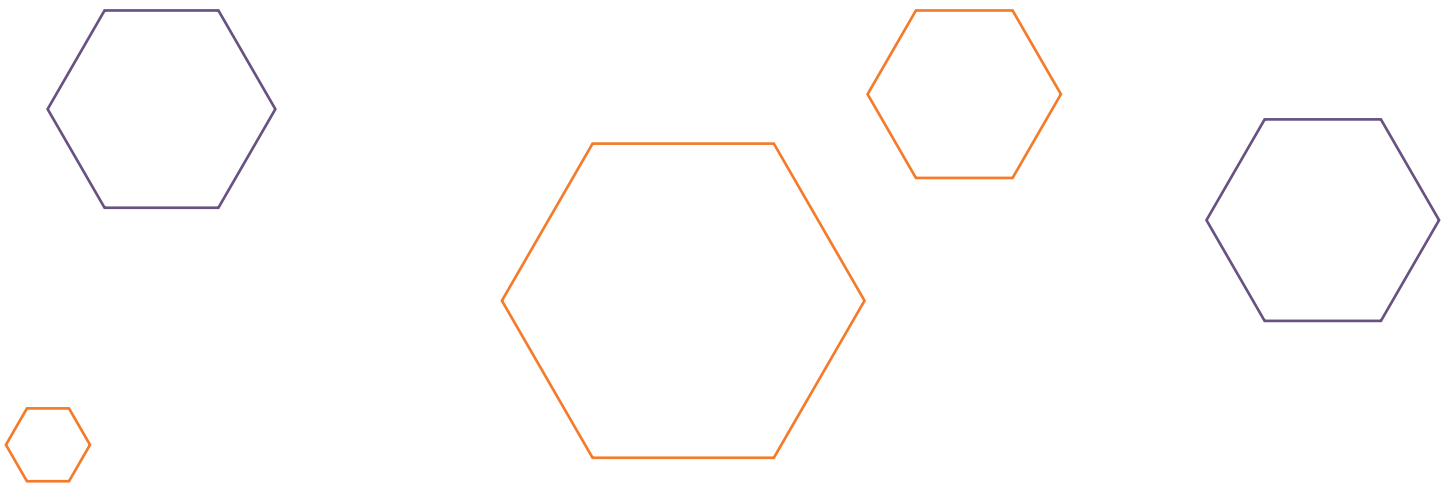
Esses desafios foram pensados de forma pioneira na estruturação da Plataforma Zika, pois o projeto propiciou arranjos de segurança da informação e pessoal especializado para receber, processar, integrar e prover acesso aos dados administrativos integrados, com qualidade e acurácia. Dados integrados que viabilizaram a condução dos estudos epidemiológicos.

Adicionalmente, o projeto viabilizou conhecer as especificidades da comunidade científica que estuda Zika no Brasil e a participação de diversos segmentos da sociedade nas edições da Feira de Soluções para Saúde, que reuniu instituições acadêmicas e de

pesquisa, empresas, organismos internacionais, instituições governamentais e sociedade civil, particularmente grupos sociais atingidos pela epidemia de Zika no Brasil.

Especificamente no que tange ao trabalho voltado à gestão, compartilhamento e abertura de dados na perspectiva da Ciência Aberta, destacamos que é uma realidade, e que o objetivo precípua do grupo de trabalho que se dedicou ao estudo desse tema, por meio dos projetos e iniciativas descritas neste relatório, foi fornecer subsídios, sob diferentes ângulos e abordagens, para que gestores, pesquisadores, autoridades e órgãos responsáveis pela ciência, tecnologia e inovação pudessem se apropriar, criticamente, dos fundamentos, conceitos e estratégias utilizadas em âmbitos internacional e nacional.

Esperamos que o trabalho iniciado na Plataforma Zika auxilie na mobilização e adesão de atores nacionais às práticas de gestão e integração de dados na perspectiva da Ciência Aberta, particularmente no setor de pesquisa em saúde, para aumentar a velocidade de circulação da informação e o reuso de dados voltados ao desenvolvimento científico e geração de evidências para políticas públicas em prol da melhoria das condições de vida e de saúde da população brasileira.



## ARTIGOS CIENTÍFICOS E PRODUTOS RESULTANTES

ALMEIDA, B; PIMENTA, D; BARRETO, M. Compartilhamento e abertura de dados na Ciência Aberta e a pesquisa em saúde pública na Lei Geral de Proteção de Dados. **Revista Bahia anál. dados**, Salvador, v. 31, n. 2, p.78-92, jul.-dez. 2021..

ALMEIDA, B; PIMENTA, D. **Percepções e experiências sobre compartilhamento e vinculação de dados para pesquisa e avaliação de políticas públicas na área da saúde**. Salvador, Cidacs, Fiocruz, 2021. Disponível em: <https://cidacs.bahia.fiocruz.br/2021/11/22/percepcoes-e-experiencias-sobre-compartilhamento-e-vinculacao-de-dados-para-pesquisa-e-avaliacao-de-politicas-publicas-na-area-da-saude/>. Acesso em set de 2022.

ALMEIDA, B; PIMENTA, D. **Seus Dados Contam e Fazem História**. 2021. Disponível em <https://cidacs.bahia.fiocruz.br/2021/11/24/material-do-projeto-seus-dados-contam-e-fazem-historia/>. Acesso em set de 2022.

FUNDAÇÃO OSWALDO CRUZ. **Formação Modular sobre Ciência Aberta**. Campus Virtual Fiocruz, 2018. Disponível em: <https://campusvirtual.fiocruz.br/portal/?q=formacao-modular/ciencia-aberta>. Acesso em set de 2022.

FUNDAÇÃO OSWALDO CRUZ. Presidência. Vice-Presidência de Educação, Informação e Comunicação. Grupo de Trabalho em Ciência Aberta. **Termo de referência: gestão e abertura de dados para pesquisa na Fiocruz**. Rio de Janeiro, 2019. 15 p. Disponível em: [https://www.arca.fiocruz.br/bitstream/handle/iciict/26803/TERMO\\_REFERENCIA\\_LANCAMENTO.pdf?sequence=2](https://www.arca.fiocruz.br/bitstream/handle/iciict/26803/TERMO_REFERENCIA_LANCAMENTO.pdf?sequence=2). Acesso em set de 2022.

FUNDAÇÃO OSWALDO CRUZ. Vice-Presidência de Educação, Informação e Comunicação. **Política de gestão, compartilhamento e abertura de dados para pesquisa: princípios e diretrizes**. - Rio de Janeiro: Fiocruz/VPEIC, 2020. 19 p. Disponível em: [https://www.arca.fiocruz.br/bitstream/handle/iciict/46408/VPEIC\\_versao\\_PORTUGUES\\_2021-03-22.pdf?sequence=2&isAllowed=y](https://www.arca.fiocruz.br/bitstream/handle/iciict/46408/VPEIC_versao_PORTUGUES_2021-03-22.pdf?sequence=2&isAllowed=y). Acesso em set de 2022.

GUANAES, P. *et al.* **Marcos legais nacionais em face da abertura de dados para pesquisa em saúde:** dados pessoais, sensíveis ou sigilosos e propriedade intelectual. Rio de Janeiro: Fiocruz, 2018. 122 p.

SANTOS, P. X. *et al.* **Livro Verde** - Ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional. Rio de Janeiro: Fiocruz, 2017. 140 p.

## REFERÊNCIAS

BOAI. **Budapest Open Access Initiative.** 2002. Disponível em: <https://www.budapestopenaccessinitiative.org/read/>. Acesso em set de 2022.

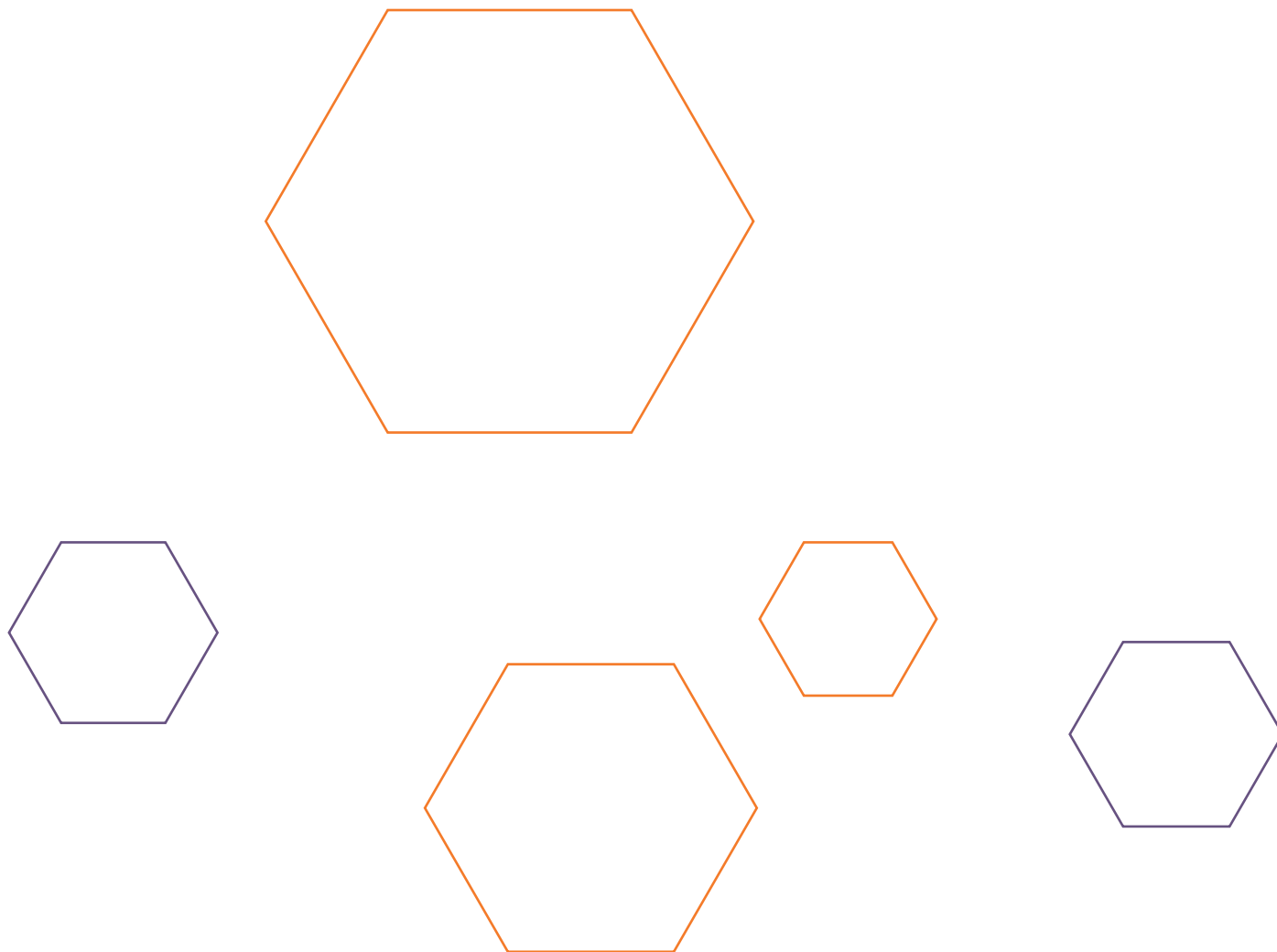
BRASIL. Lei nº 13.709, de 14 de agosto de 2018. Lei Geral de Proteção de Dados Pessoais (LGPD). **Diário Oficial da União, Brasília, DF**, 15 ago. 2018. Disponível em: [http://www.planalto.gov.br/ccivil\\_03/\\_ato2015-2018/2018/lei/L13709.htm](http://www.planalto.gov.br/ccivil_03/_ato2015-2018/2018/lei/L13709.htm). Acesso em set de 2022.

FOSTER. **Open Science Taxonomy.** 2016. Disponível em: <https://www.fosteropenscience.eu/foster-taxonomy/open-data?page=1>. Acesso em set de 2022.

OECD. **Principles and Guidelines for Access to Research Data from Public Funding.** 2007. Disponível em: <https://www.oecd.org/sti/inno/38500813.pdf>. Acesso em set de 2022.

ROYAL SOCIETY. **Data for international health emergencies: governance, operations and skills.** London: Royal Society, 2021. Disponível em: <https://royalsociety.org/-/media/about-us/international/g-science-statements/G7-data-for-international-health-emergencies-31-03-2021.pdf>. Acesso em set de 2022.

UNESCO. **Recomendação da UNESCO sobre Ciência Aberta**. 2021. Disponível em português: [https://unesdoc.unesco.org/ark:/48223/pf0000379949\\_por](https://unesdoc.unesco.org/ark:/48223/pf0000379949_por). Acesso em set de 2022.



# Considerações Finais

---



## Considerações Finais

Um surto, epidemia ou pandemia são eventos socialmente complexos e apesar de ocorrerem com crescente frequência, quer seja de forma restrita, em diferentes partes do globo (surtos ou epidemias), ou de maneira generalizada (pandemias), os fatores que determinam a emergência de tais eventos, os seus efeitos na saúde e nas mais diversas esferas da vida social e as estratégias que os diversos setores da sociedade utilizam-se para enfrentá-los, carece de esclarecimentos. Portanto, a análise em profundidade de cada surto, epidemia ou pandemia ao ser feito, constitui-se em uma imensa fonte geradora de conhecimentos sobre os seus determinantes, efeitos e os meios de mitigação, essenciais para atuar em um evento em curso, ou eventos similares no futuro.

O aprendizado gerado deve ser disseminado e partilhado de maneira mais ampla possível para que nos tornemos mais qualificadas na prevenção ou mitigação de eventos similares, que certamente ocorrerão. Enfrentar estes desafios exigem dados e equipes de investigadores provenientes de diferentes disciplinas, com capacidade para enquadrar o evento a partir de diversos ângulos e perspectivas, permitindo uma compreensão mais holística do problema e, também, na coordenação das respostas.

Nesta publicação, são relatados os estudos e investigações realizados por um diversificado grupo de pesquisadores focados na devastadora epidemia de Zika, que ocorreu no Brasil entre 2015 e 2016, porém cujos efeitos nas vidas das crianças e famílias acometidas ainda perduram. Devendo-se ainda lembrar que casos de Zika continuam a ocorrer e uma nova epidemia não pode ser descartada.

O processo de investigação tornou possível produzir uma série de conhecimentos relevantes sobre um conjunto diverso de questões, desde os desafios para governança e disponibilização de dados de diversos sistemas nacionais de informação, fundados nos princípios do movimento da Ciência Aberta, até a utilização destes dados para avaliar o impacto da epidemia no registro, incidência e mortalidade pela síndrome congênita do vírus Zika (SCZ). Graças a utilização de dados dos sistemas nacionais de

informação foi possível avaliar diferenças nas estimativas de risco entre municípios brasileiros, desfechos adversos e marcadores sociais associados aos recém-nascidos com a SCZ.

Para viabilizar a realização dos estudos longitudinais sobre a saúde das crianças de mães infectadas pelo vírus da Zika a partir dos dados integrados de diferentes sistemas nacionais, foi necessário criar infraestrutura técnica, operacional e metodológica para produção de *datasets* com qualidade e acurácia para responder as questões de investigação. Para viabilizar o recebimento, processamento, integração, preservação e acesso dos dados, em conformidade com os marcos éticos e legais vigentes no país, intensificou-se no Centro de Integração de Dados e Conhecimentos para Saúde (Cidacs/Fiocruz Bahia) o desenvolvimento de um complexo sistema para governança, processamento e análise destes dados. A experiência mostrou a quantidade e qualidade dos novos conhecimentos e tecnologias que puderam ser produzidos a partir da integração de dados rotineiramente coletados, com o esclarecimento de importantes aspectos da dinâmica da epidemia e de seus efeitos sobre a saúde da população.

Ao considerar os desafios científicos para compreensão da epidemia de Zika, processo que muito rapidamente envolveu pesquisadores, laboratórios e instituições do Brasil e de diferentes países, buscou-se contribuir para o entendimento da dinâmica dos arranjos e colaborações científicas nacionais e internacionais necessária para cobrir as diferentes questões que vão dos aspectos clínicos e epidemiológicos aos aspectos moleculares envolvidas na resposta ao vírus. Utilizando-se de diferentes bases de dados bibliométricas e métodos e abordagens cientométricas, foram mapeados áreas de conhecimento, pesquisadores, instituições e países. Exploração que possibilitou entender o ativo processo de articulação e colaborações da comunidade científica para geração de conhecimentos fundamentais na busca por respostas aos desafios colocados pelo vírus e a epidemia. Nesta jornada gerou-se também instrumentos que colaborassem com os gestores na seleção de conhecimentos e extração de competências relevantes nas bases abertas existentes, como o currículo Lattes.



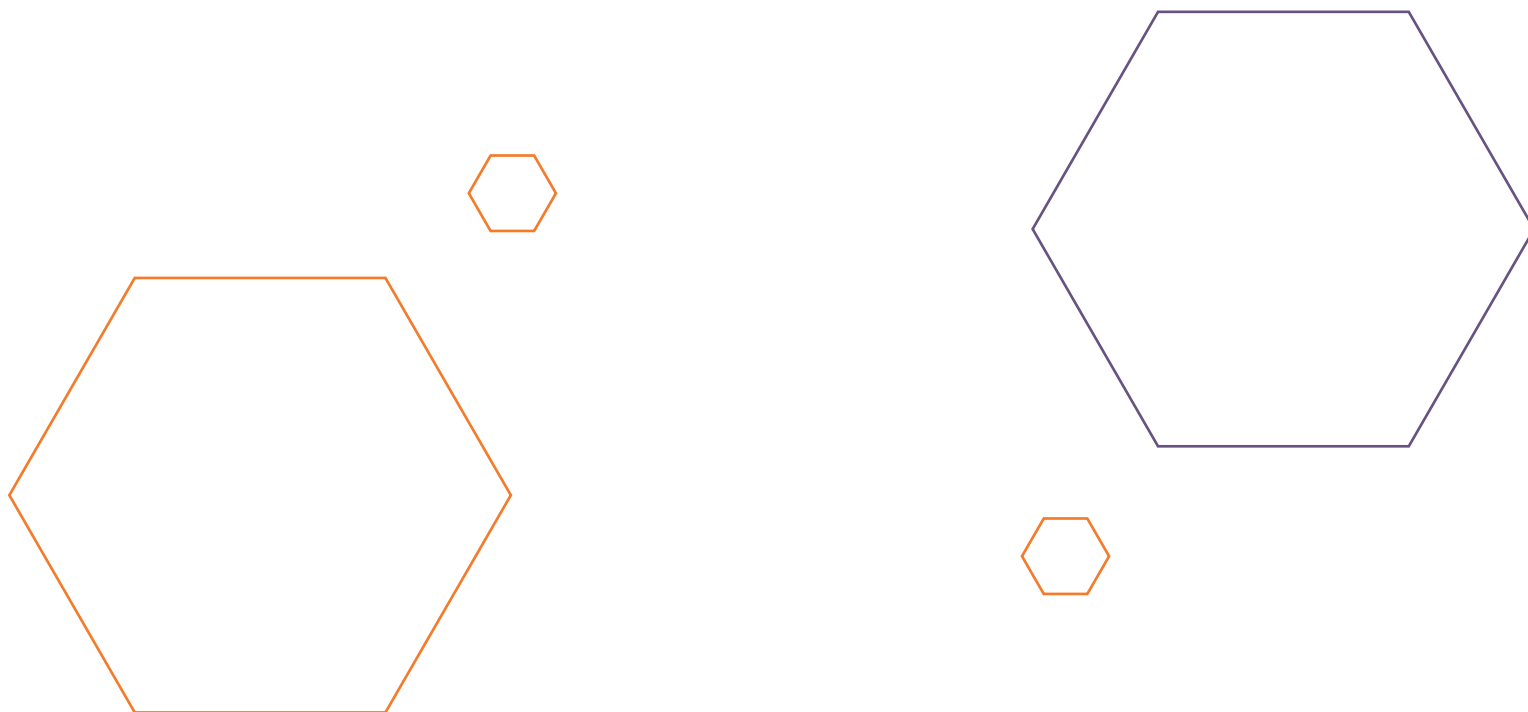
Em outro enfoque, buscou-se entender as formas que diversos outros atores, particularmente a sociedade civil, organizaram-se para lidar com os efeitos da epidemia. propor e gerar soluções. A epidemia da Zika mostrou a importância da mobilização não somente da comunidade científica, mas também da sociedade civil na elaboração e implementação ações em diversas esferas, visando mitigar os seus efeitos. A série de feiras de soluções, que teve início com uma grande feira em Salvador, focada em Zika, foi uma imensa oportunidade de conhecer e catalizar as ações e estratégias para prevenção, tratamento e cuidado, que emanam tanto da comunidade científica, mas também da sociedade civil. Esta iniciativa possibilitou o registro, a catalogação e a análise de mais de duas centenas de ações originadas na sociedade civil, demonstrando a importância da participação social em todo o processo da epidemia e os caminhos apontados para além da academia e do governo.

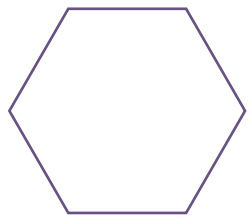
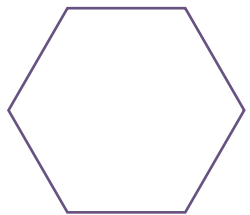
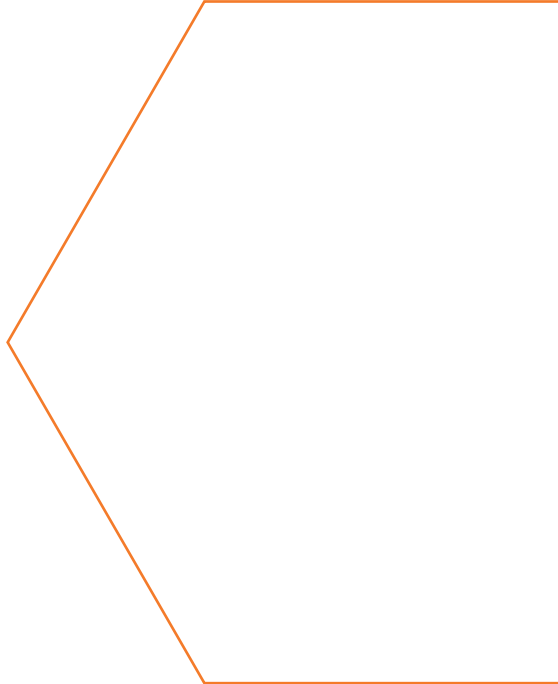
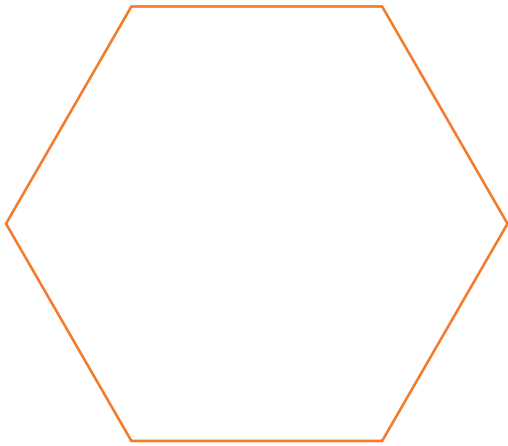
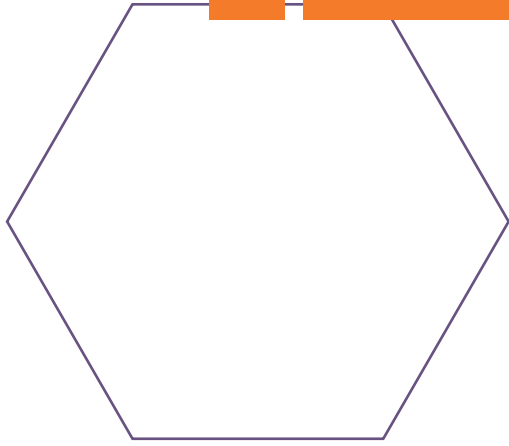
Desafio comum de todas as epidemias ou pandemias, em período recente, tem sido a utilização rápida de dados científicos e, principalmente, aqueles rotineiramente produzidos pelo sistema de saúde ou outros sistemas de registro de informações sobre as atividades humanas. A rapidez entre a produção e o uso dos dados é essencial para o estabelecimento o mais imediato possível das causas e das consequências da epidemia. Dados que precisam ser disponibilizados de forma muito rápida para responder as emergências de saúde pública, tal como foi a epidemia de Zika. Nessa perspectiva, a gestão e governança de dados precisa de investimentos contínuos em sistemas de tecnologia e informação (*hardware, software*), rapidez na coleta e atualização das bases de dados e no treinamento de recursos humanos especializados em processar, armazenar, preservar e dar acesso aos dados de forma ágil, ética e socialmente responsável.

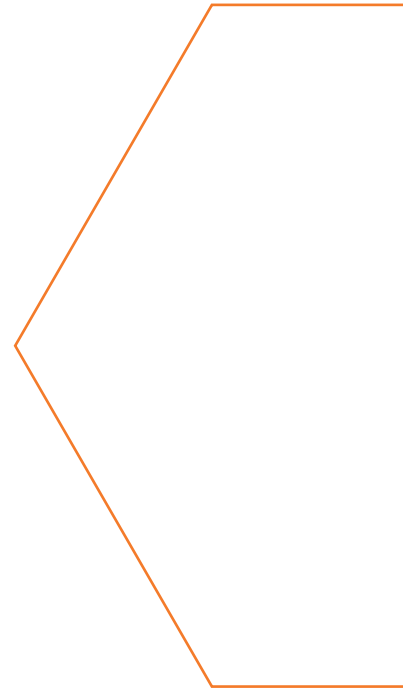
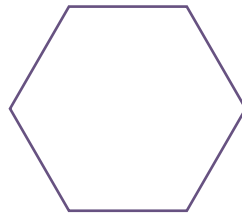
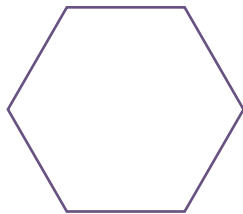
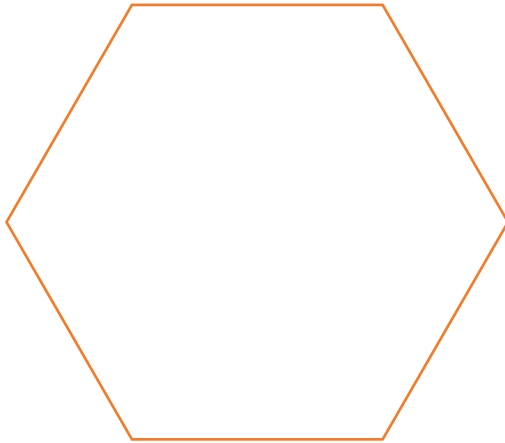
Por fim, interessava ao grupo refletir e contribuir com o debate necessário sobre a abertura dos dados e conhecimentos gerados. Isto passou por estudos sobre infraestrutura, diretrizes e políticas relacionadas ao compartilhamento de dados para pesquisa em diversos países, e acerca dos marcos legais e éticos para o compartilhamento e abertura de dados na pesquisa em saúde no Brasil. Estudos essenciais e que fundamentaram o desenvolvimento da política de gestão, compartilhamento e abertura de

dados para pesquisa no âmbito da Fiocruz. Além do acesso ao conhecimento e aos dados para pesquisa em saúde, a perspectiva de ciência aberta adotada visa que esta seja socialmente responsável e inclusiva. Portanto, disseminar experiências, fomentar discussões e incorporar novas perspectivas, que incluam a sociedade civil, fazem parte do processo iniciado na Plataforma Zika.

A síntese do processo e resultados apresentados nesta publicação refletem a missão do Centro de Integração de Dados e Conhecimentos para a Saúde (Cidacs/Fiocruz Bahia), que se especializou na governança, integração e análise de grandes bases de dados nacionais para responder à perguntas de investigação relevantes relacionadas aos problemas de saúde da população brasileira e ao Sistema Único de Saúde (SUS). Desde a sua concepção e ao longo da sua existência, iniciada em dezembro de 2016, o Cidacs busca responder aos desafios técnicos, científicos e sociais que se apresentam no cumprimento da sua missão.







As imagens de capa utilizadas nesse livro possuem licença livre.

**Foto 1:** Marilene abraça seu filho João, que nasceu com microcefalia.

**Autor:** Felipe Fittipaldi/ *Wellcome Photography Prize* 2019

**Repositório Disponível:** Wikimedia Commons - [https://commons.wikimedia.org/wiki/File:Hold\\_Me\\_Mother\\_2018\\_-\\_Wellcome\\_Photoqraphy\\_Prize\\_2019.jpg](https://commons.wikimedia.org/wiki/File:Hold_Me_Mother_2018_-_Wellcome_Photoqraphy_Prize_2019.jpg)

**Fotos 2 a 6:** Bancos de Imagens

Agência Brasil, *Creative Commons Attribution 4.0 International*, Envato Elements, Banco de Imagens Fiocruz e Freepik.

**Foto 7:** Bebê com microcefalia em sessão de fisioterapia

**Autor:** Sumaia Villela/Agência Brasil

**Repositório Disponível:** *Wikimedia Commons* - [https://commons.wikimedia.org/wiki/File:Beb%C3%AA\\_com\\_microcefalia\\_01.jpg](https://commons.wikimedia.org/wiki/File:Beb%C3%AA_com_microcefalia_01.jpg)



[cidacs.bahia.fiocruz.br/plataformazika](http://cidacs.bahia.fiocruz.br/plataformazika) | [cidacs.comunicacao@fiocruz.br](mailto:cidacs.comunicacao@fiocruz.br)  
 Rua Mundo, nº 121, sala 315 - Parque Tecnológico da Bahia - Trobogy, Salvador/BA



MINISTÉRIO DA SAÚDE

GOVERNO FEDERAL  
**BRASIL**  
 UNIÃO E RECONSTRUÇÃO

