

INCOMPLETUDE DE DADOS EM REGISTROS DE SRAG/COVID-19 E IMPLICAÇÕES PARA O MONITORAMENTO DA AGENDA 2030: EVIDÊNCIAS A PARTIR DO MORRO DA KIBON (SANTO ANDRÉ-SP)

Matheus da Silva, Bacharelado em Estatística

INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E CIÊNCIA DA COMPUTAÇÃO DA
UNIVERSIDADE DE SÃO PAULO, BRASIL

matheus.ssilva@usp.br - <https://orcid.org/0009-0009-4810-0039>

Eduardo Magalhães Rodrigues

CENTRO DE ESTUDOS DE SAÚDE COLETIVA DA FACULDADE DE MEDICINA DO
ABC, CESCO/ABC, SÃO PAULO, BRASIL

pesquisadorcesco@gmail.com - <https://orcid.org/0009-0006-4850-7371>

RESUMO

A incompletude de registros em sistemas de informação em saúde compromete diagnósticos e pode distorcer indicadores usados na gestão pública e no monitoramento da Agenda 2030. Este estudo analisou a incompletude de variáveis sociodemográficas em registros de SRAG/COVID-19 de 2020 a 2023, tomando o município de Santo André como referência e utilizando o bairro Condomínio Maracanã como recorte aproximado de um território socialmente reconhecido, porém sem delimitação administrativa formal (Morro da Kibon). Foram utilizados dados secundários de bases públicas e dados agregados obtidos via Lei de Acesso à Informação. A incompletude foi definida como a proporção de registros marcados como “Ignorado/Não informado” nas variáveis raça/cor e escolaridade (além de sexo e evolução, para comparação), estimada por ano e sintetizada em heatmap. Observou-se incompletude persistentemente mais elevada em escolaridade (aprox. 46%–72%) e relevante em raça/cor (aprox. 5%–20%), em contraste com valores próximos de zero em variáveis clínicas/administrativas. Além disso, a ausência de reconhecimento territorial inviabiliza a identificação direta do Morro da Kibon, produzindo subcobertura e limitando inferências locais. Conclui-se que lacunas em variáveis sociais e a invisibilidade administrativa do território podem introduzir vieses e fragilizar diagnósticos de desigualdades, com implicações para ações de saúde e para o acompanhamento de metas relacionadas a saúde e equidade na Agenda 2030.

PALAVRES-CHAVES: INCOMPLETUDE, INVISIBILIDADE, INDICADORES, TERRITÓRIO, COVID-19

INTRODUÇÃO

A produção e o uso de dados tornaram-se elementos estruturantes da gestão pública atualmente, sobretudo quando decisões precisam ser orientadas por evidências e monitoradas por indicadores. A pandemia do novo coronavírus que durou entre 2020 e 2023 intensificou esse cenário ao tornar mais evidente a dependência de registros consistentes para a formulação de respostas na saúde e na organização de serviços para a saúde básica, ao mesmo tempo em que expôs vulnerabilidades da qualidade dos dados, como inconsistências e campos não preenchidos, também descritos como incompletude de dados (*missing data*).

A incompletude dos dados, isso é, a ocorrência de campos em branco ou marcar como “Não informado” tende a reduzir consideravelmente a capacidade dos bancos de dados que são trabalhados, onde esses dados podem introduzir vieses e conclusões incorretas sobre qualquer território proposto. Em um cenário de estudos numa região qualquer, isso pode levar a medidas ineficazes e que não condizem com a real necessidade dos habitantes que ali vivem.

A região de escolha para a realização deste trabalho, “Morro da Kibon”, em Santo André, é uma das candidatas a receber medidas ineficazes em saúde pública devido à falta e omissão sofridas durante a coleta de dados nessa região. Além da limitação da “descrição de vida” dos habitantes dessas regiões, a incompletude dos dados amplia o enfraquecimento dos acompanhamentos das metas públicas vinculadas a Agenda 2030, cuja avaliações tendem a presumir de dados minimamente robustos e comparativos. Com isso, este trabalho tem como objetivo buscar articular a discussão sobre a coleta de dados no Morro da Kibon com as inferências para o monitoramento de indicadores relacionados aos Objetivos de Desenvolvimento Sustentável (ODS). Especificamente, pretende-se descrever a ocorrência e o padrão de incompletude/“Não informado” nos registros disponíveis para o território, discutir os efeitos potenciais dessa lacuna sobre diagnósticos e indicadores, e refletir sobre os desafios metodológicos de produzir evidência em um território socialmente reconhecido, porém com baixa visibilidade na esfera municipal andreense.

O MORRO DA KIBON

Como contextualização, o morro da Kibon localiza-se em uma área de alta vulnerabilidade social nas imediações dos bairros Condomínio Maracanã, Jardim do Mirante e Sítio Cassaquera, tendo seu núcleo predominante na porção Leste da Avenida Valentim Magalhães, no município de Santo André. Neste trabalho, o bairro do Condomínio Maracanã foi utilizado como proxy principal pela quantidade de informação proveniente da região/bairro.

O Morro da Kibon é um território socialmente reconhecido, mas sem delimitação oficial única. Isso cria também um problema de erro de mensuração territorial, onde registros podem existir, mas sem granularidade e padronização suficientes para atribuição confiável ao território, gerando subcobertura e vieses em indicadores. Um erro de mensuração pode levar o morro a ter subcobertura, onde ele apenas “some” das estatísticas sobre a região principal

MÉTODOS

Este estudo foi realizado com base nos dados secundários provenientes dos bancos de dados extraídos do OpenDataSus e que possuem fonte para o SIVEP-Gripe para os casos de Síndrome Respiratória Grave Aguda no município de Santo André em geral entre 2020 e 2023 e, via Lei de Acesso à Informação (LAI) no portal e-SIC, os casos recortados do bairro Condomínio Maracanã, em Santo André, disponibilizados via tabela agregada, que possuem fonte também no SIVEP-Gripe.

Dentre as vantagens do acesso a essas fontes, destaca-se a possibilidade de comparar os dados no âmbito de um bairro em que o “Morro da Kibon” está instalado, Condomínio Maracanã, e o município de Santo André em geral. Destaca-se ainda que, devido ao não reconhecimento administrativo do Morro da Kibon, é dificultoso que se mapeie separadamente a região do Morro da Kibon como uma parte única, o que confirma previamente também, que a incompletude se compreende desde o início, quando se não é possível reconhecer uma região.

É importante salientar que, foram procurados diversos materiais e bancos de dados que dissesse algo sobre a situação no Morro da Kibon, mas que foram dificultados e limitados pelo não reconhecimento administrativo da região, ficando condicionais aos dados interseccionados dos bairros do Condomínio Maracanã, Sítio Cassaquera e Jardim do Mirante. Essa pesquisa acima de tudo respeita a Lei Geral de Proteção de Dados (LGPD) e então, não aceitou bancos de dados com variáveis que ferem seus princípios, como variáveis que divulgam o endereço exato do paciente ou outras informações sensíveis que apontam para o estabelecimento daquele indivíduo naquela região, houve apenas a utilização de dados públicos e anonimizados.

Especificamente, para a análise dos dados de Santo André em geral, fora utilizado um banco de dados de todo o Estado de São Paulo para a Síndrome Respiratória Aguda Grave e desses dados, foram filtrados via *software RStudio* apenas os dados que interessavam, como os pacientes que residem no município de Santo André. As variáveis de interesse foram: raça/cor, escolaridade, sexo, idade e evolução do caso. As variáveis-chave para o objetivo do estudo foram raça/cor e escolaridade, por sua relevância na descrição de desigualdades.

A incompletude foi definida como a presença da categoria “Ignorado” / “Não Informado” em Raça/Cor (CS_RACA), Escolaridade (CS_ESCOL_N), Sexo (CS_SEXO) e evolução do caso (EVOLUCAO), situações alternativas como “Não se aplica” (N/A) não foram consideradas como incompletude, devido a natureza de sua regra de preenchimento e não exatamente ausência de informação

Para cada variável e recorte municipal geral, foi calculada a taxa de incompletude em porcentagem como a fórmula (i) e concluída em um heatmap.

$$\%Incompletude = \frac{I}{N} \times 100,$$

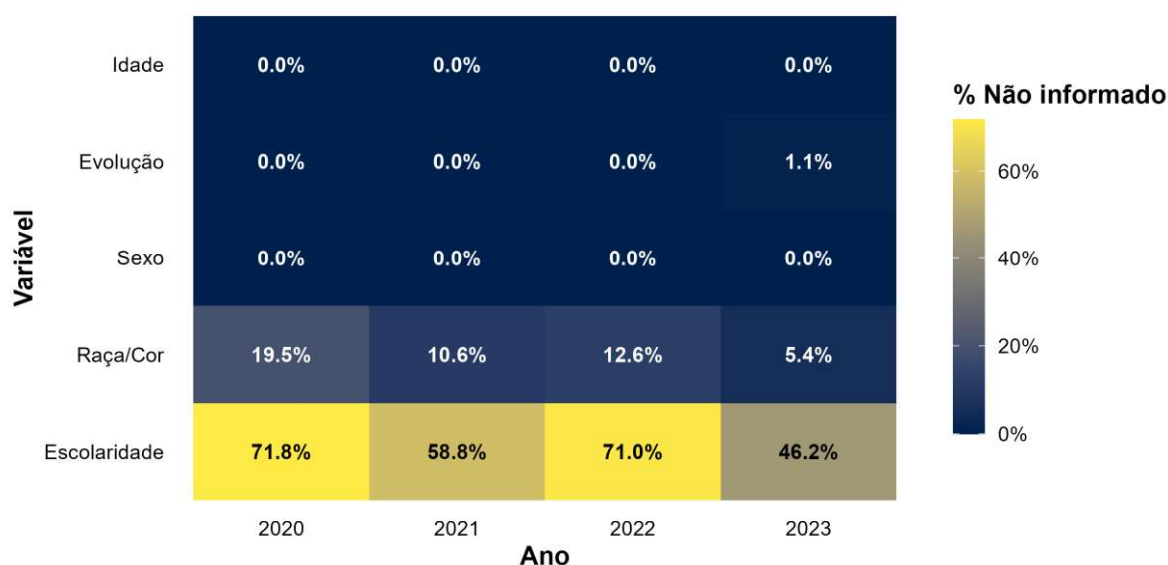
sendo I = Número de respostas “Não Informado/Ignorado” e N = número total de registros

RESULTADOS

Durante o período analisado entre 2020 e 2023 para os casos de Síndrome Respiratória Aguda Grave (SRAG) no município de Santo André, foi observado uma tendência de respostas “Não Informado” em maior quantidade em variáveis sociais como Raça/Cor e escolaridade, sendo a última em maior quantidade proporcional em comparação a outras respostas (figura 1). Tal achado pode mostrar que erros de mapeamentos de desigualdades sociais podem ascender no território de Santo André em geral e sobretudo, na região do Morro da Kibon. Com políticas públicas sendo debatidas e que não refletem de acordo com a real necessidade da população.

No heapmap (Figura 1), há uma afluência de incompletude sobretudo nas categorias sociais, com valores quase zerados em categorias administrativas, com uma leve divergência na categoria Evolução no ano 2023, com 1,1% de casos não informado. Ainda se verifica no geral que há uma maior concentração de respostas não informadas na variável “Escolaridade”, variando entre 46,2% e 71,8%, com predominâncias nos anos de 2020 e 2022, respectivamente, e com média de 61,9% nos 4 anos; Em Raça/Cor variou de 5,4% a 19,5%, com média de 12%. Já as outras categorias/Variáveis restantes obtiveram $\leq 1,1\%$

FIGURA 1 – Heapmap de respostas “Não Informado” comparado a outras respostas no Município de Santo André para casos de SRAG



FONTE: Sivep-Gripe

Verifica-se também que dados provenientes do Morro da Kibon não existem separadamente no campo da saúde pública municipal, com a solicitação dos dados referentes a COVID-19 nos três bairros que interceptam a comunidade, fica claro que os dados do núcleo virão acompanhados de dados de outras regiões, como exclusivamente o bairro Condomínio Maracanã (Tabela 1), o que compromete fortemente medidas isoladas para o Morro da Kibon, firmando ainda mais a vulnerabilidade social da região e, conseqüentemente, distanciando do cumprimento das ODS-2030. Um indício é a existência de casos confirmado para pessoas amarelas, onde não há registros de habitantes amarelos na região do Morro da Kibon, mas há para o bairro Condomínio Maracanã.

Na tabela 1, é visível também o registro em quantidade abundante de casos ignorados nos anos de 2020 e 2021, com 10 registros em 2020 e 18 registros em 2021, tendo um valor expressivamente maior em comparação às outras variáveis.

TABELA 1. DADOS DE INFECÇÃO POR COVID-19 NO CONDOMÍNIO MARACANÃ

DIMENSÃO	CATEGORIA	ANO			
		2020	2021	2022	2023
TOTAIS	CASOS	33	88	5	1
	CONFIRMADOS				
	ÓBITOS	12	28	1	0
SEXO (ÓBITOS)	FEMININO	4	14	0	0
	MASCULINO	8	14	1	0
RAÇA/COR	AMARELA	0	1	0	0
	BRANCA	9	11	1	0
	IGNORADO	1	5	0	0
	INDIGENA	0	9	0	0
	PARDA	2	2	0	0
	PRETA	2	2	0	0
ESCOLARIDADE	6º A 9º ANO	1	6	1	0
	ENSINO MÉDIO	1	4	0	0
	ANALFABETO	0	1	0	0
	IGNORADO	10	18	0	0

FONTE: Sivep-Gripe

Na tabela 1, se optou por dados de casos confirmados de COVID-19 como medida comparativa, enquanto a Figura 1 utiliza registros de SRAG. Essa divergência não é interpretada como intercambialidade entre os indicadores, mas sim como consequência de uma limitação de disponibilidade onde não foi possível obter dados de SRAG com desagregação adequada e consistência no nível do bairro. Dado que o objetivo do estudo envolve a caracterização da incompletude informacional em escala local, priorizou-se a manutenção do recorte territorial. Dessa forma, os dados de COVID-19 foram utilizados como indicador disponível para contextualização do fenômeno, reconhecendo-se potenciais vieses (p.ex., variações de testagem/notificação e diferenças conceituais entre SRAG e infecção confirmada) e restringindo-se a interpretação a comparações indicativas.

DISCUSSÃO

Os resultados deste trabalho apontam um padrão bem nítido, na qual se observa a incompletude nos registros de SRAG do município de Santo André entre 2020 e 2023, com a maior concentração de respostas “Não informado/Ignorado” aparecendo nas variáveis sociais, principalmente escolaridade e raça/cor, enquanto variáveis administrativas e de desfecho (como sexo e evolução) ficam com valores próximos de 0% no mesmo período. A escolaridade variou de 46,2% a 71,8% (média de 61,9% nos quatro anos) e raça/cor variou de 5,4% a 19,5% (média de 12%), ao passo que as demais variáveis ficaram em $\leq 1,1\%$. Isso é importante porque, são essas variáveis (raça/cor e escolaridade) que permitem enxergar diferenças dentro de territórios como o Morro da Kibon e entender quem é mais atingido e como isso se distribui

estatisticamente. Se a variável social some, a desigualdade tende a ficar invisível ou muito distorcida. Entrando um ponto crítico onde mesmo quando existe uma intenção de orientar políticas públicas por evidências e via dados, o que se mede com baixa qualidade pode virar um instrumento para decisões erradas. Um indicador pode existir, mas ele passa a representar uma média vazia, que não descreve o que acontece com grupos específicos.

Quando a discussão sai do município de Santo André em geral e tenta se aproximar da comunidade do Morro da Kibon, o cenário fica mais delicado. Este trabalho mostra que o território é socialmente reconhecido, mas não possui delimitação oficial única, e isso produz erro de mensuração territorial, onde registros podem existir, mas não com padronização suficiente para atribuição confiável ao território, levando à subcobertura e a vieses em indicadores. Em outras palavras, mesmo que a qualidade do preenchimento melhorasse, ainda existiria um obstáculo estrutural onde se o território não é reconhecido como recorte administrativo, ele não aparece como categoria de análise e, por consequência, fica difícil produzir evidência local consistente. Por isso foi adotado o bairro Condomínio Maracanã como proxy, pela maior disponibilidade de informação, mas esse caminho carrega um custo metodológico inevitável no qual os dados do núcleo vem misturados com dados de outras regiões do bairro, o que compromete medidas isoladas para o Morro e limita inferências locais.

O ponto aqui não é dizer que não é útil, e sim dizer o que ele permite e o que ele não permite. Ele permite contextualizar e levantar indícios sobre a invisibilidade do território, mas não garante que qualquer número obtido seja do Morro da Kibon, porque o recorte não é o Morro da Kibon, é um bairro que o intercepta. Essa discussão inclusive também explica por que o estudo usa, em momentos distintos, registros de SRAG para o município em geral e dados agregados de COVID-19 no Condomínio Maracanã, e deixa claro que isso não é intercambiabilidade entre indicadores, mas consequência de disponibilidade e de limitações de desagregação em nível de bairro.

Voltando ao padrão de incompletude em si, uma questão central é: por que exatamente escolaridade e raça/cor têm tantas marcações como “Ignorado/Não informado”? Aqui entram algumas hipóteses plausíveis. Pode ser falta de treinamento e padronização sobre como coletar; pode ser a dinâmica de atendimento de urgência no período da pandemia, onde prioriza-se o clínico e o desfecho, enquanto o social fica para depois e acaba não sendo registrado e pode ser também um efeito de como a informação chega, por exemplo, se o paciente não consegue responder, se não há acompanhante, se o profissional não retoma depois. O ponto não é escolher uma hipótese como verdade sem evidência, mas reconhecer que, sem mapear a causa, a incompletude vai se repetir, e isso vai continuar vulnerabilizando diagnósticos e políticas públicas re.

Por fim, é importante reconhecer limites do que este estudo consegue afirmar. Ele trabalha com dados secundários e com a definição de incompletude como presença explícita de “Ignorado/Não informado” em variáveis selecionadas. Além disso, no nível local, parte do material vem em forma agregada e com recorte proxy, o que aumenta incerteza de atribuição territorial. Mesmo assim, esse estudo cumpre um papel que é necessário mostrando que a

lacuna não é aleatória, pois ela se concentra exatamente nos marcadores sociais, e mostrar também que, para o Morro da Kibon, o problema começa muito antes do preenchimento, começando na impossibilidade de reconhecer o território dentro da estrutura de dados. Esse trabalho se encerra alegando que não exatamente pode ter faltado informações, mas com certeza, faltou visibilidade em quaisquer esferas.

REFERÊNCIAS

MINISTÉRIO DA SAÚDE (BR). OpenDataSUS. Banco de dados da Síndrome Respiratória Aguda Grave (SRAG) – 2019 a 2026. Disponível em: <https://opendatasus.saude.gov.br/> (buscar pelo conjunto “SRAG 2019 a 2026”). Acesso em: 11/02/2026

BRASIL. Governo Federal. dados.gov.br. SRAG 2019 – Banco de Dados de Síndrome Respiratória Aguda Grave. Disponível em: <https://dados.gov.br/dados/conjuntos-dados/srag-2019>. Acesso em: 11/02/2026

SECRETARIA DE ESTADO DA SAÚDE (ES). Guia do SIVEP-Gripe – Vigilância de influenza (SRAG). [PDF]. (Referência a guia integrado covid-19/influenza e rotinas de vigilância). Disponível em: <https://saude.es.gov.br/> (buscar pelo PDF “Guia do SIVEP-GRIPE”). Acesso em: 11/02/2026

BRASIL. Lei nº 12.527, de 18 de novembro de 2011. Regula o acesso a informações (Lei de Acesso à Informação). Diário Oficial da União: Brasília, DF, 2011. Disponível em: <https://www2.camara.leg.br/legin/fed/lei/2011/lei-12527-18-novembro-2011-611802-publicacaooriginal-134287-pl.html>. Acesso em: 11/02/2026

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, 2018. Disponível em: <https://www2.camara.leg.br/legin/fed/lei/2018/lei-13709-14-agosto-2018-787077-publicacaooriginal-156212-pl.html>. Acesso em: 11/02/2026

UNITED NATIONS. Transforming our world: the 2030 Agenda for Sustainable Development (A/RES/70/1). New York: United Nations, 2015. Disponível em: <https://sdgs.un.org/2030agenda>. Acesso em: 11/02/2026

ONU (BR). Objetivos de Desenvolvimento Sustentável (ODS) no Brasil. Disponível em: <https://brasil.un.org/pt-br/sdgs>. Acesso em: 11/02/2026

Ribas FV, Custódio ACD, Toledo LV, Henriques BD, Sedyama CMNO, Freitas BAC. Completeness of notifications of severe acute respiratory syndrome at the national level and in a regional health care in the state of Minas Gerais, during the COVID-19 pandemic, 2020. Epidemiol Serv Saude. 2022;31(2):e2021620. doi: 10.1590/S1679-49742022000200004. PMID: 35730813; PMCID: PMC9887980. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/35730813/>. Acesso em: 11/02/2026

CORREIA, L. O. S. et al. Métodos para avaliar a completitude dos dados dos sistemas de informação em saúde do Brasil: uma revisão sistemática. *Ciência & Saúde Coletiva*, 2014. DOI: 10.1590/1413-81232014196.08622013. Disponível em: <https://www.scielo.br/j/csc/a/HGyrfBHWLXMd3mz74HCcvpy/?lang=pt>. Acesso em: 11/02/2026

ROMERO, D. E.; CUNHA, C. B. Avaliação da qualidade das variáveis sócio-econômicas e demográficas do Sistema de Informações sobre Nascidos Vivos (SINASC), 2002. *Cadernos de Saúde Pública*, [local], 2006. Disponível em: <https://www.scielo.br/j/csp/a/HXqrdksBsMrr4R9Ydnnmqcf/?lang=pt>. Acesso em: 11/02/2026

SANTOS, H. L. P. C. dos et al. Trend of incompleteness of the race/color variable in hospitalizations for COVID-19 whose outcome was death in Brazil, 2020–2022. *Revista de Saúde Pública*, 2024. Disponível em: <https://pmc.ncbi.nlm.nih.gov/articles/PMC11548909/>. Acesso em: 11/02/2026.

LITTLE, R. J. A.; RUBIN, D. B. *Statistical Analysis with Missing Data*. 3. ed. Hoboken: Wiley, 2019.

VAN BUUREN, S. *Flexible Imputation of Missing Data*. 2. ed. Boca Raton: CRC Press, 2018.

STERNE, J. A. C. et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, v. 338, b2393, 2009. DOI: 10.1136/bmj.b2393. Disponível em: <https://www.bmj.com/content/338/bmj.b2393>. Acesso em: 11/02/2026.

OPENSHAW, S. *The Modifiable Areal Unit Problem*. Norwich: Geo Books, 1984. Disponível em: <https://www.uio.no/studier/emner/sv/iss/SGO9010/openshaw1983.pdf>. Acesso em: 11/02/2026

ONU-HABITAT. *The Challenge of Slums: Global Report on Human Settlements* 2003. London: Earthscan; UN-Habitat, 2003. Disponível em: <https://unhabitat.org/sites/default/files/download-manager-files/The%20Challenge%20of%20Slums%20-%20Global%20Report%20on%20Human%20Settlements%202003.pdf>. Acesso em: 11/02/2026.

RIBEIRO, A. I. et al. Tools for Address Georeferencing: a systematic review. *PLOS ONE*, 2014. DOI: 10.1371/journal.pone.0114130. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0114130>. Acesso em: 11/02/2026.

R CORE TEAM. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing, 2025. Disponível em: <https://www.R-project.org/>. Acesso em: 11/02/2026.

POSIT SOFTWARE, PBC. RStudio: Integrated Development Environment for R. Boston, MA: Posit, 2025. Disponível em: <https://posit.co/products/open-source/rstudio/>. Acesso em: 11/02/2026.