**Análise preditiva da fixação de profissionais do Programa Mais Médicos usando machine Learning**

**Resumo**

**Introdução**

O Brasil passa por um problema persistente em relação à carência no acesso a médicos. Em grande medida, isso ocorre, pois, a distribuição de profissionais é desigual entre as regiões do país e mesmo dentro de um mesmo município (Cury & Fonseca, 2023). Enquanto no Sudeste, existem 3,39 médicos por mil habitantes, no Norte, este número é menos da metade, com 1,45 por mil habitantes. Além deste fato, a maior parte dos médicos está concentrada em capitais dos estados (Scheffer et al., 2023).

O Programa Mais Médicos (PMM) foi criado em 2013 no intuito de reduzir essas inequidades. Na sua concepção, o PMM contou com três eixos. O primeiro, visava assegurar o provimento emergencial de médicos na Atenção Primária à Saúde (APS) em lugares desassistidos e com maior vulnerabilidade. O segundo, eixo infraestrutura, continha ações de construção, reforma e ampliação de unidades, bem como o aprimoramento tecnológico. Por fim, o eixo formação envolveu medidas de médio e longo prazo para induzir a criação de cursos de medicina e residências tanto em instituições públicas e privadas em locais de maior vulnerabilidade (Brasil, 2015).

Todavia, ao longo dos mais de dez anos do PMM, o propósito de cada eixo passou por mudanças significativas em relação ao conceito original. Isso foi ainda mais acentuado nos eixos infraestrutura e educação, com interrupção das ações já no governo Temer (Pinto et al., 2022). O eixo provimento era o que tinha mais holofotes, inclusive com mais contendas políticas (Hone et al., 2020). Com a entrada do governo Bolsonaro, foi criado um programa que coexistiu ao PMM – o Programa Médicos Pelo Brasil (PMpB) – que previa a participação de apenas médicos com registros no Brasil, mas que só teve o primeiro edital lançado em 2022. O retorno do governo Lula levou a reajustes no programa, que passou a se chamar Programa Mais Médicos para o Brasil (PMMB) e contou com a reativação de ações relacionadas ao eixo educação (Cury & Fonseca, 2023).

Diversos estudos foram realizados a fim de avaliar os efeitos do programa. Há resultados que mostram que houve aumento da oferta de médicos (Pinto Junior et al., 2020), crescimento no número de consultas (Mattos & Mazetto, 2019) e internações por condições sensíveis à APS (Maffioli et al., 2019; Russo et al., 2020). No entanto, algumas investigações alertam que o impacto do programa foi menor do que se esperava, pois houve quantidade expressiva de profissionais alocados em regiões de baixa prioridade (Hone et al., 2020; Thomas et al., 2024) e, além disso, o PMM acabou se tornando uma política focada em único insumo (médicos) em detrimento de um conjunto ampliado de fatores (ex.: outras categorias profissionais, infraestrutura e aprimoramento na coordenação das redes de atenção à saúde) (Rocha, 2025).

Além disso, reflexões feitas a partir de exemplos de outros países sugerem que ações como o PMM contribuem para suprir, a curto-prazo, as lacunas de profissionais. No entanto, desafios relacionados à sustentabilidade do programa, especialmente com a integração da força de trabalho à infraestrutura local, permanecem (Maffioli et al., 2019; Rocha, 2025). Esta última reflexão levanta a necessidade de se investigar um fenômeno subjacente às dinâmicas do programa e que é pouco evidenciado, que é a rotatividade ou fixação dos profissionais do PMM, sobretudo após a conclusão do ciclo. Afinal, medidas reativas de alavancagem da oferta de profissionais a curto prazo, como ocorrido no PMM, nem sempre podem ser garantia sustentável da disponibilidade de profissionais a longo prazo (Freer, 2017; Rees et al., 2023).

Ainda que a fixação tenha sido pouco analisada, declarações por parte dos formuladores da política de alto escalão do Ministério da Saúde têm demostrado que é um elemento que precisa ser gerenciado (Cambricoli, 2023; Teixeira & Santos, 2023). Tais argumentos estão alinhados à normatização de incentivos financeiros para aumentar a fixação dos profissionais em áreas mais vulneráveis (Brasil, 2023).

O presente estudo buscar avançar em uma análise sobre o fenômeno da fixação de profissionais ao levantar a seguinte pergunta de pesquisa: é possível prever quais profissionais vão permanecer no local para onde foram designados, mesmo após o fim do programa? Desta pergunta, emerge-se o objetivo do estudo que foi desenvolver um modelo preditivo da fixação de profissionais, após o fim do ciclo, nos locais para onde são alocados, com apoio de algoritmos de *machine learning*.

Esta pesquisa traz contribuições importantes para fins teóricos e práticos. Em relação à primeira, observa-se que a maior dos estudos que analisam o PMM adota uma abordagem econométrica, agregada, na qual o força de trabalho médica é colocada como um insumo (Hone et al., 2020; Mattos & Mazetto, 2019; Özçelik et al., 2020; Pinto Junior et al., 2020; Russo et al., 2020; Thomas et al., 2024). Os estudos geralmente não analisam o indivíduo atuante no programa e o desfecho da sua participação. Portanto, este estudo inova ao usar uma lente ampliada sobre um dos sujeitos primordiais para o sucesso da pesquisa. Na perspectiva prática, temos evidências de que a rotatividade de profissionais pode gerar descontinuidade no relacionamento com o paciente, o que pode impactar negativamente vários desfechos na saúde do indivíduo (Sabety et al., 2021). Logo, um modelo preditivo poderia contribuir para a melhor alocação dos profissionais com base na predisposição de fixação. Além disso, medidas preventivas poderiam ser tomadas a fim de minimizar a rotatividade dos profissionais com maior probabilidade de saída.

**Fundamentação Teórica**

Os resultados também indicaram a forte influência dos aspectos socioeconômicos no índice de vulnerabilidade e condições em saúde. São justamente esses os principais fatores que afetam as decisões dos médicos quanto ao seu local de trabalho. Em função disso, políticas de desenvolvimento regional e as de mitigação das desigualdades socioeconômicas que melhore as condições de vida e de saúde das distintas localidades pode ser um importante propulsor para contribuir para uma melhor alocação e fixação de médicos no território nacional. (Macedo)

**Método**

**Caracterização da pesquisa**

Esta pesquisa possui abordagem quantitativa, com uso de dados secundários e emprego de técnicas de *machine learning* para prever o desfecho do estudo: fixação/rotatividade de médicos participantes do PMM. Neste sentido, foram acessados dados do 18º Ciclo do PMM, realizado em 2019. A opção por este ciclo se deve pela presença de alguns atributos a nível indivíduo, que não estavam disponíveis em chamadas de ciclos anteriores. Já as chamadas de 2020 adiante não completaram um ciclo completo (considerando o tempo do ciclo e possível prorrogação). Todos os dados são públicos, disponíveis por meio de portarias do Ministério da Saúde que trazem resultados do PMM, assim como bases de diversas fontes. Deste modo, não foi necessária a apreciação da investigação por comitê de ética em pesquisa. A seguir, narramos os passos usados para realizar o tratamento dos dados.

**Tratamento dos dados**

O primeiro passo consistiu em acessar dados dos médicos que foram aprovados do PMM no ciclo mencionado. Os resultados da seleção foram divulgados publicamente por meio da Portaria Nº 7, de 18 de junho de 2019, e a Portaria Nº 13 de 16 de agosto de 2019. Entre os dados divulgados e que foram utilizados para análise estão: nome, município para onde o profissional é alocado, data de nascimento, mês/ano de formação e participação pregressa no programa. Nesta etapa, o total de profissionais era de 2.452 médicos.

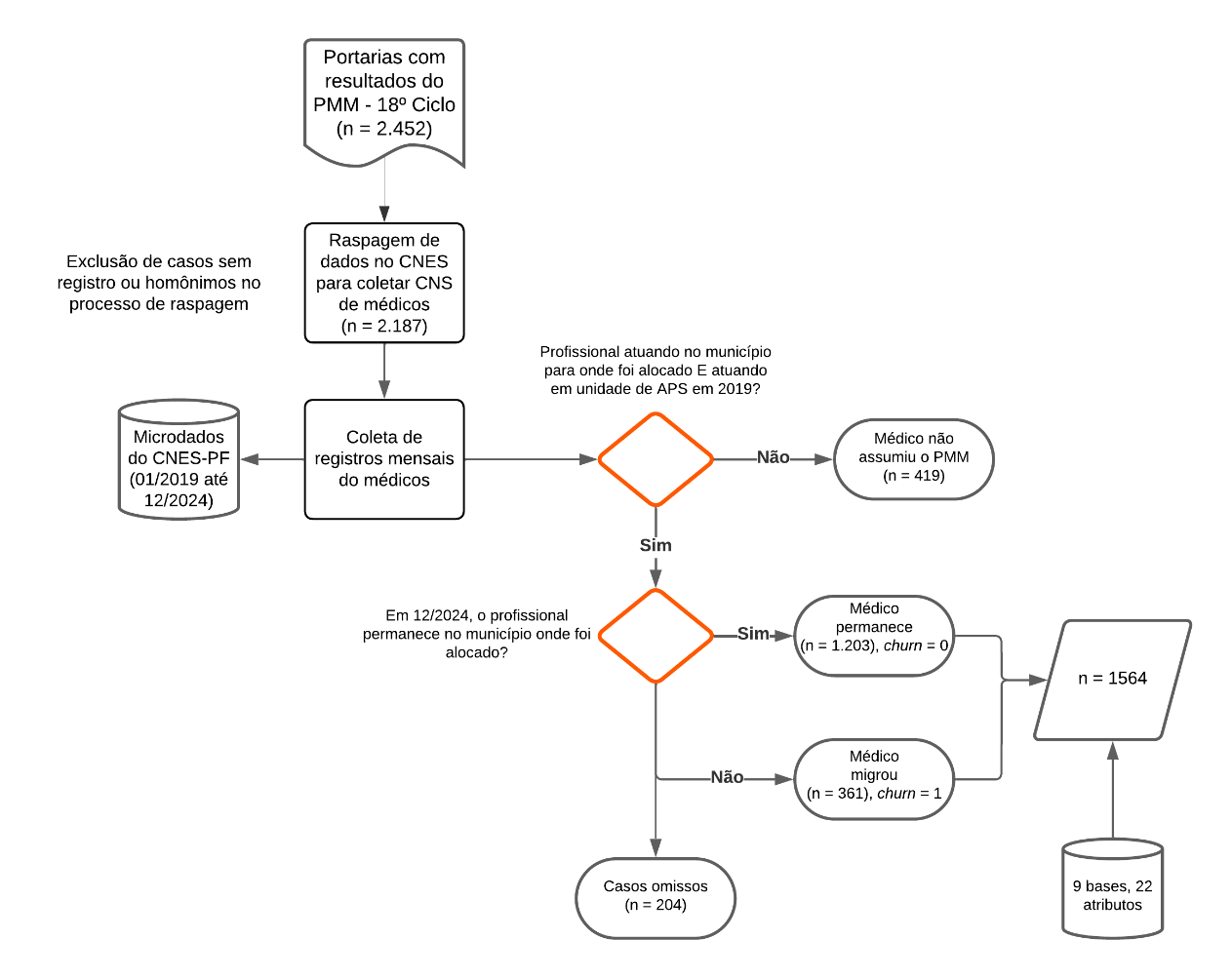
Em sequência, foi realizado um procedimento de automatização da coleta de dados dos profissionais selecionados por meio da página do Cadastro Nacional de Estabelecimentos de Saúde (CNES). Para este procedimento foi utilizada a biblioteca *RSelenium* na linguagem R. Um total de 256 médicos foram excluídos devido à falta de retorno nas consultas ou em razão da presença de homônimos nos retornos. Dessa forma, permaneceram 2.187 médicos para análise.

O próximo passo consistiu em acessar microdados da base CNES-Profissionais (CNES-PF) entre janeiro de 2019 e dezembro de 2024 que estão disponíveis via transferência de arquivos públicos do Datasus. O CNES-PF contém o registro mensal dos profissionais vinculados a estabelecimentos de saúde. Com isso, avaliamos alguns critérios para concluir a construção da base:

1. o profissional aprovado assumiu o programa? A resposta afirmativa dependia de duas condicionais: o profissional estar no município para onde o PMM o designou em algum mês de 2019 e alocado em alguma unidade de Atenção Primária à Saúde (APS). O não atendimento a este conjunto de critérios o excluía da base, pois sugeria que, mesmo sendo aprovado, o indivíduo não havia assumido a atribuição. Deste tratamento foram mantidos 1.768 médicos;
2. Passados cinco anos do programa, o indivíduo permanece no município ao qual foi alocado? Cabe lembrar que a duração do 18º ciclo do programa era de três anos, podendo ser prorrogado por mais um. A resposta para essa pergunta foi importante para definir a variável resposta do estudo. Se o indivíduo, em dezembro de 2024, permanece no mesmo município ao qual foi alocado, a variável assume o valor de “permanece”. Caso o profissional, em dezembro de 2024, esteja atuando em estabelecimentos de outros municípios, a variável assume o valor de “migrou”. Existem casos atípicos em que o profissional deixa de ter registros na base em anos/meses antes do término do programa. Esses casos foram excluídos, resultando uma amostra final de 1.564 médicos.

O tratamento completo dos dados pode ser sintetizado pela figura 1.

Figura 1 - Fluxograma de tratamento de dados



Fonte: elaborado pelos autores

De posse da amostra, foi realizado o enriquecimento incluindo outros atributos em diferentes níveis de análise, conforme apresentado na fundamentação teórica. O quadro 1 lista as dimensões, variáveis e respectivas fontes de dados.

Quadro 1 – Lista de variáveis independentes

|  |  |  |
| --- | --- | --- |
| Dimensão | Variáveis | Fonte de dados |
| Individual | Município de alocação | Portaria Nº 7, de 18 de junho de 2019; Portaria Nº 13 de 16 de agosto de 2019 |
| Idade |
| Participação prévia no PMM |
| Estabelecimento | Quantidade média de agentes comunitários de saúde nas unidades de APS na qual o médico atuou | Cadastro Nacional de Estabelecimentos de Saúde (CNES) |
| Quantidade média de técnicos e auxiliares de enfermagem nas unidades de APS na qual o médico atuou |
| Quantidade média mensal de enfermeiros nas unidades de APS na qual o médico atuou |
| Quantidade média de cirurgiões-dentistas nas unidades de APS na qual o médico atuou |
| Vínculos de trabalho | Quantidade de vínculos com outros estabelecimentos de saúde (ex.: plantões em hospitais da região) |
| Tempo de atuação prévia, anterior ao PMM, no município alocado |
| Município | População | Instituto Brasileiro de Geografia e Estatística (IBGE) |
| PIB per capita de 2019 |
| Taxa de população ocupada |
| Número de vagas em curso de medicina | Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) |
| Orçamento per capita em saúde | Índice de Desenvolvimento Sustentável das Cidades – Brasil (IDSC-BR) |
| Percentual da população atendida com esgoto sanitário |
| Taxa de homicídio |
| IDEB – Anos Iniciais |
| IDEB – Anos Finais |
| Investimento público em infraestrutura per capita |
| Número de equipamentos esportivos públicos por 100 mil habitantes |
| Número de centros culturais, espaços e casas de cultura públicos e privados por 100 mil habitantes |
| Distância da capital | Google Maps |

Fonte: elaborado pelos autores

**Análise de dados**

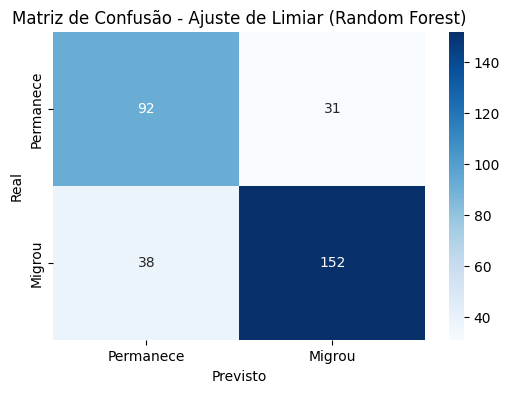
A modelagem preditiva com apoio de *machine learnin*g foi realizada utilizando linguagem Python e a biblioteca *sckit-learn*. A variável de resposta foi a fixação/rotatividade com valor um em caso de migração e zero em caso de permanência, o que torna este um problema de classificação. Percorreu-se o processo de *machine learning*, no qual, uma das primeiras etapas foi pré-processamento dos dados. No caso de variáveis numéricas, realizou-se a substituição de valores faltantes pela mediana e a padronização pelo método Z-score. No caso das variáveis categóricas foi realizada a substituição dos valores ausentes pela categoria de maior frequência e a codificação pelo método *one-hot encoding*. Os dados foram divididos em treino e teste, 80% e 20%, respectivamente.

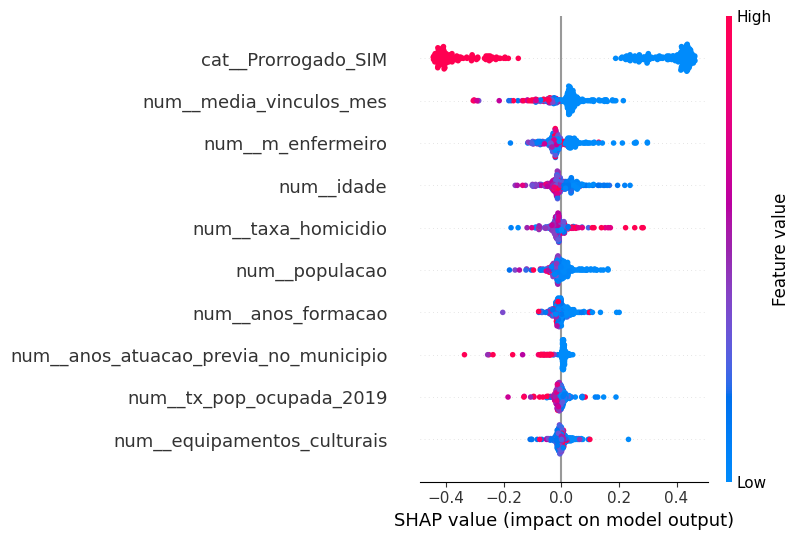
Quatro modelos foram treinados e avaliados: Regressão Logística, Árvore de Decisão, *Random Forest*, *XGBoost* e *LightGBM*. A otimização dos hiperparâmetros foi realizada com a biblioteca *Optuna*. O treinamento envolveu validação cruzada estratificada com *StratifiedKFold* em 5 partes. Devido ao desbalanceamento da variável resposta, utilizou-se a abordagem do *Synthetic Minority Oversampling Technique* (SMOTE). Ao final, utilizamos a biblioteca *Shapley Additive exPlanations* (SHAP) para adicionar uma camada de explicação das previsões do melhor modelo.

**Resultados**

Os algoritmos foram treinados considerando quatro conjuntos distintos de atributos. O primeiro apenas com variáveis a nível do indivíduo, o segundo adicionando as relacionadas aos estabelecimentos e vínculos de trabalho, a terceira agregando as relacionadas ao município e por último adicionando a variável prorrogação, de forma cumulativa. São denominados na tabela 2, respectivamente, como M1, M2, M3 e M4. Os algoritmos que obtiveram os melhores desempenhos foram a Regressão Logística (M1), XGBoost (M3) e *Random Forest* (M2 e M4).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | M1 | M2 | M3 | M4 |
| Melhor algoritmo | Regressão Logística | Random Forest | XGBoost | Random Forest |
| Acurácia | 0.667 | 0.686 | 0.706 | 0.789 |
| Precisão | 0.726 | 0.750 | 0.737 | 0.860 |
| Recall | 0.726 | 0.726 | 0.800 | 0.778 |
| F1 Score | 0.726 | 0.737 | 0.767 | 0.81 |
| ROC AUC | 0.697 | 0.749 | 0.789 | 0.860 |





Discussão

Macedo Os resultados também indicaram a forte influência dos aspectos socioeconômicos no índice de vulnerabilidade e condições em saúde. São justamente esses os principais fatores que afetam as decisões dos médicos quanto ao seu local de trabalho. Em função disso, políticas de desenvolvimento regional e as de mitigação das desigualdades socioeconômicas que melhore as condições de vida e de saúde das distintas localidades pode ser um importante propulsor para contribuir para uma melhor alocação e fixação de médicos no território nacional.

In sum, if doctors are placed in under-resourced areas and where

complementary inputs are lacking, including management capabilities,

their ability to improve health outcomes can be severely constrained.

This may also be the case if doctors are placed where substitute inputs

exist. As a corollary, the existing evidence suggests that future in­

terventions should focus not only on increasing doctor supply but also

on ensuring that substitute inputs are not present, and that all necessary

complementary inputs are available to support effective care delivery. (rocha rudi)

Considerações Finais

Limitações

Os dados foram referentes a apenas um ciclo do programa. Não foi possível utilizar de outros ciclos, pois, os mais antigos não continham dados a nível indivíduo e os mais recentes não tiveram tempo suficiente para concluir o ciclo completo. O 18º Ciclo do PMM perpassou períodos de instabilidades sobre a continuidade do programa e a pandemia de Covid-19. Portanto, é importante que estudos futuros se abasteçam de dados de ciclos subsequentes a fim de avaliar os modelos alcançados neste estudo e evoluí-los.

O trabalho contou apenas com dados públicos, que foram compilados a partir de uma coleção de fontes. Uma possível parceria com o Ministério da Saúde poderia

Estudos Futuros

Referencial

Brasil. (2023). Lei no 14.621, de 14 de Julho de 2023. In *Lei no 14.621, de 14 de Julho de 2023*.

Cambricoli, F. (2023, March 21). Mais Médicos terá desafio de fixar profissionais e versão de especialistas, diz ministra da Saúde. *Estadão/Saúde*.

Cury, G. C., & Fonseca, A. F. (2023). A retomada do Programa Mais Médicos em 2023. *Trabalho, Educação e Saúde*, *21*. https://doi.org/10.1590/1981-7746-ojs2415

Freer, J. (2017). Sustainable development goals and the human resources crisis. *International Health*, *9*(1), 1–2. https://doi.org/10.1093/inthealth/ihw042

Hone, T., Powell-Jackson, T., Santos, L. M. P., De Sousa Soares, R., De Oliveira, F. P., Sanchez, M. N., Harris, M., De Oliveira De Souza Santos, F., & Millett, C. (2020). Impact of the Programa Mais médicos (more doctors Programme) on primary care doctor supply and amenable mortality: quasi-experimental study of 5565 Brazilian municipalities. *BMC Health Services Research*, *20*(1). https://doi.org/10.1186/s12913-020-05716-2

Maffioli, E. M., Hernandes Rocha, T. A., Vivas, G., Rosales, C., Staton, C., & Nickenig Vissoci, J. R. (2019). Addressing inequalities in medical workforce distribution: Evidence from a quasi-experimental study in Brazil. In *BMJ Global Health* (Vol. 4, Issue 6). BMJ Publishing Group. https://doi.org/10.1136/bmjgh-2019-001827

Mattos, E., & Mazetto, D. (2019). Assessing the impact of more doctors’ program on healthcare indicators in Brazil. *World Development*, *123*. https://doi.org/10.1016/j.worlddev.2019.104617

Özçelik, E. A., Massuda, A., McConnell, M., & Castro, M. C. (2020). Impact of Brazil’s More Doctors Program on hospitalizations for primary care sensitive cardiovascular conditions. *SSM - Population Health*, *12*. https://doi.org/10.1016/j.ssmph.2020.100695

Pinto, H., Oliveira, F. P. de, & Soares, R. (2022). Panorama da implementação do Programa Mais Médicos até 2021 e comparação com o Programa Médicos pelo Brasil. *Revista Baiana de Saúde Pública*, *46*(1), 32–53. https://doi.org/10.22278/2318-2660.2022.v46.n1.a3616

Pinto Junior, E., Amorim, L., & Aquino, R. (2020). Programa Mais Médicos: contexto de implantação e efeito no provimento de médicos na atenção primária à saúde no Brasil, 2008 a 2016. *Revista Panamericana de Salud Pública*, *44*, 1–9. https://doi.org/10.26633/rpsp.2020.23

Rees, G. H., James, R., Samadashvili, L., & Scotter, C. (2023). Are Sustainable Health Workforces Possible? Issues and a Possible Remedy. In *Sustainability (Switzerland)* (Vol. 15, Issue 4). MDPI. https://doi.org/10.3390/su15043596

Rocha, R. (2025). More doctors, better health? Consolidating evidence from Brazil’s Mais Médicos program. *Social Science and Medicine*, *364*. https://doi.org/10.1016/j.socscimed.2024.117559

Russo, L. X., da Silva, E. N., Rosales, C., Rocha, T. A. H., & Vivas, G. (2020). Efeito do Programa Mais Médicos sobre internações. *Revista Panamericana de Salud Pública*, 1–8. https://doi.org/10.26633/rpsp.2020.25

Sabety, A. H., Jena, A. B., & Barnett, M. L. (2021). Changes in Health Care Use and Outcomes After Turnover in Primary Care. *JAMA Internal Medicine*, *181*(2), 186. https://doi.org/10.1001/jamainternmed.2020.6288

Saúde Para Os Brasileiros, M. (2015). *Mais Médicos-Dois anos*.

Scheffer, M., Guilloux, A. G. A., Miotto, B. A., & Almeida, C. de J. (2023). *Demografia Médica no Brasil 2023*. https://amb.org.br/

Teixeira, C. P., & Santos, D. V. D. dos. (2023). Programa Mais Médicos pelo Brasil: entrevista com o Prof. Dr. Felipe Proenço de Oliveira, Secretário Adjunto de Atenção Primária à Saúde do Ministério da Saúde. *Revista Portal: Saúde e Sociedade*, *8*(Especial). https://doi.org/10.28998/rpss.e02308002esp

Thomas, R. L., Millett, C., Sousa Soares, R. de, & Hone, T. (2024). More doctors, better health? A generalised synthetic control approach to estimating impacts of increasing doctors under Brazil’s Mais Medicos programme. *Social Science and Medicine*, *358*. https://doi.org/10.1016/j.socscimed.2024.117222