**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 certa medida, isso ocorre, pois, a distribuição de profissionais é desigual entre as regiões do país ou 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 – tanto brasileiros quanto estrangeiros – 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 no governo Temer (Pinto Junior 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 manifestando como 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 seja pouco evidenciada, 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 que pode apoiar o gerenciamento 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 nos locais para onde são alocados com apoio de algoritmos de *machine learning*.

A maioria dos estudos que analisam o PMM adota uma abordagem econométrica, agregada, na qual o força de trabalho médica é colocada como um insumo, a exemplo de Hone et al. (2020), Mattos e Mazetto (2019), Özçelik et al. (2020), Pinto Junior et al. (2020), Russo et al. (2020) e Thomas et al. (2024). Os estudos geralmente não analisam o indivíduo atuante no programa e o desfecho da sua participação. Portanto, uma das inovações do estudo é usar uma lente ampliada sobre o médico, que é um dos sujeitos primordiais para o sucesso da pesquisa.

Considerando o fenômeno da fixação, há amplo reconhecimento de que isso é um problema que afeta vários países, especialmente em áreas rurais e remotas (Kumar & Clancy, 2021; Wakerman et al., 2019). No entanto, as investigações que empregam abordagens preditivas com *machine learning* geralmente se concentram na saída do profissional de um estabelecimento de saúde, a exemplo de Kim et al. (2023), Lopez et al. (2023), Xu et al. (2023). Um dos poucos registros, foi um estudo que aplicou modelos de *machine learning* para prever o tempo de permanência de trabalhadores da saúde em comunidades vulneráveis na África do Sul (Moyo et al., 2018). Logo, a presente pesquisa inova ao fazer uso de uma estratégia já consolidada em um novo contexto – regional – para analisar o PMM.

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, ações preventivas poderiam ser tomadas a fim de minimizar a rotatividade dos profissionais com maior probabilidade de saída.

**Fundamentação Teórica**

A fundamentação teórica está dividida em duas subseções. Na primeira, buscamos levantar fatores associados à fixação de profissionais. Na segunda, apresentamos brevemente alguns resultados de estudos que empregaram modelagem preditiva. Porém, antes de avançar, é importante uniformizar alguns conceitos, visto que a literatura da área emprega diferentes termos para a fixação e dinâmicas correlatas: retenção, rotatividade, saída (*departure*), evasão (*churn*), migração, dentre outros.

Esta variação geralmente está ligada ao campo de estudo (ex.: saúde, administração, computação) e ao contexto do desfecho. A palavra rotatividade, por exemplo, é frequentemente usada no âmbito organizacional, enquanto os termos retenção e fixação são usados no contexto regional. Ainda que a rotatividade organizacional não seja condição suficiente para levar à migração regional, também utilizamos referenciais desse nível de análise, tendo em vista que possui um fim semelhante às dinâmicas de evasão regional. A fim de padronizar o uso dos termos, empregaremos predominantemente as palavras fixação e migração.

***Fatores determinantes da fixação***

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) – talvez ajude

***Estudos que empregaram modelagem preditiva para fixação de profissionais***

Alguns estudos já utilizaram modelagem preditiva com apoio de *machine learning* para prever a fixação da força de trabalho em saúde. A maior parte dos casos envolve a modelagem para a rotatividade de profissionais em contextos organizacionais. Uma das poucas exceções foi o estudo de Moyo et al. (2018).

Kim et al. (2023) usaram regressão logística, árvore de decisão e *random forest* para prever a rotatividade de enfermeiros na Coréia do Sul. O algoritmo de *random forest* apresentou uma acurácia de 92% e uma AUC-ROC de 0,97. De acordo com os autores, as quatro variáveis mais importantes foram: salário, idade, uso de moradia proporcionada pela organização e tipo de equipe (atendimento ambulatorial, unidade de terapia intensiva, enfermaria e departamento de administração).

Além dos algoritmos utilizados no último estudo, Xu et al. (2023) empregaram XGBoost para prever a rotatividade de enfermeiros nos EUA. Como a base tinha um alto grau de desbalanceamento da variável de desfecho, eles também aplicaram SMOTE. O algoritmo que apresentou os melhores resultados foi o *XGBoost* com acurácia de 82,19% e AUC-ROC de 0,80. As variáveis mais importantes para a previsão foram idade, carga de trabalho semanal (40 horas semanais ou acima), tipo de vínculo de trabalho e renda individual.

Lopez et al. (2023) aplicaram uma modelagem para prever a rotatividade de médicos em uma amostra de 319 médicos americanos. Por utilizarem registros eletrônicos, os pesquisadores conseguiram levantar um amplo conjunto de variáveis, desde características sociodemográficas até a dados sobre produtividade do profissional. O algoritmo que obteve os melhores resultados foi o XGBoost com acurácia de 97% e AUC-ROC de 0,82. As variáveis que exerceram maior importância foram tempo de serviço, idade, complexidade do perfil do paciente e demanda pelos médicos.

Por fim, o estudo de Moyo et al. (2018) utilizaram os algoritmos de regressão logística multinomial, árvore de decisão e *nayve-bayes* para prever o tempo de fixação de profissionais de diferentes categorias em regiões desassistidas na África do Sul. O problema foi de classificação, pois os pesquisadores recodificaram a variável dependente em quatro categorias: permanência abaixo de um ano, entre um e dois anos, entre dois e três anos, acima de três anos. Os três algoritmos tiveram desempenho parecido, com a regressão logística tendo um valor um pouco maior de acurácia de 47% e AUC-ROC de 0.66. O estudo teve um número de atributos limitado – nacionalidade, estado civil, categoria profissional e gênero – e não registrou a importância das variáveis.

**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 anteriores. 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.

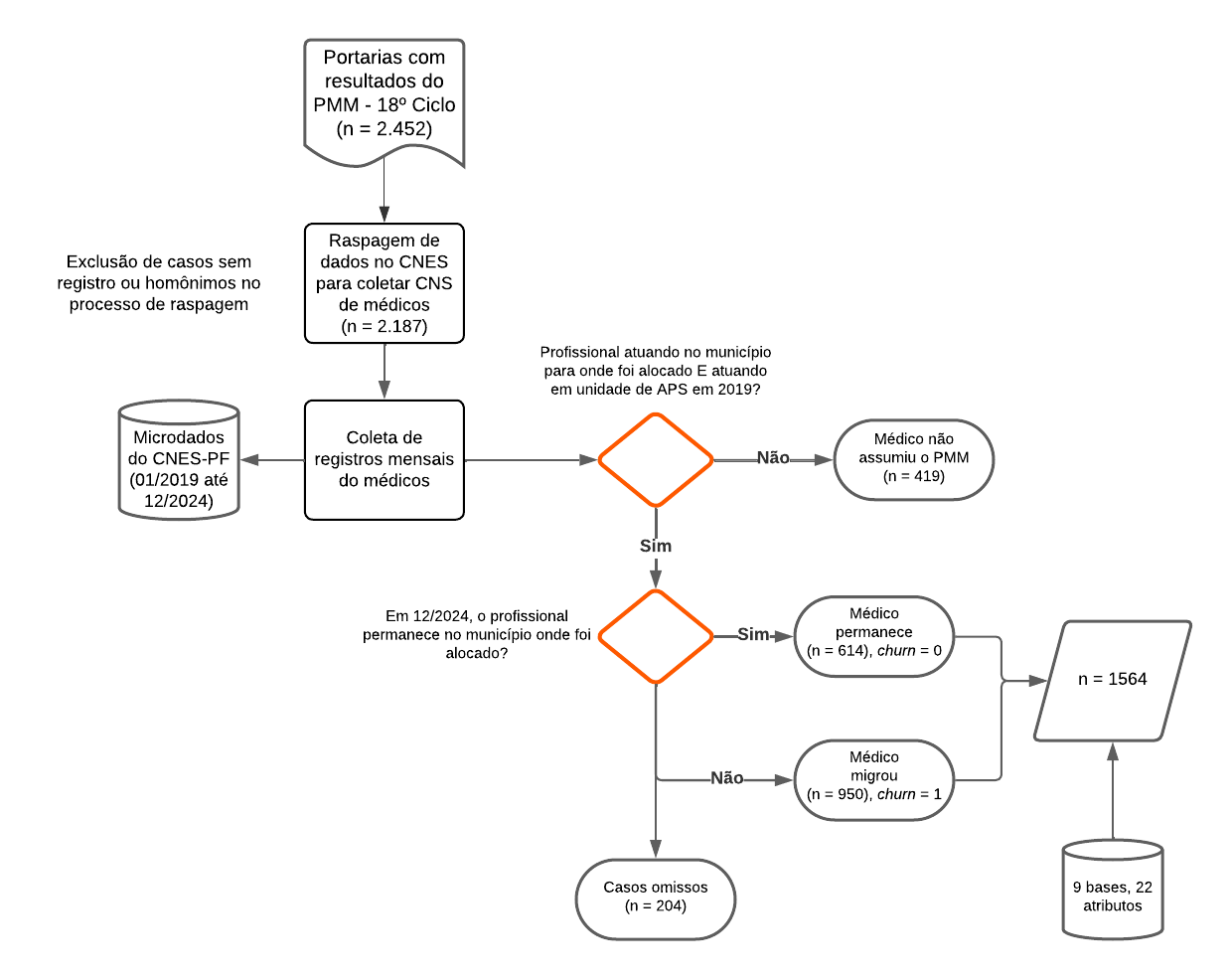
Em sequência, foi realizado um procedimento de automatização da coleta de dados dos profissionais selecionados por meio da página do Cadastro Nacional de Estabelecimentos de Saúde (CNES). Para este procedimento foi utilizada a biblioteca *RSelenium* na linguagem R. Um total de 256 médicos foram excluídos 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 | Idade, participação prévia no PMM | Portaria Nº 7, de 18 de junho de 2019; Portaria Nº 13 de 16 de agosto de 2019 |
| Estabelecimento | Quantidade média de agentes comunitários de saúde, enfermeiros, técnicos/auxiliares de enfermagem nas unidades de APS na qual o médico atuou | Cadastro Nacional de Estabelecimentos de Saúde (CNES) |
| 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, PIB per capita de 2019, Taxa de população ocupada | Instituto Brasileiro de Geografia e Estatística (IBGE) |
| Número de vagas em curso de medicina | Instituto Nacional de Estudos e Pesquisas Educacionais Anísio Teixeira (INEP) |
| Orçamento per capita em saúde, Percentual da população atendida por esgoto sanitário, taxa de homicídios, IDEB (anos finais e iniciais), 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, | Índice de Desenvolvimento Sustentável das Cidades – Brasil (IDSC-BR) |
| 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.

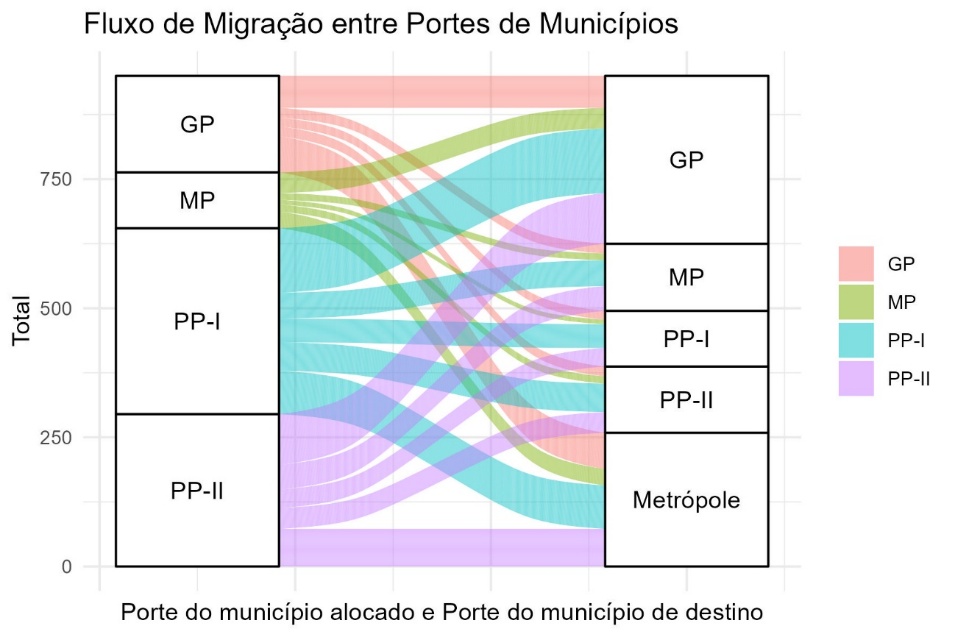
Quatro modelos foram treinados e avaliados: Regressão Logística, Árvore de Decisão, *Random Forest*, *XGBoost* e *LightGBM*. A otimização dos hiperparâmetros foi realizada com a biblioteca *Optuna*. O treinamento envolveu validação cruzada estratificada com *StratifiedKFold* em 5 partes. Devido ao desbalanceamento 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.

Todos os resultados foram desenvolvidos utilizando linguagens de programação e dados secundários e podem ser disponibilizados como material suplementar como parte das premissas da ciência aberta.

**Resultados**

A seguir, apresentamos os resultados em termos descritivos e depois passamos para a modelagem inferencial. O primeiro resultado importante a saber é que 950 (60%) dos 1564 médicos da amostra migraram ao final do programa. A figura 2 mostra os principais fluxos considerando a classificação de porte do município, segundo classificação do IBGE, para onde foi alocado no PMM (esquerda) e o município para onde migrou (direita).

Figura 2 - Fluxos de migração



Fonte: elaborado pelos autores

Na figura é possível observar que os municípios de pequeno porte 1 (PP-I), com menos de 20 mil habitantes, e os de pequeno porte 2 (PP-II), entre 20 mil e 50 mil habitantes foram as alocações do PMM mais frequentes dentro da amostra. Os fluxos migratórios mais prevalentes foram: saindo de PP-I para grande porte (GP[[1]](#footnote-1)) (13,15%); saindo de PP-II para GP (10,31%); saindo de PP-I para metrópole (8,84%); saindo de PP-II para metrópole (7,68%); saindo de GP para metrópole (7,36%). Estes cinco fluxos juntos totalizam cerca de 47% das migrações.

A tabela 1 traz o resumo de algumas variáveis selecionadas, com um recorte sobre a permanência e a migração dos municípios.

Tabela 1 – Análise descritiva de variáveis selecionadas.

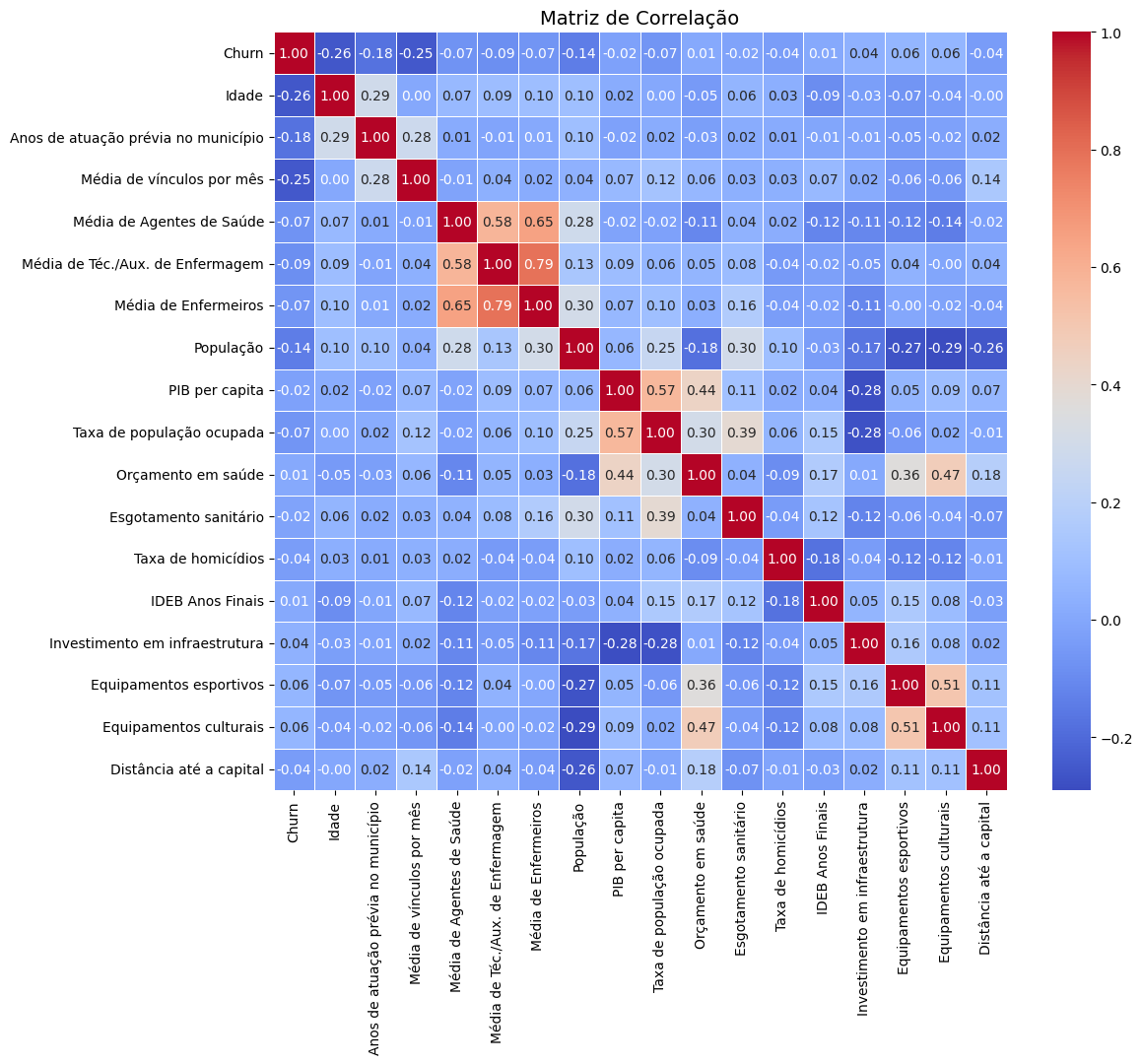
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Variáveis | | Permaneceram  n = 614 | Migraram  n = 950 | p-valor |
| Idade | | 43,9  (10,2) | 38,7  (8,84) | < 0,01a |
| Ano de formação | | 6,16  (9,29) | 3,72  (7,17) | < 0,01a |
| Média de vínculos de trabalho | | 1,31  (0,56) | 1,10  (0,25) | <0,01a |
| Quantidade média de enfermeiros | | 1,90  (1,43) | 1,72  (1,14) | <0,01ª |
| Quantidade média de técnicos/auxiliares de enfermagem | | 3,30  (3,06) | 2,83  (2,23) | <0,01a |
| Prorrogação | Prorrogaram | 476  (77,5%) | 211  (22,2%) | <0,01b |
| Não prorrogaram | 138  (22,5%) | 739  (77,8%) |
| Região do município para onde o profissional foi destinado | Centro-Oeste | 59  (9,60%) | 83  (8,73%) | <0,01b |
| Nordeste | 339  (55,21%) | 562  (59,15%) |
| Norte | 94  (15,30%) | 90  (9,47%) |
| Sudeste | 79  (12,86%) | 136  (14,31%) |
| Sul | 43  (7,00%) | 79  (8,31%) |

Fonte: elaborado pelos autores

Nota: a. Teste de Mann-Whitney; b. Teste de qui-quadrado; valores entre parêntese representam o desvio-padrão (variáveis numéricas) ou proporções (variáveis categóricas)

A figura 3 apresenta a matriz de correlação das variáveis numéricas, permitindo visualizar o grau de associação entre as variáveis do modelo. Os maiores níveis de correlação estão entre os profissionais de saúde (agente de saúde, enfermeiros e técnicos e auxiliares de enfermagem) que variam de 58% a 79%. Algumas variáveis a nível municipal também apresentam grau de associação, taxa de população ocupada e PIB per capita com 57% e número de equipamentos esportivos e culturais por 100 mil habitantes com 51%.

Figura 3 - Matriz de correlação



Fonte: elaborado pelos autores

Na parte de modelagem preditiva, adotou-se uma estratégia de construção de quatro modelos. O primeiro (M1) possuía apenas variáveis a nível do indivíduo. No segundo (M2) foram adicionadas os atributos relacionados aos estabelecimentos e vínculos de trabalho. O terceiro (M3) foram incluídas variáveis relacionadas ao município. Por fim, o último modelo (M4) manteve todos os atributos anteriores e só adicionou variável prorrogação. Os algoritmos que obtiveram os melhores desempenhos foram a *Regressão Logística* (M1), e *Random Forest* (M2, M3 e M4).

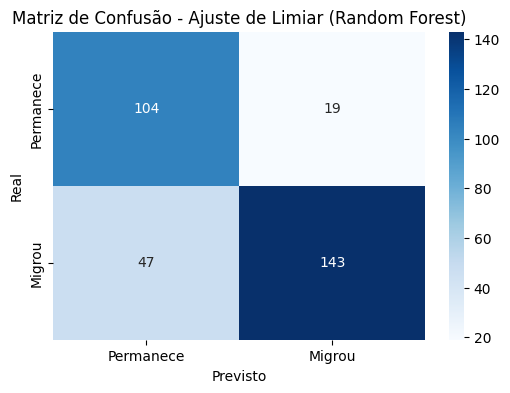
Tabela 2 – Resultados de modelagem preditiva

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | M1 | M2 | M3 | M4 |
| Melhor algoritmo | Regressão  Logística | Random Forest | Random Forest | Random Forest |
| Acurácia | 0.665 | 0.697 | 0.709 | 0.792 |
| Precisão | 0.715 | 0.751 | 0.759 | 0.861 |
| Recall | 0.742 | 0.747 | 0.763 | 0.784 |
| F1 Score | 0.729 | 0.749 | 0.761 | 0.821 |
| ROC AUC | 0.695 | 0.758 | 0.781 | 0.857 |

Fonte: elaborado pelos autores

A figura 4 apresenta a matriz de confusão do modelo M4. A matriz de confusão é gerada, reunindo informações quanto ao número de predições que o modelo conseguiu acertar e quantos foram preditos de modo equivocado. No eixo horizontal tem-se a classificação do algoritmo e no vertical os valores reais. Portanto, dos 123 médicos da base de teste que permaneceram nas cidades designadas (a soma do primeiro e segundo quadrante), o modelo classificou 104 corretamente e 19 como migraram incorretamente. O modelo classificou corretamente a migração de 143 casos dos 190 que de fato migraram.

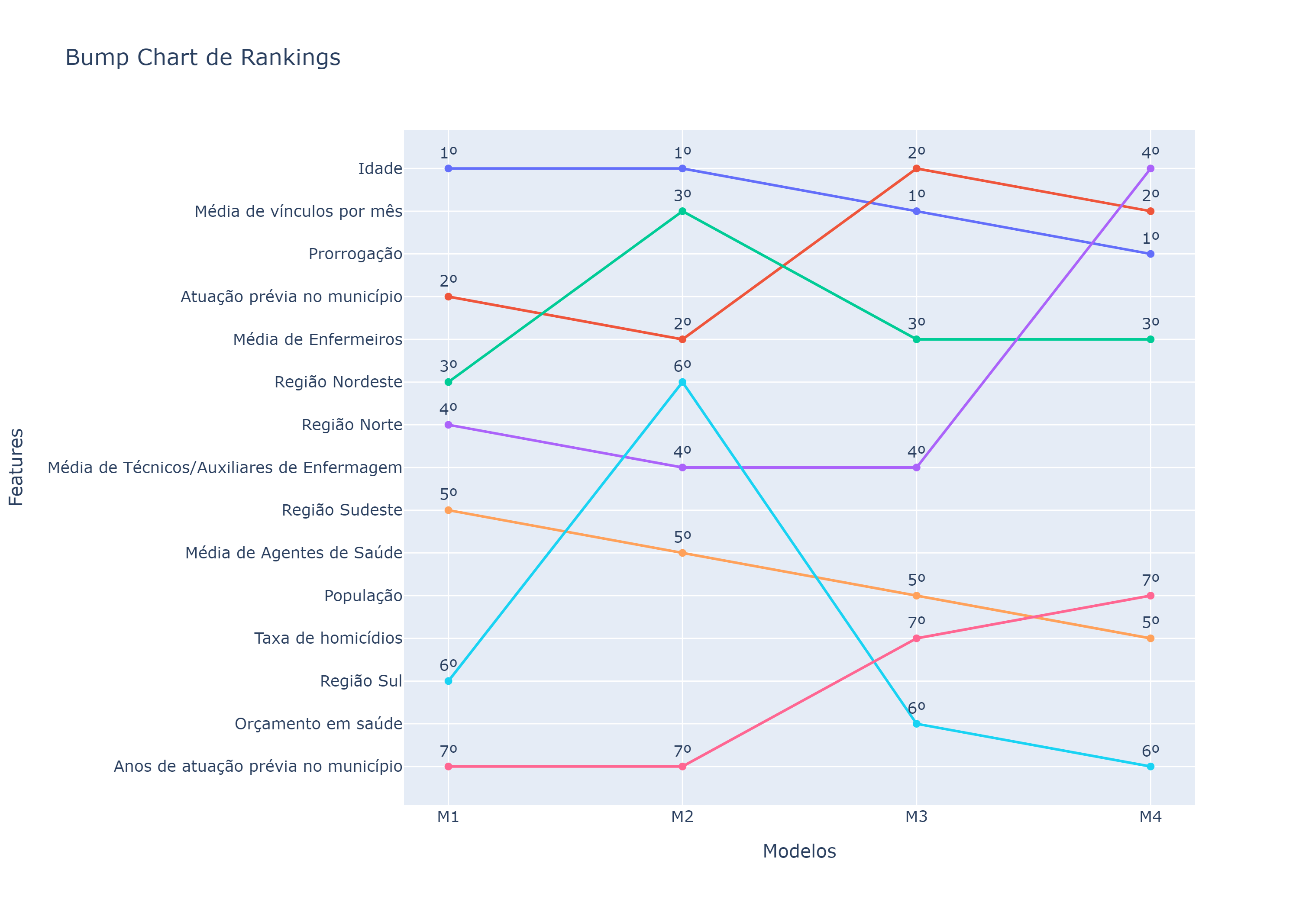
Figura 4 - Matriz de confusão



Fonte: elaborado pelos autores

A figura 5 apresenta as sete variáveis que exerceram maior importância para a previsão dos modelos M1 a M4. Entre os destaques está a idade que é a variável que esteve entre as mais importantes em todos os modelos: 1º lugar (M1 e M2), 2º lugar (M3) e 4º lugar (M4). Anos de atuação prévia no município é outra variável a nível indivíduo que persistiu no ranking até o modelo final, 7º lugar (M1 e M2) e 6º lugar (M4). A média de vínculos por mês esteve entre as mais importantes, 3º lugar (M2), 1º lugar (M3) e 2º lugar (M4). A média de enfermeiros nas unidades de APS onde o profissional atuou também foi importante em todos os modelos: 2º lugar (M2) e 3º lugar (M3 e M4). Ao inserir as variáveis relacionados aos municípios, tem-se: a população em 5º lugar (M3) e 7º lugar (M4); e a taxa de homicídio em 7º lugar (M3) e 5º lugar (M4). E por último, ao inserir prorrogação no modelo final (M4), aparece como a variável mais importante para predição do modelo, demonstrando que a decisão de adesão ao edital de prorrogação é um indicativo de permanência ou migração.

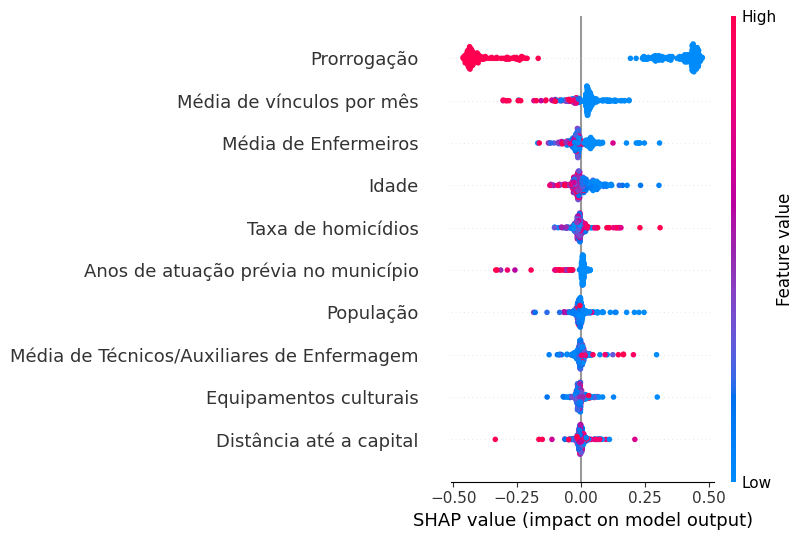
Figura 5 - Importância das variáveis (M1 ao M4)



Fonte: elaborado pelos autores

Por fim, a figura 6 traz o gráfico resumido do SHAP. Este recurso é utilizado para agregar uma camada de explicação ao modelo, permitindo compreender o porquê o modelo está tomando determinadas decisões. Diferente do gráfico da figura 5, aqui é possível verificar a direção dos relacionamentos. Algumas variáveis são destaque: o ato de prorrogar se associa à fixação; maior número de vínculos em outros estabelecimentos além da APS também tende a se associar à retenção; profissionais com idade mais elevada tende a fixar; uma maior média de enfermeiros nas unidades de APS onde atuam está associada a maior fixação.

Figura 6 - Explicabilidade do modelo



Fonte: elaborado pelos autores

**Discussão**

Os resultados da presente pesquisa se mostraram promissores, com métricas satisfatórias de previsão. Comparando com investigações prévias, observa-se que os resultados foram superiores a um estudo que contou com desfecho de fixação a nível regional (Moyo et al., 2018) e se aproximou de alguns estudos que empregaram análises de rotatividade em contexto organizacional (Kim et al., 2023; Lopez et al., 2023; Y. Xu et al., 2023).

Analisando as variáveis de importância, observa-se que a idade é uma variável que foi importante desde em todos os experimentos (M1, M2, M3 e M4). A idade também figurou entre os atributos mais importantes em estudos prévios (Kim et al., 2023; Lopez et al., 2023; Y. Xu et al., 2023).

Outra variável que se mostrou importante tanto em análises preditivas prévias (Kim et al., 2023; Y. Xu et al., 2023) quanto em levantamentos de fatores associados foi a renda (Soares, 2022; Steinmetz et al., 2014; K. Xu et al., 2024). Estudos prévios mostraram que os rendimentos possuem papel moderador entre o engajamento do trabalho e a intenção de sair (K. Xu et al., 2024). A remuneração é fator motivador que pode reduzir descontentamento e, consequentemente, minimizar a rotatividade (Liu et al., 2023). Ainda que no presente estudo, a renda não tenha sido aferida, pois seria necessário ter dados sobre as demais relações de trabalho, temos uma variável que pode ser considerada uma *proxy* da renda, que é o número médio de vínculos.

Esta variável foi uma das mais importantes para a previsão, indicando que maior quantidade de vínculos pode estar ligada a maior retenção. No entanto, esse ponto merece atenção, pois, se por um lado, ter vínculos de trabalho pode significar maior renda, por outro, o profissional acaba tendo uma carga de trabalho mais elevada. A carga de trabalho acima da média foi observada como uma variável relacionada à rotatividade (Steinmetz et al., 2014).

Estudos preditivos sobre fixação da força de trabalho em saúde com apoio de machine learning ainda estão

Estudos preditivos 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**

Os resultados mostram que é possível criar um modelo com algo grau de assertividade para prever da fixação de profissionais do PMM. Esse tipo de apoio gerencial poderia trazer alguns benefícios para a execução da política, como melhor definição de critérios para participantes do programa, melhor alocação de profissionais e adoção de medidas antecipadas para minimizar as chances de migração. Os resultados foram satisfatórios, ainda que a quantidade de variáveis seja restrita àquelas disponíveis publicamente. Pode-se supor que as métricas preditivas podem alcançar valores ainda melhores caso exista a disponibilidade de um número maior de atributos (ex.: estado civil, local de nascimento, local de formação) e observações de outros ciclos.

Além das contribuições relacionadas ao modelo preditivo, o estudo trouxe dois achados importantes. Primeiro, com base na amostra deste estudo, há uma grande evasão daqueles profissionais dos municípios. Os fluxos mostram que depois do programa, uma grande parcela vai para municípios de grande porte ou metrópoles. Portanto, enquanto houver o estímulo por parte do programa, pode haver garantia de preenchimento das lacunas de profissionais nos lugares de carência de médicos. Todavia, na ausência do recurso, o problema poderá ressurgir.

O segundo achado diz respeito às variáveis que foram mais importantes para a previsão. Constata-se que variáveis relacionadas à equipe de enfermagem, oportunidades de trabalho complementares ao PMM e condições da comunidade também devem ser alvo de ações por parte do programa a fim de garantir maior fixação dos profissionais nas regiões mais vulneráveis. Há de se reconhecer que alguns destes estímulos possuem maior complexidade, dado que são dependentes de outras pastas (ex.: educação e segurança), e que possivelmente mudariam apenas a longo prazo. Todavia, a inação pode levar ao mesmo problema mencionado anteriormente: na ausência do recurso do programa, fatores de fixação do local não seriam suficientes para reter o profissional.

Estudos passados trouxeram evidências que demonstram a importância do programa. No entanto, ao longo dos seus mais de dez anos, continua sendo um programa reativo e emergencial. No momento da sua concepção, havia eixos que previam ações de longo prazo destinadas à promoção de uma infraestrutura das unidades e formação de pessoal em vazios assistenciais para garantir maior sustentabilidade na oferta de profissionais. Porém, tais medidas foram impactadas por turbulências políticas ou execuções imprecisas. Além disso, os resultados também mostraram que o problema da retenção perpassa por outras questões, como a disponibilidade de outras categorias profissionais, oportunidades de trabalho e condições ambientais. Logo, o aprimoramento da política demanda ações integradas sob uma perspectiva multiprofissional e relacionadas a outras pastas, como segurança, desenvolvimento regional, educação, dentre outras.

Este estudo possui algumas limitações. A primeira repousa sobre a qualidade dos dados, que é uma crítica frequente nos estudos sobre gestão do trabalho na saúde. Bases como o CNES-PF podem apresentar inconsistências, especialmente em relação à atualização. Todavia, apesar desta observação, é um conjunto de dados amplamente utilizado em estudos que analisam aspectos relacionados à força de trabalho em saúde.

A segunda limitação diz respeito à avaliação de apenas um ciclo do programa. Não foi possível utilizar 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.

**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

Kim, S. K., Kim, E. J., Kim, H. K., Song, S. S., Park, B. N., & Jo, K. W. (2023). Development of a Nurse Turnover Prediction Model in Korea Using Machine Learning. *Healthcare (Switzerland)*, *11*(11). https://doi.org/10.3390/healthcare11111583

Kumar, S., & Clancy, B. (2021). Retention of physicians and surgeons in rural areas-What works? *Journal of Public Health (United Kingdom)*, *43*(4), E689–E700. https://doi.org/10.1093/pubmed/fdaa031

Liu, Y., Duan, Y., & Guo, M. (2023). Turnover intention and its associated factors among nurses: a multi-center cross-sectional study. *Frontiers in Public Health*, *11*. https://doi.org/10.3389/fpubh.2023.1141441

Lopez, K., Li, H., Paek, H., Williams, B., Nath, B., Melnick, E. R., & Loza, A. J. (2023). Predicting physician departure with machine learning on EHR use patterns: A longitudinal cohort from a large multi-specialty ambulatory practice. *PLoS ONE*, *18*(2 February). https://doi.org/10.1371/journal.pone.0280251

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

Moyo, S., Doan, T. N., Yun, J. A., & Tshuma, N. (2018). Application of machine learning models in predicting length of stay among healthcare workers in underserved communities in South Africa. *Human Resources for Health*, *16*(1). https://doi.org/10.1186/s12960-018-0329-1

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/

Soares, C. (2022). Análise dos fatores de Atração e retenção de profissionais médicos da estratégia da saúde da família na região oeste de Minas Gerais. *APS EM REVISTA*, *4*(1), 12–18. https://doi.org/10.14295/aps.v4i1.233

Steinmetz, S., Vries, D. H. D., & Tijdens, K. G. (2014). Should I stay or should I go? The impact of working time and wages on retention in the health workforce. *Human Resources for Health*, *12*(1). https://doi.org/10.1186/1478-4491-12-23

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

Wakerman, J., Humphreys, J., Russell, D., Guthridge, S., Bourke, L., Dunbar, T., Zhao, Y., Ramjan, M., Murakami-Gold, L., & Jones, M. P. (2019). Remote health workforce turnover and retention: What are the policy and practice priorities? In *Human Resources for Health* (Vol. 17, Issue 1). BioMed Central Ltd. https://doi.org/10.1186/s12960-019-0432-y

Xu, K., Lei, L., Guo, Z., Liu, X., Shi, Y., Han, G., Lin, K., Cai, W., Lu, C., Li, X., Li, Y., & Peng, K. (2024). Turnover intention among healthcare workers in Shenzhen, China: the mediating effect of job satisfaction and work engagement. *BMC Health Services Research*, *24*(1). https://doi.org/10.1186/s12913-024-11872-6

Xu, Y., Park, Y., Park, J. D., & Sun, B. (2023). Predicting Nurse Turnover for Highly Imbalanced Data Using the Synthetic Minority Over-Sampling Technique and Machine Learning Algorithms. *Healthcare (Switzerland)*, *11*(24). https://doi.org/10.3390/healthcare11243173

1. Entre 100 mil e 900 mil [↑](#footnote-ref-1)