

# **O USO DO MODELO DE REGRESSÃO DE COX PARA ESTIMAR FATORES PROGNÓSTICOS NA SOBREVIDA DE PACIENTES ONCOLÓGICOS NO ESTADO DO AMAZONAS: UM ESTUDO RETROSPECTIVO**

**THE USE OF THE COX REGRESSION MODEL TO ESTIMATE PROGNOSTIC  
FACTORS IN THE SURVIVAL OF CANCER PATIENTS IN THE STATE OF  
AMAZONAS: A RETROSPECTIVE STUDY**

Ciências Exatas e da Terra, Ciências da Saúde • 09/08/2026

REGISTRO DOI: [10.70773/revistatopicos/786321204](https://doi.org/10.70773/revistatopicos/786321204)

---

Bianca de Oliveira Holanda<sup>1</sup>

Carlos Cezar da Costa Negreiros<sup>2</sup>

Alex Leal Mota<sup>3</sup>

---

## RESUMO

O câncer constitui um dos principais desafios da saúde pública brasileira, e a identificação de fatores prognósticos é essencial para orientar estratégias de prevenção, diagnóstico e tratamento. Este artigo estima os fatores prognósticos associados à sobrevida de pacientes oncológicos no estado do Amazonas, no período de 2010 a 2018. Trata-se de estudo de coorte retrospectiva construído a partir de 30.732 registros do Registro Hospitalar de Câncer, analisados no software R. A sobrevida global foi descrita pelo estimador de Kaplan-Meier e comparada entre grupos pelo teste de Log-Rank; as covariáveis sociodemográficas idade ao diagnóstico, raça/cor, estado civil e escolaridade foram avaliadas por modelos de riscos proporcionais de Cox univariados e multivariados, com resultados expressos em razão de risco (HR) e intervalo de confiança de 95%. Os sítios tumorais mais frequentes foram colo do útero (C539 e D069; n = 6.852), mama (C509; n = 2.827) e próstata (C61; n = 2.297), para os quais os modelos foram ajustados. A idade avançada e a baixa escolaridade destacaram-se como os principais determinantes de pior prognóstico, enquanto o estado civil exerceu influência moderada e a raça/cor não apresentou associação estatisticamente significativa após o ajuste. Os achados reforçam a necessidade de políticas públicas regionais que incorporem determinantes sociais e ampliem o acesso ao diagnóstico precoce e ao tratamento especializado na região amazônica.

**Palavras-chave:** Análise de sobrevivência; Modelo de riscos proporcionais de Cox; Neoplasias; Fatores prognósticos; Amazonas.

## ABSTRACT

Cancer is one of the main public health challenges in Brazil, and the identification of prognostic factors is essential to guide prevention, diagnosis and treatment strategies. This article estimates the

prognostic factors associated with the survival of cancer patients in the state of Amazonas between 2010 and 2018. It is a retrospective cohort study based on 30,732 records from the Hospital Cancer Registry, analysed with the R software. Overall survival was described by the Kaplan-Meier estimator and compared between groups by the Log-Rank test; the sociodemographic covariates age at diagnosis, race/skin colour, marital status and educational level were assessed through univariate and multivariate Cox proportional hazards models, with results expressed as hazard ratios (HR) and 95% confidence intervals. The most frequent tumour sites were cervix uteri (C539 and D069; n = 6,852), breast (C509; n = 2,827) and prostate (C61; n = 2,297), for which the models were fitted. Advanced age and low educational level stood out as the main determinants of a worse prognosis, whereas marital status exerted a moderate influence and race/skin colour showed no statistically significant association after adjustment. The findings reinforce the need for regional public policies that incorporate social determinants and expand access to early diagnosis and specialised treatment in the Amazon region.

**Keywords:** Survival analysis; Cox proportional hazards model; Neoplasms; Prognostic factors; Amazonas.

## 1. INTRODUÇÃO

O câncer representa um dos maiores desafios globais para a saúde pública, sendo responsável por elevados índices de morbidade e mortalidade. Nos últimos anos, a incidência e a prevalência dessa doença aumentaram significativamente e, segundo o Instituto Nacional de Câncer, são estimados 704 mil novos casos anuais no Brasil durante o triênio 2023-2025, com expectativa de ultrapassar 25 milhões de novos casos no país até 2030. Especificamente no

Amazonas, são esperados 15 mil novos casos no mesmo período (INCA, 2022).

Diante desse cenário, compreender os fatores que influenciam a sobrevida dos pacientes com câncer é crucial para aprimorar as estratégias de prevenção, diagnóstico e tratamento, além de fornecer subsídios que impactam diretamente os padrões da prática clínica e a formulação de políticas públicas de saúde. Os fatores prognósticos, definidos como características ou variáveis associadas ao curso clínico de uma doença, têm papel determinante na relação entre incidência e mortalidade. Essa relação reflete a gravidade de cada tipo de câncer, sendo influenciada tanto por características biológicas do tumor, como tamanho, grau de diferenciação e envolvimento de linfonodos, quanto por fatores individuais do paciente, como idade, gênero e estado de saúde geral.

Na era da informação, a manipulação, a análise e a interpretação de dados tornaram-se parte essencial da saúde pública. Desde o entendimento das doenças e o desenvolvimento de novos tratamentos até o diagnóstico e o gerenciamento de pacientes individuais, o uso de dados e tecnologia tornou-se parte integral da assistência à saúde. Nesse sentido, a conversão dessa abundância de informações em conhecimento útil é essencial para a prestação de cuidados de alta qualidade ao paciente, haja vista que a democratização da análise de dados, impulsionada pelo movimento de software de código aberto, está permitindo que uma ampla gama de pessoas se envolva nesse processo, sem a necessidade de software especializado caro (Harrison; Pius, 2020). Nesse contexto, a análise de sobrevivência emerge como uma ferramenta estatística poderosa para investigar a relação entre variáveis prognósticas e desfechos clínicos ao longo de um intervalo de tempo.

Uma das abordagens mais amplamente utilizadas para esse tipo de análise é o modelo de regressão de Cox, também conhecido como modelo semiparamétrico de riscos proporcionais (Cox, 1972). Esse modelo permite estimar os efeitos de múltiplas covariáveis sobre a probabilidade de ocorrência de um evento, como o óbito, sem a necessidade de assumir uma distribuição específica para o tempo de sobrevivência. Considerado semiparamétrico, ele pressupõe que as covariáveis influenciam multiplicativamente o risco, constituindo ferramenta robusta para o estudo de fatores prognósticos em câncer (Carvalho *et al.*, 2011).

Portanto, o presente estudo analisa os fatores prognósticos que influenciam a sobrevida de pacientes oncológicos no Amazonas entre os anos de 2010 e 2018, utilizando dados provenientes do Instituto Nacional de Câncer. Por meio da aplicação do modelo de regressão de Cox, espera-se identificar padrões até então não percebidos, contribuindo para a compreensão do comportamento da doença na região amazônica e fornecendo subsídios para o aprimoramento das estratégias de cuidado, das práticas clínicas e das políticas públicas.

Além de seu impacto científico, este estudo pode gerar repercussões significativas no contexto social e clínico, ao promover uma gestão mais eficiente e abrangente do câncer, ampliando a qualidade de vida dos pacientes afetados por essa patologia tão desafiadora. Por meio de uma metodologia robusta e embasada em dados, busca-se não apenas avançar o conhecimento científico sobre o câncer, mas também abrir novas perspectivas para pesquisas futuras e intervenções mais eficazes, direcionando esforços para a melhoria do cuidado oncológico no Amazonas.

## **1.1. Objetivos**

O objetivo geral consiste em estimar os fatores prognósticos de diferentes tipos de câncer na população amazonense no período de 2010 a 2018, por meio do modelo de riscos proporcionais de Cox. Como objetivos específicos, o estudo busca:

- a. conhecer o número de casos de câncer na população amazonense no período de 2010 a 2018;
- b. inferir a incidência de diferentes tipos de câncer no Amazonas;
- c. verificar quais são os principais fatores prognósticos de diferentes tipos de câncer no Amazonas.

## **2. FUNDAMENTAÇÃO TEÓRICA**

A análise de sobrevivência é uma abordagem estatística desenvolvida para estudar o tempo até a ocorrência de um evento de interesse, como óbito, recidiva de doença ou falha de um sistema. Esse tipo de análise é amplamente utilizado na pesquisa oncológica, na qual se busca compreender os fatores que influenciam a sobrevida de pacientes com câncer. Dentre as características principais desses dados, destaca-se a variável tempo, que mede o intervalo entre o início do acompanhamento e a ocorrência do evento. Além disso, há a presença de censura, que ocorre quando não é possível observar o tempo completo até o evento. Outro aspecto importante são as covariáveis, ou variáveis explanatórias, que representam fatores capazes de influenciar diretamente a sobrevida dos indivíduos, sendo fundamentais para a modelagem dos dados.

Essa área de estudo existe, sobretudo, devido à necessidade de lidar com dados censurados, uma limitação que impede a aplicação de métodos estatísticos tradicionais. A censura ocorre quando o evento não pode ser observado em sua totalidade, pois o estudo pode ser encerrado antes de sua ocorrência ou os indivíduos podem sair do acompanhamento. Em estudos de câncer, os dados são censurados se o paciente ainda estiver vivo ao final do período de observação ou falecer por causas não relacionadas ao evento de interesse (Bustamante-Teixeira; Faerstein; Latorre, 2002). Dessa forma, a análise de sobrevivência se torna essencial, pois possibilita estudar o comportamento dos dados mesmo na presença de informações incompletas e censuradas.

Existem diferentes tipos de censura, sendo a censura à direita a mais comum. Nesse caso, o evento não é observado durante o período do estudo. A censura à direita pode ser subdividida em três categorias principais. A primeira é a censura fixa, em que se define uma janela de tempo específica para o acompanhamento, assumindo que o evento de interesse ocorrerá dentro desse intervalo. A segunda é a censura baseada na quantidade de falhas, em que o estudo se encerra assim que um número predeterminado de eventos é observado, como um número específico de óbitos. Já a terceira subdivisão, conhecida como censura à direita aleatória, ocorre quando há uma janela fixa de tempo, mas os indivíduos podem sair do estudo, falecer por outras causas ou não apresentar o evento de interesse (Carvalho *et al.*, 2011). Além disso, existem também a censura à esquerda, quando o evento ocorre antes do início do acompanhamento, e a censura intervalar, que ocorre quando o evento se dá dentro de um intervalo de tempo, mas o momento exato permanece desconhecido.

A função de sobrevivência é uma das principais ferramentas na análise de sobrevivência, pois estima a probabilidade de um indivíduo sobreviver além de um tempo específico  $T > t$ . Complementarmente, a função de risco descreve a taxa de falha instantânea em um determinado tempo, dado que o indivíduo sobreviveu até aquele momento. Já a função de risco acumulado mede o risco cumulativo ao longo do tempo, fornecendo uma visão mais abrangente do comportamento dos eventos no período estudado. Essas funções são fundamentais para entender não apenas a probabilidade de ocorrência do evento, mas também os momentos em que ele é mais provável (Deo; Deo; Sundaram, 2021).

Para modelar esses dados, existem três classes principais de métodos. Os modelos paramétricos assumem uma distribuição específica para os tempos de sobrevivência, enquanto os não paramétricos, como o estimador de Kaplan-Meier, não fazem suposições sobre a distribuição dos dados. Por outro lado, os modelos semiparamétricos, como o modelo de regressão de Cox, combinam características de ambos e são amplamente utilizados devido à sua flexibilidade. Proposto por Cox (1972), o modelo de regressão de Cox permite analisar o efeito das covariáveis sobre a sobrevivência sem a necessidade de especificar uma distribuição para o tempo até o evento. A equação geral do modelo é expressa como:

$$h(t|\mathbf{X}) = h_0(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p)$$

em que  $h(t|\mathbf{X})$  é a função de risco no tempo  $t$  para um vetor de covariáveis  $\mathbf{X}$ ,  $h_0(t)$  é a função de risco basal e  $\exp(\beta)$  representa a razão de risco (HR), que quantifica o impacto das covariáveis no tempo de sobrevivência. A razão de risco é interpretada como o fator multiplicativo pelo qual o risco de ocorrência do evento muda a

cada unidade de incremento na covariável. Quando  $HR > 1$ , há aumento no risco associado à covariável; quando  $HR < 1$ , o risco diminui (Patil; Dessai, 2019).

A avaliação do modelo de Cox é fundamental para garantir sua validade. Nesse sentido, a principal suposição do modelo é a proporcionalidade dos riscos, que determina que a relação entre as covariáveis e a taxa de falha deve permanecer constante ao longo do tempo. Métodos gráficos, como a análise dos resíduos de Schoenfeld, e testes estatísticos são amplamente utilizados para verificar essa suposição. Caso a proporcionalidade dos riscos não seja atendida, o modelo pode necessitar de modificações ou de abordagens alternativas (Deo; Deo; Sundaram, 2021).

### **3. METODOLOGIA**

Trata-se de estudo observacional analítico, de coorte retrospectiva, conduzido com dados do Instituto Nacional de Câncer referentes ao estado do Amazonas, abrangendo o período de 2010 a 2018. Os dados foram analisados com o software R-Project 4.3.3, no qual as variáveis foram selecionadas com base em análises univariadas e multivariadas, permanecendo no modelo final apenas aquelas com  $p < 0,05$ . A razão de risco (HR) foi estimada para identificar os fatores prognósticos, e o ajuste do modelo foi avaliado por meio da análise dos resíduos de Cox-Snell (Carvalho *et al.*, 2011).

As curvas de Kaplan-Meier foram utilizadas para descrever a sobrevivência ao longo do tempo, permitindo visualizar a probabilidade de sobrevida em diferentes momentos do acompanhamento, e o teste de Log-Rank foi empregado para comparar os tempos de sobrevivência entre os grupos definidos

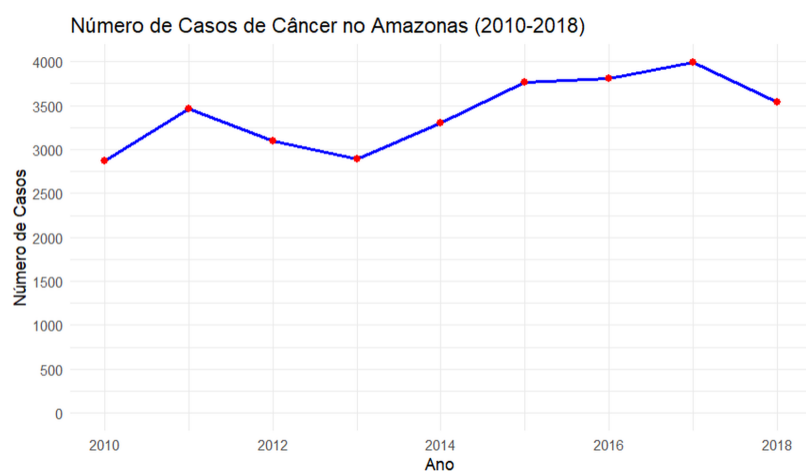
pelas covariáveis sociodemográficas: idade ao diagnóstico, raça/cor, estado civil e escolaridade. Para o modelo de regressão de Cox, a significância das covariáveis foi avaliada por meio dos testes estatísticos gerados automaticamente no ajuste — razão de verossimilhança, escore (Log-Rank) e Wald —, assegurando maior robustez às estimativas. A suposição de riscos proporcionais foi verificada pelo teste de Schoenfeld; quando violada, a covariável correspondente foi incluída no modelo como fator de estratificação.

## **4. RESULTADOS E DISCUSSÃO**

### **4.1. Perfil dos Registros Oncológicos no Amazonas**

Como parte do objetivo específico de conhecer o número de casos de câncer na população amazonense entre 2010 e 2018, foram identificados 30.732 registros de pacientes com diagnóstico oncológico no estado do Amazonas, conforme dados provenientes do Registro Hospitalar de Câncer (RHC). A Figura 1 apresenta a evolução anual desses casos, evidenciando oscilações ao longo do período. Observa-se crescimento progressivo a partir de 2013, com pico máximo registrado em 2017, ano em que o número de casos atingiu 3.990. Em contrapartida, o ano de 2010 representou o menor registro da série, com 2.866 casos. Essa variação pode refletir tanto a ampliação da cobertura de notificação quanto alterações no acesso da população amazonense aos serviços de saúde oncológica e no diagnóstico.

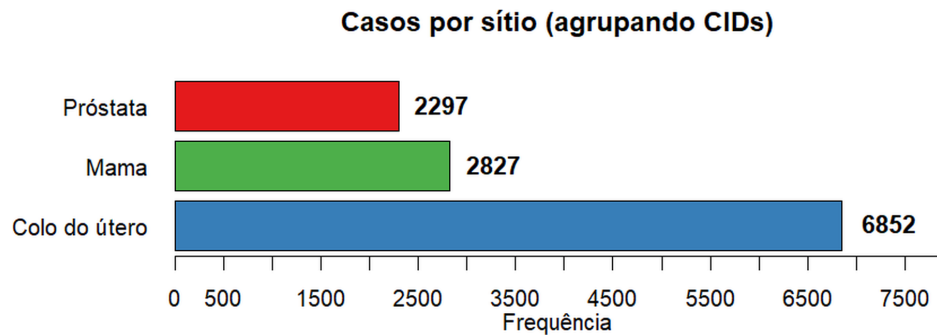
**Figura 1.** Número de casos de câncer no Amazonas (2010-2018)



Fonte: Elaborada pelos autores com base em dados do RHC/INCA.

Com base nos dados analisados, foi possível identificar os tipos de câncer com maior frequência no estado do Amazonas entre 2010 e 2018. A Figura 2 evidencia as classificações de doença (CID) mais incidentes nesse período. A neoplasia intraepitelial cervical grau III não especificada, ou neoplasia *in situ* do colo do útero (CID D069), geralmente considerada lesão precursora do câncer cervical, ocupa o primeiro lugar em número de casos, com 3.542 registros. Em seguida, a neoplasia maligna do colo do útero não especificada (CID C539) aparece com 3.310 ocorrências. Para fins de análise, ambas as categorias foram agrupadas sob a denominação C539 + D069, totalizando 6.852 casos. Em terceiro lugar está a neoplasia maligna da mama (CID C509), com 2.827 casos. Por fim, a neoplasia maligna da próstata (CID C61) fecha a lista, com 2.297 registros.

**Figura 2.** Cânceres mais frequentes no estado do Amazonas



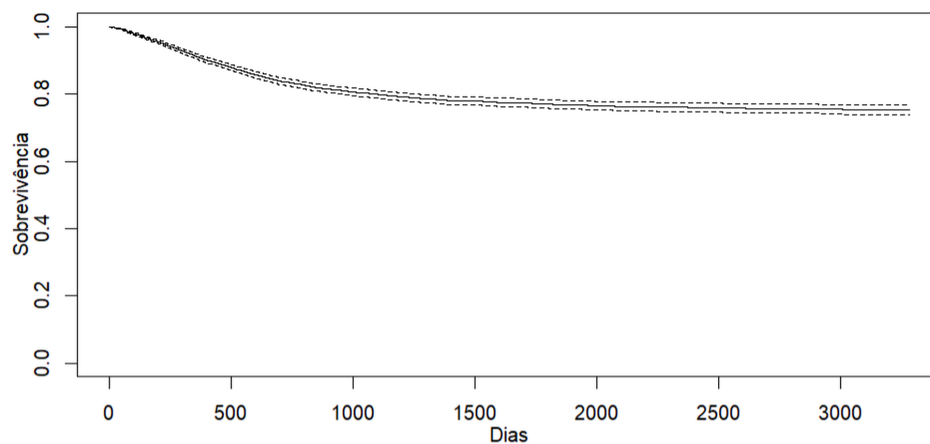
Fonte: Elaborada pelos autores com base em dados do RHC/INCA.

A seguir, o modelo de Cox foi ajustado considerando esses três sítios tumorais mais frequentes, com o objetivo de conhecer os fatores prognósticos influentes na sobrevida dos pacientes oncológicos.

#### **4.2. Carcinoma do Colo do Útero (D069 e C539)**

Foram analisadas 5.570 pacientes com câncer do colo do útero (CID-10: D069 e C539) registradas em Manaus, com seguimento até 31 de dezembro de 2018. O desfecho (evento de interesse) foi o óbito por câncer do colo do útero. Observaram-se 1.081 falhas e 4.489 censuras. As variáveis observadas foram idade ao diagnóstico (até 50 anos e 51 anos ou mais), raça/cor (branca, preta e outras), estado civil (solteira, casada e outras) e escolaridade (sem escolaridade, ensino fundamental, ensino médio, ensino superior e sem informação). A Figura 3 mostra a curva de sobrevivência global estimada pelo estimador de Kaplan-Meier, com intervalo de confiança de 95%. Nota-se que essa curva tende a se estabilizar entre 75% e 80% após aproximadamente 3.000 dias (cerca de oito anos). Esse comportamento evidencia que, apesar dos óbitos registrados, proporção significativa das pacientes permanece viva mesmo após longos períodos de acompanhamento. A utilização conjunta desses códigos abrange desde lesões precursoras até neoplasias invasivas, o que impacta diretamente os desfechos observados.

**Figura 3.** Curva de sobrevivência global (Kaplan-Meier) para neoplasias do colo uterino



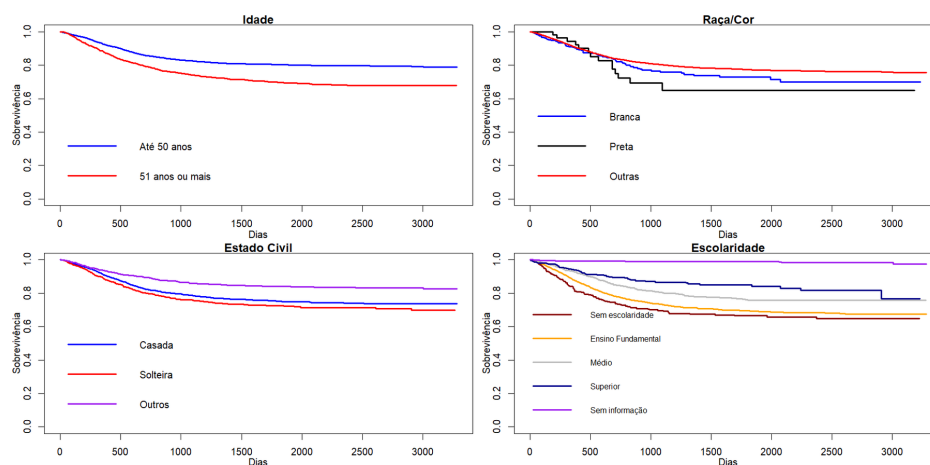
Fonte: Elaborada pelos autores.

A Figura 4 mostra as curvas de sobrevivência de Kaplan-Meier para as covariáveis observadas. No painel superior esquerdo, referente à idade, verifica-se que as pacientes com até 50 anos apresentam tendência de maior sobrevida ao longo do tempo em comparação àquelas com 51 anos ou mais. No painel superior direito, referente à raça/cor, observa-se pior prognóstico entre pacientes autodeclaradas pretas em relação às brancas e ao grupo classificado como outras raças. A curva das pacientes brancas apresenta maior estabilidade, enquanto a das pretas evidencia queda mais acentuada da sobrevivência ao longo do tempo observado. O painel inferior esquerdo, referente ao estado civil, indica que mulheres casadas apresentam maior sobrevida em relação às solteiras. Por fim, o painel inferior direito, referente à escolaridade, mostra que pacientes sem escolaridade e com ensino fundamental apresentam as piores taxas de sobrevida, enquanto as com nível superior alcançam maior probabilidade de sobrevivência. O grupo sem informação destaca-se por apresentar sobrevida superior.

Com o objetivo de comparar os tempos de sobrevivência entre os grupos, foi realizado o teste de Log-Rank. Os resultados indicaram

diferenças estatisticamente significativas ao nível de 5% para as variáveis idade, estado civil e escolaridade, mas não para raça/cor. Isso significa que, para idade, estado civil e escolaridade, os grupos apresentam tempos de sobrevivência significativamente diferentes. Por outro lado, do ponto de vista estatístico, não houve evidência suficiente para afirmar que a variável raça/cor influencia o tempo de sobrevivência.

**Figura 4.** Curvas de sobrevivência de Kaplan-Meier para os dados de câncer do colo do útero



Fonte: Elaborada pelos autores.

Na análise de regressão de Cox univariada, todas as covariáveis, com exceção de raça/cor ( $p = 0,10$ ), mostraram-se fatores prognósticos significativamente associados à sobrevida das pacientes com câncer do colo do útero ( $p < 0,05$ ). A suposição de proporcionalidade dos riscos foi violada apenas para a variável escolaridade, sugerindo que o efeito dessa covariável sobre o risco de morte varia ao longo do tempo. O modelo estimou razão de risco de 1,64 (IC95%: 1,45-1,85;  $p < 0,001$ ) para mulheres com 51 anos ou mais, indicando risco de morte 64% maior em comparação com aquelas com até 50 anos. Em relação ao grupo de mulheres brancas, as mulheres pretas apresentaram HR de 1,23 (IC95%: 0,69-2,18;  $p = 0,484$ ), sugerindo tendência a maior risco de morte, enquanto o grupo classificado

como outras apresentou HR de 0,82 (IC95%: 0,65-1,05;  $p = 0,115$ ), indicando tendência a menor risco. No entanto, ambas as associações não atingiram significância estatística. Além disso, o modelo apresentou baixo poder discriminativo para essa variável (concordância = 0,506). Comparadas às mulheres casadas, aquelas solteiras apresentaram risco de morte 16% maior (HR = 1,16; IC95%: 1,01-1,33;  $p = 0,036$ ), enquanto as mulheres classificadas como outras (viúvas, divorciadas e demais situações) apresentaram risco 38% menor (HR = 0,62; IC95%: 0,53-0,72;  $p < 0,001$ ).

No modelo de regressão de Cox multivariado, isto é, considerando simultaneamente as covariáveis idade ao diagnóstico, estado civil e raça/cor, e estratificado por escolaridade, verificou-se que a idade avançada manteve-se como preditor independente de menor sobrevida. Pacientes com 51 anos ou mais apresentaram risco de óbito 28% superior em relação àquelas com até 50 anos (HR = 1,28; IC95%: 1,12-1,47;  $p < 0,001$ ). Tais dados sugerem que mulheres mais velhas podem apresentar diagnóstico tardio, maior carga de comorbidades e menor acesso a estratégias de rastreamento precoce (INCA, 2022). O estado civil também se mostrou associado de forma significativa ao desfecho. Em comparação às mulheres casadas, as solteiras apresentaram aumento de 20% no risco de morte (HR = 1,20; IC95%: 1,04-1,37;  $p = 0,011$ ) e aquelas classificadas no grupo outros (viúvas, separadas judicialmente e com informação ausente) tiveram risco 39% maior (HR = 1,39; IC95%: 1,17-1,65;  $p < 0,001$ ). Esses achados corroboram evidências de que o suporte emocional, a presença de um parceiro e o contexto familiar contribuem positivamente para a adesão ao tratamento e para o enfrentamento da doença, com impacto direto na sobrevida oncológica (Pinquart; Duberstein, 2010). Embora a literatura aponte desigualdades raciais no acesso à saúde no Brasil (INCA, 2022), a variável raça/cor não

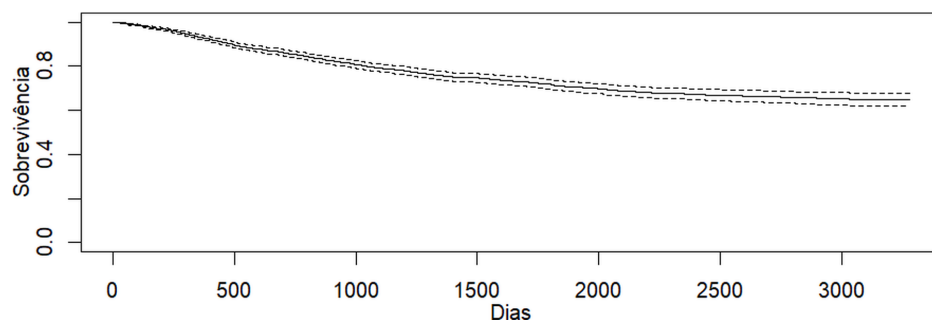
apresentou associação estatisticamente significativa no modelo final. Mulheres pretas exibiram HR de 1,08 (IC95%: 0,61-1,91;  $p = 0,79$ ) e o grupo outras, HR de 0,92 (IC95%: 0,74-1,14;  $p = 0,53$ ), ambos em relação às mulheres brancas. A variável escolaridade foi incluída no modelo como fator de estratificação, em razão da violação da suposição de proporcionalidade dos riscos detectada pelo teste de Schoenfeld ( $p = 0,032$ ). Essa abordagem permitiu que o modelo estimasse funções de risco basal específicas para cada estrato de escolaridade, evitando vieses nas estimativas e preservando a robustez dos resultados das demais covariáveis. O modelo multivariado atendeu à suposição de riscos proporcionais (teste global de Schoenfeld:  $p = 0,14$ ) e apresentou ajuste global significativo (teste de log-verossimilhança:  $p < 0,001$ ). A estatística de concordância ( $C = 0,549$ ) indicou capacidade discriminatória moderada, sugerindo que, embora as variáveis incluídas contribuam para explicar parte da variação na sobrevida, outros fatores não mensurados — como estágio clínico ao diagnóstico, tipo de tratamento e presença de comorbidades — podem influenciar de maneira relevante o prognóstico dessas pacientes.

#### **4.3. Neoplasia Maligna da Mama (C509)**

Foi avaliada a sobrevida de 2.492 pacientes com câncer de mama registradas em Manaus, acompanhadas até 31 de dezembro de 2018. Foram identificados 566 óbitos (eventos) e 1.926 censuras. As variáveis analisadas incluíram idade (até 50 anos; 51 anos ou mais), raça/cor (branca, preta, outras), estado civil (casada, solteira, outras) e escolaridade (sem escolaridade, fundamental, médio, superior, sem informação).

A Figura 5 mostra a curva de sobrevivência global para pacientes com câncer de mama (C509), estimada pelo método de Kaplan-Meier com intervalo de confiança de 95%. Observa-se queda progressiva da probabilidade de sobrevida ao longo do tempo, sobretudo nos primeiros anos após o diagnóstico. Com o prolongamento do acompanhamento, a curva tende a se estabilizar em torno de 65%, sugerindo que proporção expressiva das pacientes permanece viva mesmo após vários anos. Esse padrão reflete tanto o impacto dos óbitos relacionados ao câncer quanto a heterogeneidade clínica dos casos incluídos, que abrangem estágios distintos da doença.

**Figura 5.** Curva de sobrevivência global (Kaplan-Meier) para os dados de câncer de mama



Fonte: Elaborada pelos autores.

Na análise por idade, as curvas de Kaplan-Meier apresentadas na Figura 6 mostraram sobrevida semelhante entre mulheres de até 50 anos e aquelas com 51 anos ou mais. O teste de Log-Rank não foi estatisticamente significativo ( $p = 0,20$ ), e o modelo de Cox univariado confirmou ausência de associação estatisticamente significativa com o tempo de sobrevivência ( $HR = 0,89$ ;  $IC_{95\%}$ : 0,75-1,05;  $p = 0,15$ ). Isso reforça que a idade, isoladamente, não foi determinante da sobrevida, resultado consistente com a literatura

de análise de sobrevida, que enfatiza a necessidade de avaliar fatores clínicos e sociais em conjunto (Bustamante-Teixeira; Faerstein; Latorre, 2002).

A análise por raça/cor indicou menor sobrevivência para mulheres do grupo outras (pardas, indígenas, amarelas e sem informação) em comparação às brancas (Log-Rank  $p = 0,02$ ). A regressão de Cox revelou que a variável raça/cor está marginalmente associada ao tempo até o evento. Especificamente, pacientes classificadas como outras raças/cores apresentaram risco de ocorrência do evento 37,3% maior em comparação às pacientes brancas (HR = 1,373; IC95%: 1,046-1,803;  $p = 0,022$ ). Por outro lado, não foi observada diferença estatisticamente significativa entre pacientes pretas e brancas (HR = 1,063;  $p = 0,906$ ), sugerindo que, nesse conjunto de dados, a raça/cor preta não influencia significativamente o tempo de sobrevivência em relação à branca.

O estado civil apresentou associação significativa com a sobrevida. O grupo composto por mulheres solteiras teve risco de morte 27% maior em comparação às casadas (HR = 1,27; IC95%: 1,05-1,53;  $p = 0,014$ ), enquanto o grupo outras (viúvas, separadas, sem informação) apresentou risco 26% menor (HR = 0,74; IC95%: 0,59-0,92;  $p = 0,008$ ). O teste de Log-Rank global foi altamente significativo ( $p < 0,001$ ). A literatura demonstra que a rede de suporte social exerce efeito protetor no prognóstico oncológico, favorecendo a adesão terapêutica e o enfrentamento da doença (Pinquart; Duberstein, 2010).

Os resultados da regressão de Cox indicam que o nível de escolaridade está significativamente associado à sobrevivência de mulheres com câncer de mama. A escolaridade foi categorizada em

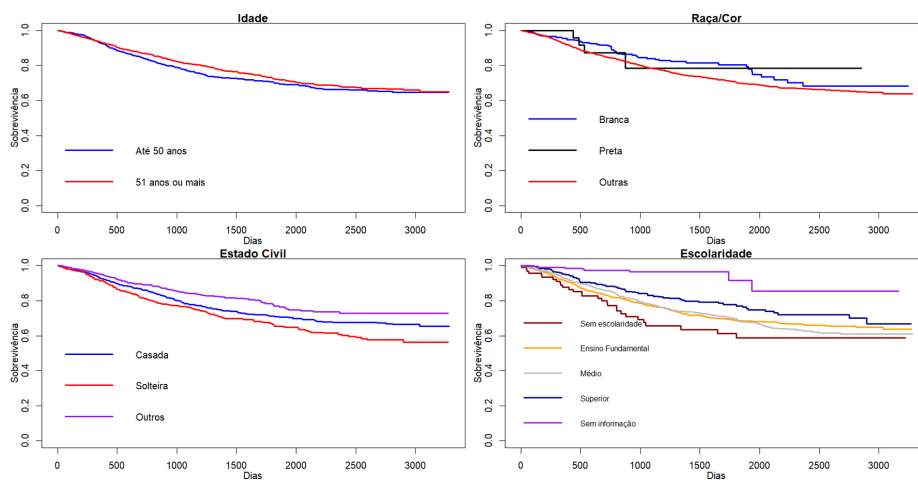
cinco níveis, tendo como grupo de referência as mulheres sem escolaridade formal. Observou-se tendência de diminuição do risco de morte com o aumento da escolaridade, conforme descrito a seguir:

- a. mulheres com ensino fundamental apresentaram risco de morte 26% menor em relação ao grupo sem escolaridade (HR = 0,74; IC95%: 0,50-1,08), embora essa associação não tenha atingido significância estatística ( $p = 0,118$ );
- b. para aquelas com ensino médio completo, o risco de morte foi 26% menor (HR = 0,74; IC95%: 0,50-1,10), também sem significância estatística ( $p = 0,137$ );
- c. mulheres com nível superior apresentaram redução de risco estatisticamente significativa de 46% em comparação àquelas sem escolaridade (HR = 0,54; IC95%: 0,35-0,84;  $p = 0,005$ );
- d. a categoria sem informação apresentou redução ainda maior do risco (HR = 0,12; IC95%: 0,056-0,27;  $p < 0,0001$ ); no entanto, essa associação deve ser interpretada com cautela, pois a categoria pode incluir dados ausentes ou distorcidos, o que dificulta uma interpretação válida.

Esses achados são consistentes com a literatura, que indica que maiores níveis de escolaridade estão associados a melhores desfechos de saúde, incluindo a sobrevida no câncer. Isso pode refletir diferenças no acesso ao diagnóstico precoce, na adesão ao tratamento e na compreensão sobre a doença, além de melhores condições socioeconômicas e maior suporte social geralmente associados à escolaridade mais elevada. Do ponto de vista da saúde pública, os resultados reforçam a importância de considerar o nível

educacional como determinante social da saúde. Mulheres com menor escolaridade podem demandar estratégias específicas de rastreamento, acompanhamento e educação em saúde, visando mitigar desigualdades no tratamento e nos desfechos do câncer de mama.

**Figura 6.** Curvas de sobrevivência de Kaplan-Meier para os dados de câncer de mama



Fonte: Elaborada pelos autores.

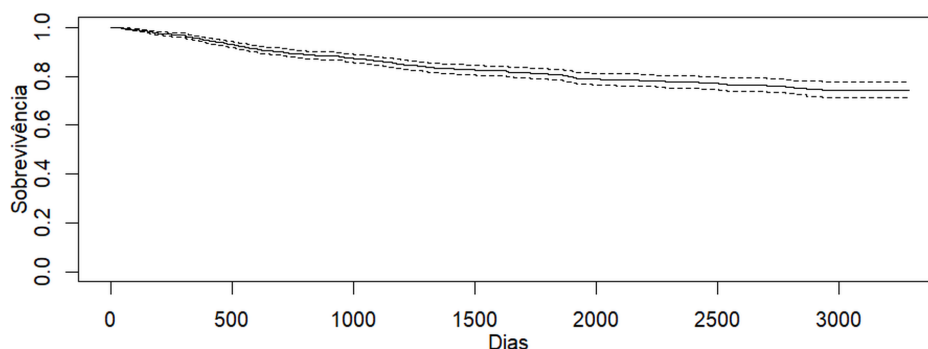
No modelo de Cox multivariado, que incluiu simultaneamente as variáveis idade, raça/cor, estado civil e escolaridade, o ajuste global foi estatisticamente significativo, mas o poder discriminatório mostrou-se apenas moderado ( $C = 0,586$ ). A suposição de riscos proporcionais não foi violada ( $p = 0,184$ ). Os resultados indicam que estado civil e escolaridade foram os fatores significativamente associados à sobrevida: pacientes solteiras apresentaram maior risco de evento em relação às casadas ou em união estável, enquanto aquelas com ensino superior tiveram risco quase 50% menor em comparação às pacientes sem escolaridade. Já idade e raça/cor não mostraram associação significativa com o desfecho no modelo ajustado.

#### 4.4. Neoplasia Maligna da Próstata (C61)

Foram analisados 1.875 pacientes com câncer de próstata (C61) registrados em Manaus e acompanhados até 31 de dezembro de 2018. Nesse período, foram observados 296 óbitos (eventos), com 18 observações excluídas por ausência de dados.

A Figura 7 apresenta a curva de sobrevivência global dos pacientes com câncer de próstata, estimada pelo método de Kaplan-Meier com intervalo de confiança de 95%. Observa-se queda gradual da sobrevida ao longo do seguimento, sobretudo nos primeiros anos após o diagnóstico. A probabilidade de sobrevida inicia próxima de 100% e declina para aproximadamente 98% no primeiro ano (cerca de 365 dias), situa-se em torno de 95% após três anos e estabiliza-se próxima de 90% entre oito e nove anos de acompanhamento. Esse comportamento indica que, embora haja mortalidade associada à doença, a maioria dos pacientes sobrevive por longos períodos após o diagnóstico, o que reflete tanto o curso clínico relativamente mais indolente de muitos tumores de próstata quanto a possibilidade de detecção precoce em estágios menos agressivos. Ainda assim, a progressiva redução da curva evidencia o impacto da doença na mortalidade ao longo do tempo.

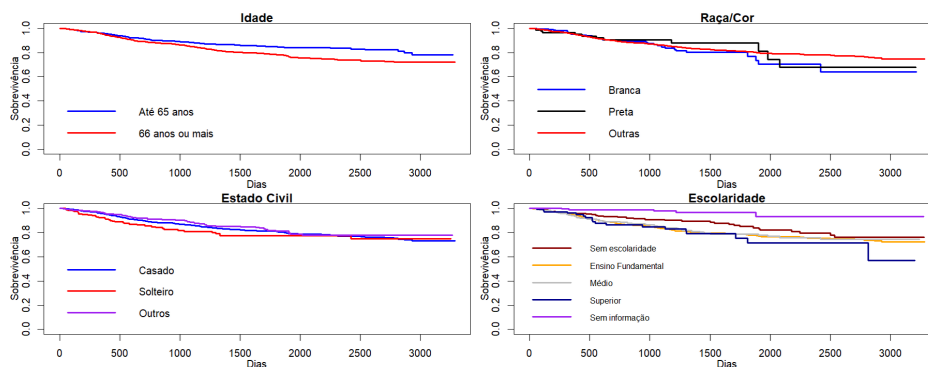
**Figura 7.** Curva de sobrevivência global (Kaplan-Meier) para neoplasias de próstata



Fonte: Elaborada pelos autores.

A Figura 8 apresenta as curvas de sobrevivência de Kaplan-Meier para as covariáveis observadas. Para a variável idade, a categorização adotada foi de até 65 anos versus 66 anos ou mais. O teste de Log-Rank evidenciou diferenças significativas entre os grupos ( $p = 0,004$ ). O modelo de Cox univariado confirmou que pacientes com 66 anos ou mais apresentaram risco de morte 43% maior em relação aos mais jovens ( $HR = 1,43$ ;  $IC95\%: 1,12-1,83$ ;  $p = 0,004$ ). A escolha desse ponto de corte justifica-se pelo fato de o câncer de próstata ser reconhecido como doença da terceira idade, com cerca de 75% dos diagnósticos ocorrendo em homens a partir dos 65 anos (INCA, 2022). Assim, a divisão reflete não apenas um marco epidemiológico amplamente utilizado em estudos de sobrevida, mas também um limiar biológico e social que distingue uma população de menor risco de outra em que a incidência e a mortalidade se elevam de forma expressiva (Bustamante-Teixeira; Faerstein; Latorre, 2002). Esse resultado reforça o impacto da idade avançada no prognóstico, frequentemente associada à presença de comorbidades, ao diagnóstico tardio e ao menor acesso a terapias curativas, conforme descrito por Carvalho *et al.* (2011).

**Figura 8.** Curvas de sobrevivência de Kaplan-Meier para os dados de câncer de próstata



Fonte: Elaborada pelos autores.

A variável raça/cor, categorizada em branca, preta e outras (parda, amarela, indígena e sem informação), não apresentou associação significativa. O teste de Log-Rank não mostrou diferenças estatísticas ( $p = 0,70$ ) e, no modelo de Cox, as categorias preta ( $HR = 0,83$ ;  $IC95\%: 0,42-1,64$ ;  $p = 0,58$ ) e outras ( $HR = 0,86$ ;  $IC95\%: 0,59-1,25$ ;  $p = 0,44$ ) não diferiram da branca. Esses achados sugerem ausência de efeito direto da raça/cor sobre a sobrevida, ainda que a literatura destaque o papel das desigualdades sociais e do acesso aos serviços de saúde como determinantes indiretos da mortalidade (Bustamante-Teixeira; Faerstein; Latorre, 2002).

O estado civil apresentou tendência, embora não significativa. Pacientes solteiros tiveram risco 22% maior de morte em comparação aos casados ( $HR = 1,22$ ;  $IC95\%: 0,86-1,74$ ;  $p = 0,26$ ), enquanto o grupo outros apresentou risco reduzido em 13% ( $HR = 0,87$ ;  $IC95\%: 0,66-1,15$ ;  $p = 0,32$ ). Apesar da ausência de significância estatística no modelo de Cox, o Log-Rank apontou valor próximo ao limiar ( $p = 0,07$ ). Esse achado é coerente com evidências de que o suporte social e familiar pode influenciar a adesão ao tratamento e melhorar o prognóstico (Pinquart; Duberstein, 2010).

A escolaridade foi a variável que mais impactou a sobrevida. O teste de Log-Rank revelou diferenças significativas entre os grupos ( $p <$

0,001). No modelo de Cox univariado, indivíduos com ensino fundamental (HR = 1,51; IC95%: 1,03-2,20; p = 0,03) e médio (HR = 1,59; IC95%: 1,01-2,51; p = 0,046) apresentaram risco maior de morte em comparação aos sem escolaridade. Pacientes com ensino superior também mostraram tendência a risco elevado (HR = 1,80; IC95%: 1,00-3,24; p = 0,051). O grupo sem informação apresentou risco reduzido (HR = 0,20; IC95%: 0,08-0,54; p = 0,001), possivelmente refletindo viés de registro. Diante desse resultado e considerando a literatura metodológica, a escolaridade foi utilizada como variável de estratificação, em consonância com as recomendações de Deo, Deo e Sundaram (2021) e de Patil e Dessai (2019).

No modelo de Cox multivariado, ajustando simultaneamente idade, raça/cor, estado civil e escolaridade, a idade igual ou superior a 66 anos permaneceu como fator de risco independente (HR = 1,48; IC95%: 1,14-1,91; p = 0,003). A escolaridade também manteve impacto significativo: ensino fundamental (HR = 1,51; p = 0,032), médio (HR = 1,59; p = 0,047) e superior (HR = 1,80; p = 0,051) associaram-se a maior risco de morte, enquanto o grupo sem informação apresentou efeito protetor espúrio (HR = 0,21; p = 0,001). Já raça/cor e estado civil não permaneceram como preditores independentes. O modelo global mostrou ajuste satisfatório e estatisticamente significativo (Log-Rank = 40,85; p < 0,001).

Em síntese, na população estudada, a idade avançada e a baixa escolaridade foram os principais determinantes sociodemográficos associados à menor sobrevida no câncer de próstata, enquanto raça/cor e estado civil não apresentaram efeito independente após o ajuste. Esses achados reforçam que fatores sociais — como acesso desigual ao diagnóstico precoce e ao tratamento especializado — precisam ser considerados no planejamento de políticas públicas

em oncologia (Bustamante-Teixeira; Faerstein; Latorre, 2002; Pinquart; Duberstein, 2010; Deo; Deo; Sundaram, 2021).

## **5. CONCLUSÃO**

O presente estudo identificou os fatores prognósticos associados à sobrevida de pacientes oncológicos no estado do Amazonas, utilizando o modelo de regressão de Cox em uma coorte retrospectiva. Foram analisados os principais sítios tumorais — colo do útero (C539 e D069), mama (C509) e próstata (C61) —, o que possibilitou avaliação abrangente dos padrões de sobrevivência e das variáveis sociodemográficas envolvidas.

De maneira geral, a idade avançada destaca-se como o fator mais relevante de pior prognóstico. Esse achado é consistente com evidências que associam o envelhecimento a maior carga de comorbidades, diagnóstico tardio e acesso limitado a terapias de alta complexidade (Carvalho *et al.*, 2011; Deo; Deo; Sundaram, 2021). Em contextos de vulnerabilidade social, como na Amazônia, tais barreiras tornam-se ainda mais expressivas, ampliando desigualdades já existentes.

A escolaridade mostra-se igualmente determinante, com efeito significativo nos cânceres de mama e de próstata, ainda que em direções distintas conforme o sítio tumoral, o que sugere a influência de vieses de registro e de padrões diferenciados de acesso ao rastreamento.

O estado civil, como indicador de suporte social, também demonstra relevância. Pacientes casados ou em união estável apresentam tendência a melhores desfechos, o que corrobora a literatura que aponta o suporte familiar e conjugal como fator protetor,

influenciando positivamente a adesão ao tratamento e o enfrentamento da doença (Pinquart; Duberstein, 2010).

Em relação à raça/cor, não se observou associação estatisticamente significativa após o ajuste nos modelos multivariados. Esse resultado não afasta a existência de barreiras estruturais, como a desigualdade no acesso ao diagnóstico precoce e ao tratamento especializado, que podem atuar de forma indireta sobre a mortalidade (Pinquart; Duberstein, 2010), e reforça a necessidade de estudos que incorporem variáveis clínicas e de acesso aos serviços.

A aplicação do modelo de Cox mostra-se fundamental, pois possibilita avaliar o efeito simultâneo de múltiplas covariáveis sem necessidade de pressupor distribuição paramétrica para os tempos de sobrevida (Cox, 1972; Carvalho *et al.*, 2011). A suposição de proporcionalidade dos riscos foi atendida nos modelos finais, garantindo a robustez das estimativas e reforçando a importância da análise de sobrevivência como ferramenta epidemiológica e de saúde pública (Patil; Dessai, 2019; Deo; Deo; Sundaram, 2021).

Como limitações, destacam-se a ausência, na base utilizada, de variáveis clínicas relevantes — estadiamento ao diagnóstico, tipo de tratamento e comorbidades —, a proporção de registros sem informação em covariáveis sociodemográficas e a capacidade discriminatória apenas moderada dos modelos ajustados. Estudos futuros que incorporem essas informações podem qualificar as estimativas apresentadas.

Em síntese, este estudo demonstra que idade avançada, escolaridade e estado civil são determinantes relevantes da sobrevida oncológica no Amazonas, em um contexto marcado por

desigualdades sociais e estruturais. Os achados evidenciam a urgência de estratégias regionais que fortaleçam o rastreamento precoce, ampliem o acesso ao diagnóstico e ao tratamento especializado e incorporem determinantes sociais na formulação de políticas públicas. O uso do modelo de Cox reforça seu papel como ferramenta essencial para subsidiar decisões baseadas em evidências, contribuindo para um cuidado oncológico mais justo, equitativo e eficaz na região amazônica.

## REFERÊNCIAS BIBLIOGRÁFICAS

BUSTAMANTE-TEIXEIRA, M. T.; FAERSTEIN, E.; LATORRE, M. R. Técnicas de análise de sobrevida. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 18, n. 3, p. 579-594, 2002.

CARVALHO, M. S. *et al.* **Análise de sobrevivência**: teoria e aplicações em saúde. 2. ed. rev. e ampl. Rio de Janeiro: Editora Fiocruz, 2011. 432 p.

COX, D. R. Regression models and life-tables. **Journal of the Royal Statistical Society. Series B, Statistical Methodology**, v. 34, n. 2, p. 187-202, 1972.

DEO, S. V.; DEO, V.; SUNDARAM, V. Survival analysis – part 2: Cox proportional hazards model. **Indian Journal of Thoracic and Cardiovascular Surgery**, v. 37, n. 2, p. 229-233, 2021.

HARRISON, E.; PIUS, R. **R for health data science**. 1. ed. Nova Iorque: Chapman and Hall/CRC, 2020.

INCA. **Estimativa 2023**: incidência de câncer no Brasil. Rio de Janeiro: Instituto Nacional de Câncer, 2022. 160 p. Disponível em:

<https://www.inca.gov.br/publicacoes/livros/estimativa-2023-incidencia-de-cancer-no-brasil>. Acesso em: 1 mar. 2024.

PATIL, V.; DESSAI, S. Testing and interpreting assumptions of COX regression analysis. **Cancer Research, Statistics, and Treatment**, v. 2, n. 1, p. 108, 2019.

PINQUART, M.; DUBERSTEIN, P. R. Associations of social networks with cancer mortality: a meta-analysis. **Critical Reviews in Oncology/Hematology**, v. 75, n. 2, p. 122-137, 2010. DOI: <https://doi.org/10.1016/j.critrevonc.2009.06.003>.

---

<sup>1</sup> Discente do Curso de Medicina da Universidade Federal do Amazonas (UFAM). Endereço postal: Av. General Rodrigo Octavio Jordão Ramos, 1200 – Coroado I, Manaus – AM, 69067-005. E-mail: [acesse o artigo original para visualizar o e-mail](#).

<sup>2</sup> Discente do Curso de Engenharia de Software da Universidade Federal do Amazonas (UFAM). Endereço postal: Av. General Rodrigo Octavio Jordão Ramos, 1200 – Coroado I, Manaus – AM, 69067-005. E-mail: [acesse o artigo original para visualizar o e-mail](#).

<sup>3</sup> Docente da Universidade Federal do Amazonas (UFAM); orientador do projeto. Endereço postal: Av. General Rodrigo Octavio Jordão Ramos, 1200 – Coroado I, Manaus – AM, 69067-005. E-mail: [acesse o artigo original para visualizar o e-mail](#).