



O uso do modelo de regressão de Cox para estimar fatores prognósticos na sobrevida de pacientes oncológicos no estado do Amazonas: um estudo retrospectivo

Bianca de Oliveira Holanda¹, Carlos Cezar da Costa Negreiros², Alex Leal Mota³ (PQ)

¹Graduanda em Medicina pela Universidade Federal do Amazonas. Rua Afonso Pena, 1053, Centro, Manaus-AM, Brasil.

²Graduando em Engenharia de Software pela Universidade Federal do Amazonas. Av. Rodrigo Otávio Jordão Ramos, 6200, Coroado I, Manaus-AM, Brasil.

³Universidade Federal do Amazonas, Departamento de Estatística. Av. Rodrigo Otávio Jordão Ramos, 6200, Coroado I, Manaus-AM, Brasil.

* bianca.oliveira.holanda@hotmail.com

Palavras-Chave: análise de sobrevivência, dados de câncer, modelo de cox.

Introdução

O câncer permanece entre as principais causas de morbimortalidade no mundo. Para o triênio 2023–2025, o INCA estima 704 mil novos casos anuais no país, com cerca de 15 mil no estado do Amazonas [1], cenário que reforça a urgência de estratégias de prevenção, diagnóstico e tratamento pautadas por evidências. Nesse ínterim, a análise de sobrevivência oferece um arcabouço matemático capaz de quantificar, ao longo do tempo, a probabilidade de permanência sem o evento de interesse, ao mesmo tempo em que lida adequadamente com observações censuradas [2]. Em aplicações clínicas e epidemiológicas, o estimador não paramétrico de Kaplan–Meier tem sido amplamente utilizado para estimar a curva de sobrevivência [2], enquanto o teste de Log-Rank compara formalmente as curvas entre grupos [2]. Para avaliar o efeito simultâneo de múltiplas covariáveis sobre o risco, o modelo de Cox constitui a abordagem de referência [3, 4], pois não supõe uma distribuição paramétrica para o tempo até o evento e por causa da interpretabilidade via razão de riscos (HR). Em contextos de grandes distâncias geográficas e marcadas desigualdades estruturais, como na região amazônica, a mensuração do impacto de determinantes sociodemográficos — idade, escolaridade, raça/cor e suporte social — sobre a sobrevida é particularmente relevante para orientar políticas públicas e a priorização do cuidado [5, 6]. Assim, este estudo analisa fatores prognósticos associados à sobrevida de pacientes oncológicos no estado do Amazonas entre 2010 e 2018, empregando curvas de Kaplan–Meier, comparações por Log-Rank e o modelo de regressão de Cox, além disso, discute as implicações dos resultados para prática clínica e gestão.

Material e Métodos

Trata-se de um estudo de coorte retrospectivo, utilizando dados do Registro Hospitalar de Câncer (RHC/INCA) referentes a casos diagnosticados no estado do Amazonas entre 2010 e 2018, com seguimento até 31 de dezembro de 2018. O evento de interesse foi o óbito por câncer, sendo as censuras atribuídas ao término do período de seguimento, morte devido à outras causas ou à perda de

acompanhamento. O tempo foi mensurado em dias desde o diagnóstico até o evento ou censura. As análises concentraram-se nos sítios mais frequentes — colo do útero (C539 e D069), mama (C509) e próstata (C61). As covariáveis incluíram idade no diagnóstico (categorizada em $\leq 50/\geq 51$ anos para colo/mama e $\leq 65/\geq 66$ anos para próstata), raça/cor (branca, preta, outras), estado civil (casada(o), solteira(o), outros) e escolaridade (sem escolaridade, fundamental, médio, superior, sem informação). Procedeu-se a descrição inicial das frequências e tendência temporal dos casos, seguida de estimação da curva de sobrevivência por Kaplan–Meier com intervalos de confiança de 95% e comparações por Log-Rank. Em seguida ajustaram-se modelos de Cox uni e multivariados. A suposição de proporcionalidade dos riscos foi avaliada global e por covariável usando resíduos de Schoenfeld. Quando necessário, adotou-se estratificação para permitir riscos basais distintos por níveis de determinada variável, preservando a validade das estimativas dos demais coeficientes. O ajuste global foi testado por razão de verossimilhança, enquanto o poder discriminatório foi quantificado pelo C-index de Harrell.

Resultados e Discussão

Foram identificados 30.732 registros oncológicos na série histórica, com tendência de crescimento a partir de 2013 e pico em 2017, conforme podemos ver na Figura 1.

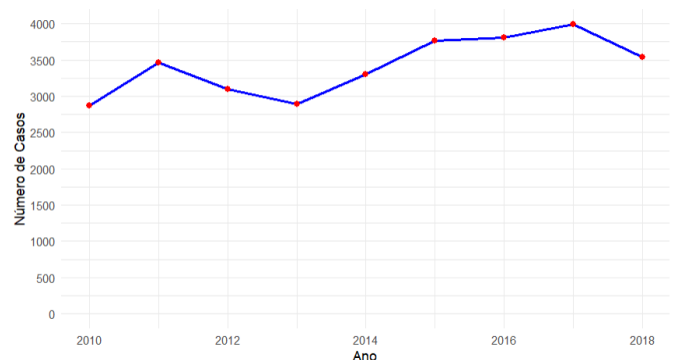


Figura 1: Número de casos de câncer no Amazonas (2010–2018).

A Figura 2 mostra que os sítios mais incidentes no período estudado foram colo do útero (C539 + D069, $n=6.852$), mama

(C509, n=2.827) e próstata (C61, n=2.297), motivando análises específicas por sítio.

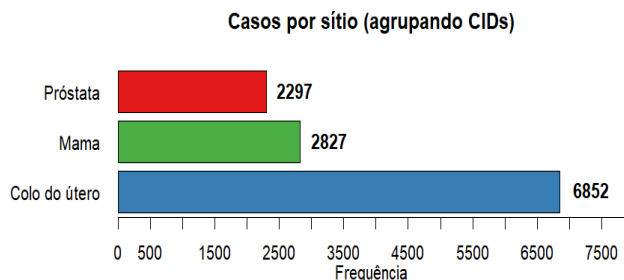


Figura 2: Cânceres mais frequentes no estado do Amazonas.

No conjunto de pacientes com câncer do colo do útero, a curva de sobrevivência global estimada pelo método de Kaplan–Meier apresentou estabilização entre 75% e 80%. Esse comportamento evidencia que, apesar dos óbitos registrados, uma proporção significativa das pacientes permanece viva mesmo após longos períodos de acompanhamento. O teste de Log-Rank mostrou diferenças significativas na sobrevida entre grupos por idade, estado civil e escolaridade, enquanto raça/cor não apresentou diferença estatística ao nível de 5%. Na modelagem de Cox univariada, idade ≥ 51 anos associou-se a maior risco de óbito ($HR \approx 1,64$), e o estado civil solteira mostrou HR discretamente acima de 1 frente a casada. Raça/cor não atingiu significância, mas exibiu tendências compatíveis com literatura sobre determinantes sociais [6]. No multivariado, com estratificação por escolaridade por violação da proporcionalidade, a idade manteve-se como preditor independente ($HR \approx 1,28$), e o estado civil permaneceu associado ao desfecho, sugerindo que suporte social e conjugal tende a favorecer adesão terapêutica e enfrentamento da doença [6]. O ajuste global foi significativo e o C-index indicou discriminação modesta, o que é esperado na ausência de variáveis clínicas cruciais como estadiamento, tratamento e comorbidades [7].

No câncer de mama, a curva de sobrevida global apresentou declínio mais acentuado nos primeiros anos, com estabilização em torno de 65% ao longo do seguimento. A idade, isoladamente, não discriminou tempos de sobrevida de forma significativa, um resultado coerente com a noção de que a idade atua em interação com fatores biológicos e de acesso [4]. A variável raça/cor mostrou associação marginal no grupo “outras” em comparação a brancas, com $HR > 1$ e intervalo de confiança excluindo a nulidade, enquanto não se observou diferença significativa para pretas frente a brancas. O estado civil exibiu efeito consistente, com maior risco para solteiras e menor risco para o grupo “outros” em relação às casadas. Observou-se gradiente educacional, com redução de risco acompanhando maior escolaridade, atingindo significância estatística no nível superior. A categoria “sem informação” apresentou HR artificialmente protetor, recomendando interpretação cautelosa por possível viés de registro. No modelo multivariado, escolaridade e estado civil mantiveram associação independente com o desfecho, enquanto idade e raça/cor perderam significância, e o poder discriminatório permaneceu moderado.

Para próstata, as curvas de Kaplan–Meier sugeriram alta sobrevida de longo prazo, aproximando-se de 90% aos oito a nove anos. A idade ≥ 66 anos esteve associada a maior risco tanto nos modelos univariados quanto no multivariado, reforçando o impacto da senescência, comorbidades e potenciais atrasos diagnósticos [3, 4]. Raça/cor e estado civil não permaneceram como preditores independentes após

ajuste, e a escolaridade diferenciou significativamente as curvas, com indicação de não proporcionalidade dos riscos. O ajuste global foi estatisticamente significativo e a discriminação situou-se em patamar moderado.

Tomados em conjunto, os achados indicam que idade avançada e menor escolaridade constituem determinantes consistentes de pior prognóstico, enquanto o estado civil casado/união estável tende a apresentar efeito protetor compatível com a literatura [6]. A raça/cor, por sua vez, exibe associações que se atenuam após ajuste por covariáveis sociodemográficas, sugerindo que diferenças observadas podem ser mediadas por determinantes sociais e barreiras de acesso [1, 6]. As análises de sensibilidade demonstraram estabilidade das estimativas frente à remoção de observações influentes, recategorização de idade e inclusão de termos dependentes do tempo quando necessários [8, 9]. Não se identificaram colinearidades relevantes, e os resíduos de Cox–Snell não sinalizaram falta grave de ajuste. Em termos de implicações, os resultados apoiam a priorização de estratégias de rastreamento organizado, navegação do paciente, redução de atrasos diagnóstico-terapêuticos e focalização em grupos vulneráveis por escolaridade e idade, além do investimento na qualidade e completude dos registros [6].

Conclusões

Verificou-se que idade avançada e baixa escolaridade se associam independentemente a pior sobrevida, enquanto o estado civil indica benefício compatível com o papel do suporte social; raça/cor apresentou associação limitada após ajuste, sugerindo mediação por determinantes sociais e acesso. A combinação de Kaplan–Meier, Log-Rank e regressão de Cox, com diagnóstico explícito dos pressupostos e, quando necessário, estratificação por variáveis não proporcionais, mostrou-se adequada para dados censurados e heterogêneos [3, 4, 8]. Tais evidências apoiam políticas públicas regionais que integrem determinantes sociais à gestão oncológica, ampliem o rastreamento e o acesso a diagnóstico e tratamento oportuno e qualifiquem a informação em saúde.

Referências

- [1] Instituto Nacional de Câncer (Brasil). Estimativa 2023: incidência de câncer no Brasil. Rio de Janeiro: INCA; 2022.
- [2] Bustamante-Teixeira MT, Faerstein E, Latorre MRDO. Técnicas de análise de sobrevida. Cad Saúde Pública. 2002;18(3):579–94.
- [3] Cox DR. Regression models and life-tables. J R Stat Soc B. 1972;34(2):187–220.
- [4] Carvalho MS, Andreozzi VL, Codeço CT, Campos DP, Barbosa MTS, Shimakura SE. Análise de sobrevivência: teoria e aplicações em saúde. 2ª ed. Rio de Janeiro: Fiocruz; 2011.
- [5] Harrison E, Pius R. R for Health Data Science. New York: Chapman and Hall/CRC; 2020.
- [6] Pinquart M, Duberstein PR. Associations of social networks with cancer mortality: a meta-analysis. Crit Rev Oncol Hematol. 2010;75(2):122–37.
- [7] Carvalho MS, Shimakura SE, Barreto ML. Modelagem estatística em epidemiologia: conceitos, métodos e aplicações. Cad Saúde Pública. 2005;21(2):159–72.
- [8] Patil V, Dessai S. Testing and interpreting assumptions of Cox regression analysis. Cancer Res Stat Treat. 2019;2(1):108–12.
- [9] Deo SV, Deo V, Sundaram V. Survival analysis – part 2: Cox proportional hazards model. Indian J Thorac Cardiovasc Surg. 2021;37:229–33.