

---

# Associação entre os marcadores CK19, EpCAM e CD44 e óbito por carcinoma hepatocelular: coorte retrospectiva

DOCUMENTO: SAR-2023-028-AN-v01

De: Felipe Figueiredo Para: Aline Nascimento

2023-11-30

## SUMÁRIO

1	LISTA DE ABREVIATURAS.....	2
2	CONTEXTO.....	3
2.1	Objetivos.....	3
3	METODOLOGIA.....	3
4	RESULTADOS.....	4
4.1	População do estudo e acompanhamento.....	4
4.2	Risco de óbito por CHC.....	11
5	OBSERVAÇÕES E LIMITAÇÕES.....	13
6	CONCLUSÕES.....	14
7	REFERÊNCIAS.....	14
8	APÊNDICE.....	15
8.1	Análise exploratória de dados.....	15
8.2	Estratégia de modelagem.....	17
8.2.1	Tentativas de especificação de modelos.....	17
8.2.2	Proporcionalidade de riscos.....	18
8.2.3	Estimativas de modelos convergentes.....	18
8.3	Análises associadas.....	19
8.4	Disponibilidade.....	19
8.5	Dados utilizados.....	20

---

# Associação entre os marcadores CK19, EpCAM e CD44 e óbito por carcinoma hepatocelular: coorte retrospectiva

## Histórico do documento

Versão	Alterações
01	Versão inicial

## 1 LISTA DE ABREVIATURAS

- CHC: carcinoma hepatocelular
- DP: desvio padrão
- HR: taxa de risco (*hazards ratio*)
- IC: intervalo de confiança

## 2 CONTEXTO

### 2.1 Objetivos

Identificar fatores de risco que alteram a associação entre marcadores de carcinoma hepatocelular e óbito em pacientes do INCA.

## 3 METODOLOGIA

Os procedimentos de limpeza de dados, desenho do estudo e métodos de análise usados neste relatório estão descritos no documento anexo **SAP-2023-028-AN-v01**.

Esta análise foi realizada utilizando-se o software R versão 4.3.0.

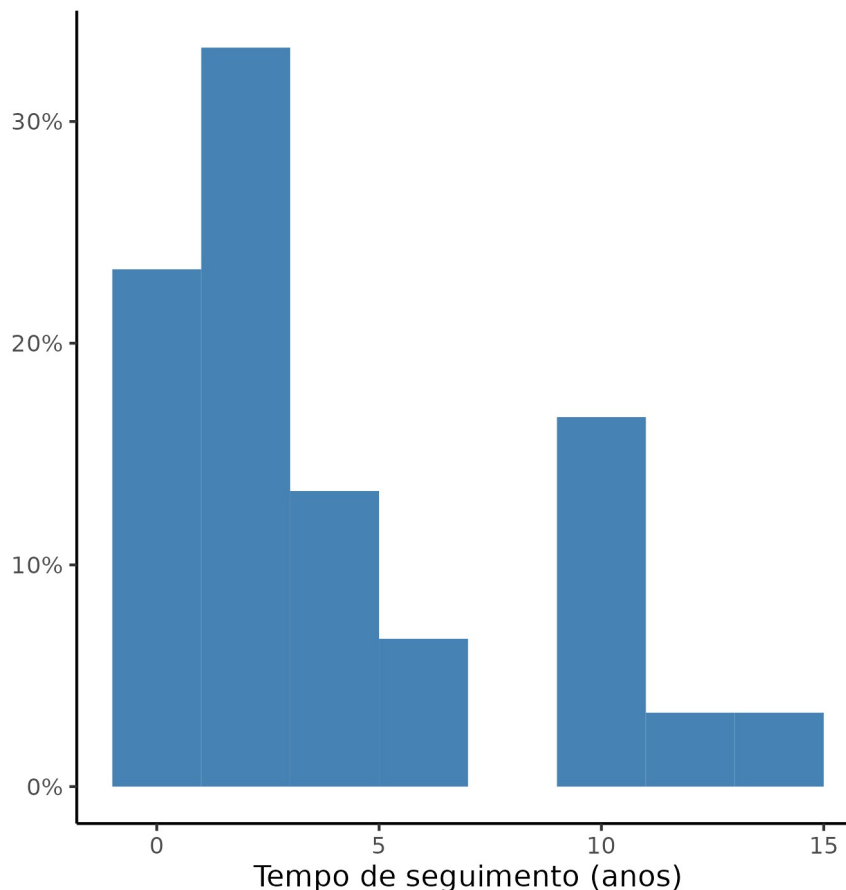
## 4 RESULTADOS

### 4.1 População do estudo e acompanhamento

A base de dados original possuía 31 observações. Após remoção de um participante duplicado 30 observações foram incluídas na análise.

A amostra de estudo foi composta por participantes com idades entre 25 e 77 anos, dos quais 70% tinham 60 anos ou mais no momento da inclusão na coorte, e 23 (77%) eram homens.

O tempo mediano de acompanhamento na coorte (Figura 1) foi de 2.61 anos, com amplitude entre 11 dias e 13.54 anos.



**Figura 1** Distribuição do tempo de seguimento na coorte.

As prevalências dos marcadores CK19, EpCAM e CD44 na coorte foram, respectivamente, 33%, 23% e 50%. A incidência de óbito por CHC foi 6 (20%), dos quais 3 (50%) eram CK19 positivos, 2 (33%) EpCAM positivos e 3 (50%) CD44 positivos.

Relatório de Análise Estatística (SAR)

A Tabela 1 mostra as características dos participantes por desfecho.

**Tabela 1** Características epidemiológicas e clínicas dos participantes, por desfecho.

Características dos participantes	Total, N = 30	Censurado, N = 12	CHC, N = 6	Outra, N = 12
<b>CK19, n (%)</b>	10 (33%)	1 (8.3%)	3 (50%)	6 (50%)
<b>EpCAM, n (%)</b>	7 (23%)	3 (25%)	2 (33%)	2 (17%)
<b>CD44, n (%)</b>	15 (50%)	8 (67%)	3 (50%)	4 (33%)
<b>Tempo de seguimento (anos), Mediana (Amplitude)</b>	2.6 (0.0 - 13.5)	5.0 (0.7 - 13.5)	2.4 (1.6 - 10.5)	1.0 (0.0 - 10.8)
<b>Sexo, n (%)</b>				
F	7 (23%)	5 (42%)	1 (17%)	1 (8.3%)
M	23 (77%)	7 (58%)	5 (83%)	11 (92%)
<b>Idade, Mediana (Amplitude)</b>	64 (25 - 77)	66 (34 - 77)	59 (25 - 70)	63 (28 - 69)
<b>Transplante, n (%)</b>	21 (70%)	8 (67%)	1 (17%)	12 (100%)
<b>Etiologia primária, n (%)</b>				
ALCOOL	3 (10%)	0 (0%)	0 (0%)	3 (25%)
DESC	1 (3.3%)	0 (0%)	1 (17%)	0 (0%)
HAI	1 (3.3%)	0 (0%)	0 (0%)	1 (8.3%)
NASH	5 (17%)	3 (25%)	1 (17%)	1 (8.3%)
VHB	2 (6.7%)	1 (8.3%)	1 (17%)	0 (0%)
VHC	18 (60%)	8 (67%)	3 (50%)	7 (58%)
<b>Etiologia secundária, n (%)</b>				

# Consultoria em Estatística Médica e Epidemiologia Clínica

## Relatório de Análise Estatística (SAR)

Características dos participantes	Total, N = 30	Censurado, N = 12	CHC, N = 6	Outra, N = 12
def alfa1 antitripsina	1 (50%)	1 (100%)	0 (NA%)	0 (0%)
VHB	1 (50%)	0 (0%)	0 (NA%)	1 (100%)
(Faltante)	28	11	6	11
<b>Terapia locorregional, n (%)</b>	17 (57%)	7 (58%)	4 (67%)	6 (50%)
<b>Método da terapia locorregional, n (%)</b>				
alcoholizacao	2 (12%)	0 (0%)	0 (0%)	2 (33%)
desconhecido	8 (47%)	4 (57%)	2 (50%)	2 (33%)
lipiodol	1 (5.9%)	0 (0%)	1 (25%)	0 (0%)
radiofrequencia	1 (5.9%)	0 (0%)	1 (25%)	0 (0%)
TACE	5 (29%)	3 (43%)	0 (0%)	2 (33%)
(Faltante)	13	5	2	6
<b>Momento da terapia locorregional, n (%)</b>				
antes	6 (75%)	1 (50%)	1 (50%)	4 (100%)
depois	2 (25%)	1 (50%)	1 (50%)	0 (0%)
(Faltante)	22	10	4	8
<b>Recidiva, n (%)</b>	7 (23%)	3 (25%)	4 (67%)	0 (0%)
<b>Quanto tempo apresentou recidiva (meses), n (%)</b>				
4	1 (14%)	1 (33%)	0 (0%)	0 (NA%)

# Consultoria em Estatística Médica e Epidemiologia Clínica

## Relatório de Análise Estatística (SAR)

Características dos participantes	Total, N = 30	Censurado, N = 12	CHC, N = 6	Outra, N = 12
12	3 (43%)	1 (33%)	2 (50%)	0 (NA%)
24	2 (29%)	1 (33%)	1 (25%)	0 (NA%)
96	1 (14%)	0 (0%)	1 (25%)	0 (NA%)
(Faltante)	23	9	2	12
<b>Tratamento com Sorafenibe, n (%)</b>	3 (10%)	1 (8.3%)	2 (33%)	0 (0%)
<b>Número de lesões, Mediana (Amplitude)</b>	2.00 (1.00 - 8.00)	3.00 (1.00 - 6.00)	2.00 (1.00 - 6.00)	2.00 (1.00 - 8.00)
<b>Tamanho da maior lesão, Média (Desvio Padrão)</b>	4.7 (3.8)	5.8 (4.8)	7.1 (3.1)	2.3 (0.9)
<b>Classificação macroscópica, n (%)</b>				
VN	13 (43%)	6 (50%)	0 (0%)	7 (58%)
EN	8 (27%)	5 (42%)	1 (17%)	2 (17%)
NP	2 (6.7%)	0 (0%)	2 (33%)	0 (0%)
MC	6 (20%)	1 (8.3%)	3 (50%)	2 (17%)
INF	1 (3.3%)	0 (0%)	0 (0%)	1 (8.3%)
exofítico	0 (0%)	0 (0%)	0 (0%)	0 (0%)
<b>Cirrose, n (%)</b>	23 (77%)	8 (67%)	3 (50%)	12 (100%)
<b>Padrão histológico predominante, n (%)</b>				
misto	12 (40%)	5 (42%)	2 (33%)	5 (42%)
trabecular	13 (43%)	7 (58%)	2 (33%)	4 (33%)

Relatório de Análise Estatística (SAR)

Características dos participantes	Total, N = 30	Censurado, N = 12	CHC, N = 6	Outra, N = 12
macrotrabecular	1 (3.3%)	0 (0%)	1 (17%)	0 (0%)
acinar	2 (6.7%)	0 (0%)	0 (0%)	2 (17%)
células claras	1 (3.3%)	0 (0%)	1 (17%)	0 (0%)
esteato-hepatítico	1 (3.3%)	0 (0%)	0 (0%)	1 (8.3%)
<b>Padrão histológico secundário, n (%)</b>				
não tem	12 (40%)	6 (50%)	2 (33%)	4 (33%)
misto	0 (0%)	0 (0%)	0 (0%)	0 (0%)
trabecular	3 (10%)	0 (0%)	1 (17%)	2 (17%)
macrotrabecular	1 (3.3%)	0 (0%)	1 (17%)	0 (0%)
células claras	11 (37%)	5 (42%)	1 (17%)	5 (42%)
cirroso	1 (3.3%)	0 (0%)	1 (17%)	0 (0%)
esteato-hepatítico	2 (6.7%)	1 (8.3%)	0 (0%)	1 (8.3%)
<b>Tipo histológico, n (%)</b>				
Convencional	28 (93%)	12 (100%)	5 (83%)	11 (92%)
Esteato-hepatítico	1 (3.3%)	0 (0%)	0 (0%)	1 (8.3%)
macrotrabecular maciço	1 (3.3%)	0 (0%)	1 (17%)	0 (0%)
<b>Grau de diferenciação segundo Edmonson predominante, n (%)</b>				
Indeterminado	0 (0%)	0 (0%)	0 (0%)	0 (0%)

# Consultoria em Estatística Médica e Epidemiologia Clínica

## Relatório de Análise Estatística (SAR)

Características dos participantes	Total, N = 30	Censurado, N = 12	CHC, N = 6	Outra, N = 12
I	1 (3.3%)	1 (8.3%)	0 (0%)	0 (0%)
II	22 (73%)	9 (75%)	5 (83%)	8 (67%)
III	7 (23%)	2 (17%)	1 (17%)	4 (33%)
IV	0 (0%)	0 (0%)	0 (0%)	0 (0%)
<b>Grau de diferenciação segundo Edmonson secundário, n (%)</b>				
não tem	13 (43%)	6 (50%)	3 (50%)	4 (33%)
I	5 (17%)	2 (17%)	1 (17%)	2 (17%)
II	3 (10%)	1 (8.3%)	0 (0%)	2 (17%)
III	4 (13%)	1 (8.3%)	0 (0%)	3 (25%)
IV	5 (17%)	2 (17%)	2 (33%)	1 (8.3%)
<b>Grau de diferenciação segundo OMS predominante, n (%)</b>				
pouco	0 (0%)	0 (0%)	0 (0%)	0 (0%)
moderadamente	28 (93%)	11 (92%)	6 (100%)	11 (92%)
bem	2 (6.7%)	1 (8.3%)	0 (0%)	1 (8.3%)
<b>Grau de diferenciação segundo OMS secundário, n (%)</b>				
não tem	15 (50%)	5 (42%)	3 (50%)	7 (58%)
pouco	6 (20%)	2 (17%)	2 (33%)	2 (17%)
moderadamente	1 (3.3%)	0 (0%)	0 (0%)	1 (8.3%)



# Consultoria em Estatística Médica e Epidemiologia Clínica

## Relatório de Análise Estatística (SAR)

Características dos participantes	Total, N = 30	Censurado, N = 12	CHC, N = 6	Outra, N = 12
bem	8 (27%)	5 (42%)	1 (17%)	2 (17%)
<b>Grau de diferenciação segundo OMS terciário, n (%)</b>				
não tem	28 (93%)	11 (92%)	6 (100%)	11 (92%)
pouco	1 (3.3%)	0 (0%)	0 (0%)	1 (8.3%)
moderadamente	0 (0%)	0 (0%)	0 (0%)	0 (0%)
bem	1 (3.3%)	1 (8.3%)	0 (0%)	0 (0%)
<b>Invasão vascular microscópica, n (%)</b>	19 (63%)	6 (50%)	5 (83%)	8 (67%)
<b>Invasão vascular macroscópica, n (%)</b>	3 (10%)	0 (0%)	1 (17%)	2 (17%)
<b>Margem cirúrgica, n (%)</b>				
comprometida	2 (6.7%)	0 (0%)	2 (33%)	0 (0%)
livre	9 (30%)	6 (50%)	3 (50%)	0 (0%)
tx	19 (63%)	6 (50%)	1 (17%)	12 (100%)
<b>Infiltrado inflamatório peri/intra tumoral, n (%)</b>	30 (100%)	12 (100%)	6 (100%)	12 (100%)
<b>Linfócitos, n (%)</b>	30 (100%)	12 (100%)	6 (100%)	12 (100%)
<b>Neutrófilos, n (%)</b>	19 (63%)	8 (67%)	2 (33%)	9 (75%)
<b>Esteatose associada ao tumor, n (%)</b>	15 (50%)	7 (58%)	0 (0%)	8 (67%)
<b>Necrose tumoral, n (%)</b>	11 (37%)	5 (42%)	4 (67%)	2 (17%)
<b>Necrose pós terapia locorregional, n (%)</b>	16 (53%)	7 (58%)	3 (50%)	6 (50%)

Relatório de Análise Estatística (SAR)

Características dos participantes	Total, N = 30	Censurado, N = 12	CHC, N = 6	Outra, N = 12
<b>Proliferação ductular, n (%)</b>	13 (43%)	5 (42%)	2 (33%)	6 (50%)
<b>Metástase, n (%)</b>	5 (17%)	1 (8.3%)	4 (67%)	0 (0%)
<b>Local da metástase, n (%)</b>				
atrio direito, peritoneo	1 (20%)	1 (100%)	0 (0%)	0 (NA%)
peritoneo	1 (20%)	0 (0%)	1 (25%)	0 (NA%)
pulmão	2 (40%)	0 (0%)	2 (50%)	0 (NA%)
tibia, pulmão, vertebral L1	1 (20%)	0 (0%)	1 (25%)	0 (NA%)
(Faltante)	25	11	2	12
<b>Quanto tempo depois apresentou metástase (anos), Mediana (Amplitude)</b>	3.00 (1.00 - 9.00)	5.00 (5.00 - 5.00)	2.50 (1.00 - 9.00)	NA (Inf - -Inf)
(Faltante)	25	11	2	12

## 4.2 Risco de óbito por CHC

As estimativas de associação entre os três marcadores e a incidência de óbito por CHC são apresentadas na tabela 2. Nenhum marcador é significativamente associado à mortalidade, mas sob a perspectiva descritiva, a estimativa bruta indica riscos aumentados para os três marcadores, com CK19 indica um aumento em duas vezes no risco de óbito (HR=2.16), EpCAM um risco 9% maior (HR=1.09) e CD44 20% maior (HR=2.16). Todos os IC têm grande amplitude, variando desde efeito extremamente protetor até risco aumentado em 13 vezes, indicando baixa precisão nas estimativas.

**Tabela 2** Risco de óbito por CHC.

Características	Estimativa bruta			Interações			Estimativa ajustada		
	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>12</sup>	95% IC <sup>1</sup>	Valor p
<b>CK19</b>	2.16	0.35 to 13.3	0.407	5.35	0.42 to 67.5	0.194	5.73	0.42 to 77.5	0.189

Relatório de Análise Estatística (SAR)

Características	Estimativa bruta			Interações			Estimativa ajustada		
	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>12</sup>	95% IC <sup>1</sup>	Valor p
EpCAM	1.09	0.09 to 13.9	0.947	2.63	0.19 to 36.9	0.472	4.55	0.24 to 86.6	0.313
CD44	1.20	0.13 to 11.0	0.871	0.69	0.05 to 9.06	0.780	0.56	0.04 to 7.09	0.653
CK19 * EpCAM				0.09	0.00 to 7.06	0.280	0.05	0.00 to 6.37	0.224

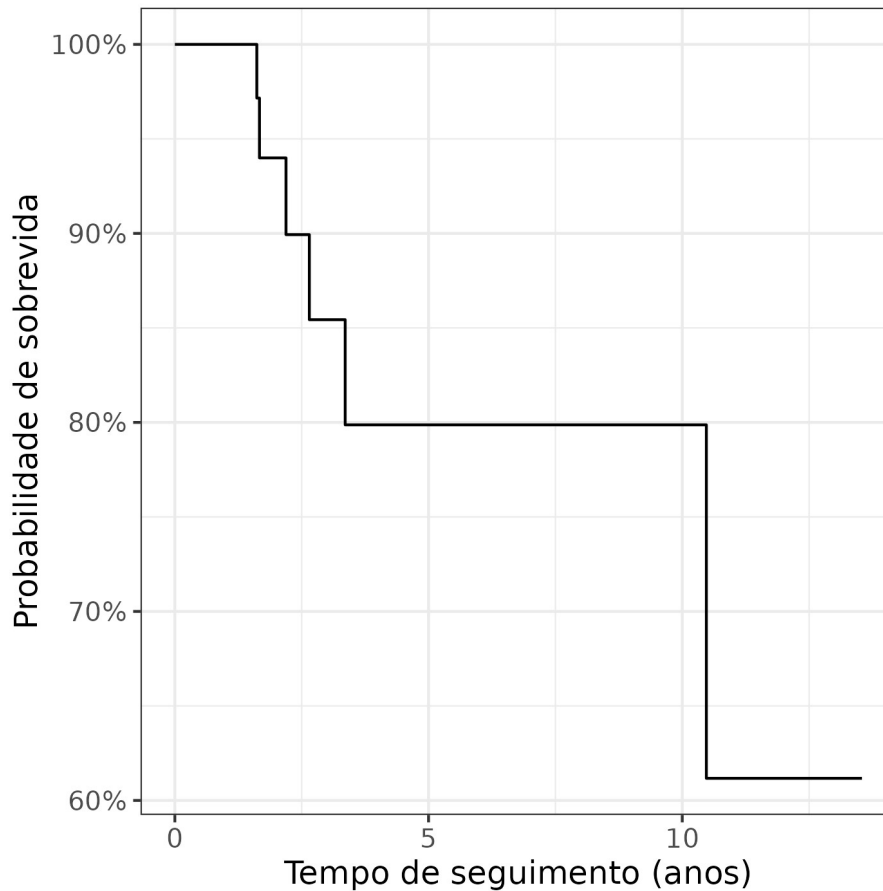
<sup>1</sup>HR = Taxa de risco, IC = Intervalo de confiança

<sup>2</sup>Ajustado para sexo e invasão vascular microscópica

Considerar as interações entre os marcadores resulta em alterações drásticas nas estimativas, o que indica que as interações devem ser consideradas em análises desses marcadores. Sob a perspectiva descritiva, CK19 indica um aumento em cinco vezes de mortalidade (HR=5.35) e EpCAM um risco duas vezes maior (HR=2.63), quando considerados isoladamente. Quando considerados juntos, estes marcadores formam um fator protetor (HR=0.09). O marcador CD44 teve uma inversão na direção do efeito em relação à estimativa bruta, agora indicando um fator protetor (HR=0.69). Os IC são ainda maiores que na estimativa bruta, atingindo riscos da ordem de 70 vezes. Nenhum fator foi significativamente associado com o desfecho.

Ajustando para o sexo e a invasão vascular microscópica as estimativas mantêm as direções dos efeitos do modelo anterior, e com magnitudes semelhantes, o que indica alguma robustez nas estimativas, considerando as interações entre os marcadores. CK19 indica um risco quase seis vezes maior de óbito (HR=5.73) e EpCAM um risco quatro vezes maior (HR=4.55), quando considerados isoladamente. Quando considerados juntos, estes marcadores formam um fator protetor (HR=0.05). O marcador CD44 indica um fator protetor (HR=0.56). Os IC são muito amplos e comparáveis aos IC respectivos da estimativa com interações. Nenhum fator foi significativamente associado com o desfecho.

A figura 2 mostra o prognóstico de um participante hipotético com todos os três marcadores, assumindo valores para as variáveis incluídas no modelo: sexo masculino seguindo o perfil epidemiológico da amostra, e com presença de invasão vascular microscópica. A curva de sobrevivência deste perfil não atinge 50% de mortalidade, portanto não é possível estimar o tempo mediano de óbito.



**Figura 2** Sobrevivência do perfil selecionado (Homem, positivo para os três marcadores, com invasão vascular micro).

## 5 OBSERVAÇÕES E LIMITAÇÕES

Nesta análise dezesseis modelos foram avaliados, incluindo especificações manuais e automatizadas (Apêndice), mas muitas foram rejeitadas por critérios técnicos (eg. convergência e premissa de proporcionalidade de riscos). As amplitudes dos IC dos modelos selecionados para a análise de sensibilidade tornam as estimativas de difícil interpretação e não-aplicáveis; não é recomendado o uso destas estimativas na prática clínica.

O fato de as estimativas do segundo e terceiro modelos serem comparáveis é indício de robustez, tanto nas direções dos efeitos como nas amplitudes da incerteza em torno das estimativas. Por outro lado, na presença de interações, a inversão do efeito do marcador CD44 em relação à estimativa bruta exige cautela na interpretação. Isto poderia indicar que há modificação de efeito, e o termo de interação deve ser considerado na análise,

mas também pode indicar que há confundimento residual mesmo após as tentativas de ajuste na análise. Não foi possível determinar qual destas hipóteses é mais verossímil nesta amostra.

É notável que o algoritmo stepwise, aplicado na direção “para trás” (backward stepwise selection) removeu todos os termos do modelo (seção 8.2.1), que certamente é concordante com a hipótese nula de não associação.

Uma hipótese para a dificuldade encontrada em ajustar os modelos aos dados pode ser formulada sobre o desenho do estudo. Não há clareza sobre qual é a população alvo e quanto a amostra de estudo é representativa desta população, bem como os critérios de seleção dos prontuários coletados. A análise exploratória (Figuras 1 na seção 4.1 e A3 na seção 8.1) indica que o tempo até os desfechos são bimodais; outras características parecem ser consistentes entre os participantes (Tabela 1). Ambas as situações dificultam determinar quais características discriminam o risco de ocorrência do desfecho. Adicionalmente, o número de eventos observado é baixo, relativo ao tamanho da amostra, o que pode ter dificultado a convergência do modelo de Cox. Uma análise de tempo-até-evento se beneficia de um número de eventos “razoável” e tempos “bem comportados”. A bimodalidade do tempo pode ser um efeito de amostra pequena, mas não é possível testar esta hipótese com a amostra disponível.

### Guideline recomendado de relato completo

A adoção dos guidelines de relato da rede EQUATOR (<http://www.equator-network.org/>) têm sido ampliada em revistas científicas. Recomenda-se que todos os estudos observacionais sigam o guideline STROBE (von Elm et al, 2014).

Em particular quando um estudo retrospectivo é conduzido usando registros hospitalares de rotina, incluindo informações de prontuários, é recomendado que a extensão RECORD do guideline STROBE seja considerado (Benchimol et al, 2015).

## 6 CONCLUSÕES

Os marcadores CK19, EpCAM e CD44 não são associados ao óbito por CHC nesta amostra. Nenhum fator de risco foi identificado como alterando esta associação.

## 7 REFERÊNCIAS

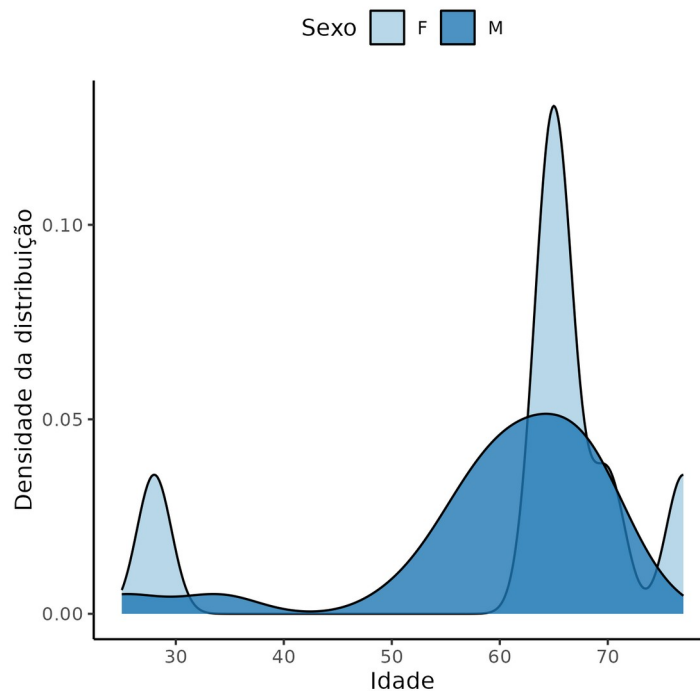
- **SAP-2023-028-AN-v01** – Plano Analítico para Associação entre os marcadores CK19, EpCAM e CD44 e óbito por carcinoma hepatocelular: coorte retrospectiva
- von Elm E, Altman DG, Egger M, Pocock SJ, Gøtzsche PC, Vandenbroucke JP; STROBE Initiative. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement: guidelines for reporting observational

studies. Int J Surg. 2014 Dec;12(12):1495-9  
(<https://doi.org/10.1016/j.ijssu.2014.07.013>).

- Benchimol EI, Smeeth L, Guttman A, Harron K, Moher D, Petersen I, Sørensen HT, von Elm E, Langan SM; RECORD Working Committee. The Reporting of studies Conducted using Observational Routinely-collected health Data (RECORD) statement. PLoS Med. 2015 Oct 6;12(10):e1001885  
(<https://doi.org/10.1371/journal.pmed.1001885>).

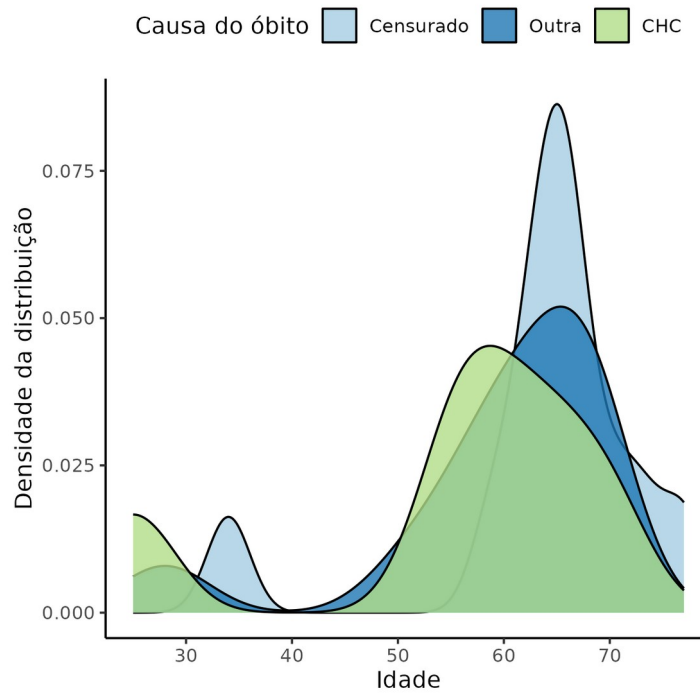
## 8 APÊNDICE

### 8.1 Análise exploratória de dados

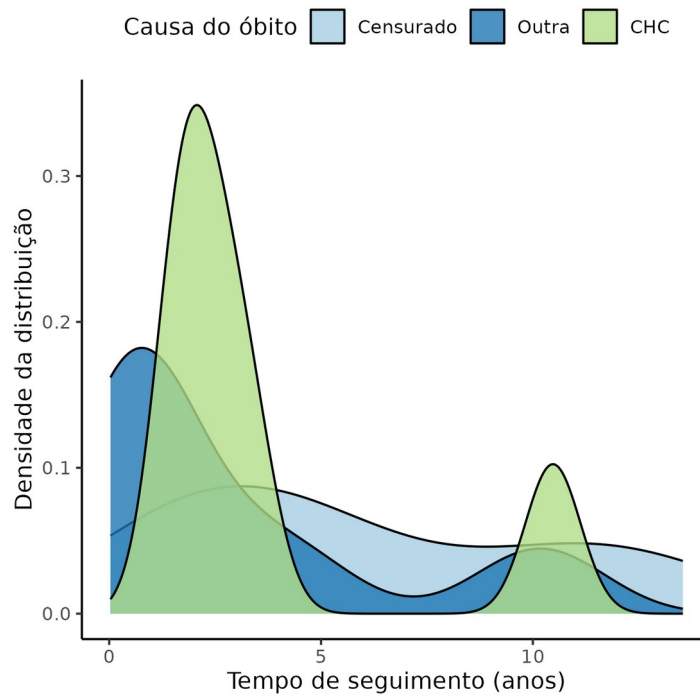


**Figura A1** Distribuição da idade na população do estudo, por sexo.

Relatório de Análise Estatística (SAR)



**Figura A2** Distribuição da idade na população do estudo, por desfecho.



**Figura A3** Distribuição do tempo até o evento, por desfecho.

## 8.2 Estratégia de modelagem

### 8.2.1 Tentativas de especificação de modelos

Foram ajustados dois modelos iniciais, um com os marcadores (modelo 1) e um que incluía as interações de segunda ordem entre os marcadores (modelo 2). O modelo com as interações não convergiu, e um novo modelo foi ajustado removendo os termos de interações que impediam o ajuste.

Um novo modelo com Sexo, Idade foi ajustado com sucesso e aceito para a próxima etapa (modelo 7).

A partir desse modelo, as variáveis Etiologia primária, Metástase, Recidiva, Cirrose, Grau de diferenciação segundo OMS predominante não resultaram em convergência e foram rejeitadas.

A variável Invasão vascular microscópica resultou em modelo convergente, mas tornou a idade não-proporcional, sendo assim rejeitado também.

A listagem a seguir mostra os 16 modelos que foram testados e seus resultados.

1. marcadores
  - modelo selecionado como Estimativa bruta
2. acrescentar todas as interações
  - modelo não convergente
3. remover as interações conflitantes
  - Modelo selecionado como Interações
4. backward stepwise com todas as variáveis
  - modelo nulo (sem variáveis)
5. forward stepwise com todas as variáveis
  - acrescentadas "mtx", "recidiva" e "idade" (contínua)
  - modelo não tem resíduos, portanto a proporcionalidade de riscos não pode ser testada
6. acrescentar sexo
  - modelo convergente e com riscos proporcionais
7. stepwise forward para acrescentar "idade" ou "idade cortada"
  - "idade cortada" acrescentada
  - modelo convergente e com riscos proporcionais
  - modelo com coeficientes não interpretáveis
8. stepwise forward para acrescentar maior\_tamanho ou maior\_tamanho\_cat (não convergente)
  - "tamanho contínuo" acrescentada
  - modelo não convergente
9. acrescentar etiologia\_principal ao modelo 7



- modelo não convergente
- 10. acrescentar recidiva ao modelo 7
  - modelo não convergente
- 11. acrescentar cirrose ao modelo 7
  - modelo não convergente
- 12. acrescentar dif\_oms ao modelo 7
  - modelo não convergente
- 13. acrescentar inv\_vascular\_micro ao modelo 7
  - modelo convergente mas idade viola premissa de riscos proporcionais
- 14. acrescentar inv\_vascular\_micro ao modelo 6 (sem idade)
  - modelo convergente e com riscos proporcionais
  - modelo selecionado como final
- 15. acrescentar metástase ao modelo 7
  - modelo não convergente
- 16. acrescentar terapia\_locorreg ao modelo 7
  - modelo não convergente

Tentativas de incluir interações com idade, sexo e ambas resultaram em modelos não-convergentes que não puderam ser testados quanto à proporcionalidade de riscos (dados não mostrados).

### 8.2.2 Proporcionalidade de riscos

A tabela A1 mostra o teste de Schoenfeld nos modelos convergentes. A idade e a taxa de risco basal (GLOBAL) tiveram risco significativamente não-proporcional no modelo 9.

**Tabela A1** Teste de Schoenfeld de proporcionalidade de riscos nos modelos convergentes.

modelo	ck19	ep_cam	cd44	GLOBAL	ck19:ep_ca m	sexo	idade_cat	inv_vascula r_micro
modelo 1	>0.9	0.8	0.15	0.4	NA	NA	NA	NA
modelo 3	0.9	0.6	0.13	0.5	0.3	NA	NA	NA
modelo 6	>0.9	0.5	0.2	0.6	0.3	0.3	NA	NA
modelo 7	>0.9	0.6	0.12	0.071	0.3	0.8	0.051	NA
modelo 13	>0.9	0.8	0.11	0.046	0.4	0.7	0.040	0.6
modelo 14	0.9	0.6	0.2	0.6	0.4	0.3	NA	0.4

### 8.2.3 Estimativas de modelos convergentes

A tabela A2 mostra os modelos convergentes. Os modelos 1, 3 e 14 atendem a todos os critérios foram os modelos selecionados para a análise dos resultados (Tabela 2).

Relatório de Análise Estatística (SAR)

**Tabela A2** Estimativas dos modelos convergentes.

Características	modelo 1 (est. bruta)			modelo 3 (interações)			modelo 6			modelo 7			modelo 13			modelo 14 (est. ajustada)			
	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	HR <sup>1</sup>	95% IC <sup>1</sup>	Valor p	
<b>CK19</b>	2.16	0.35 to 13.3	0.407	5.35	0.42 to 67.5	0.194	5.96	0.46 to 78.0	0.174	75.4	1.02 to 5,569	<b>0.049</b>	76.3	0.97 to 5,995	0.052	5.73	0.42 to 77.5	0.189	
<b>EpCAM</b>	1.09	0.09 to 13.9	0.947	2.63	0.19 to 36.9	0.472	5.36	0.33 to 88.0	0.239	21.5	0.50 to 928	0.110	17.6	0.35 to 884	0.151	4.55	0.24 to 86.6	0.313	
<b>CD44</b>	1.20	0.13 to 11.0	0.871	0.69	0.05 to 9.06	0.780	0.67	0.06 to 7.75	0.750	0.09	0.00 to 2.88	0.174	0.07	0.00 to 2.64	0.153	0.56	0.04 to 7.09	0.653	
<b>CK19 * EpCAM</b>				0.09	0.00 to 7.06	0.280	0.05	0.00 to 5.20	0.206	0.00	0.00 to 1.76	0.069	0.00	0.00 to 2.29	0.080	0.05	0.00 to 6.37	0.224	
<b>Sexo</b>																			
F							—	—		—	—		—	—		—	—		
M							4.41	0.43 to 45.5	0.213	12.1	0.44 to 331	0.141	8.09	0.24 to 278	0.246	3.69	0.28 to 49.0	0.322	
<b>Idade (Maior ou igual a 60 anos)</b>										0.05	0.00 to 0.85	<b>0.038</b>	0.05	0.00 to 0.82	<b>0.036</b>				
<b>Invasão vascular microscópica</b>													2.36	0.15 to 37.3	0.542	1.99	0.19 to 20.5	0.562	

<sup>1</sup>HR = Taxa de risco, IC = Intervalo de confiança

### 8.3 Análises associadas

Esta análise é parte de um projeto maior e é suportada por outras análises, disponíveis abaixo.

#### Associação entre fatores diagnósticos e marcadores de CHC: estudo caso-controle

<https://philsf-biostat.github.io/SAR-2023-031-AN/>

### 8.4 Disponibilidade

Todos os documentos gerados nessa consultoria foram incluídos no portfólio do consultor.

O portfólio pode ser visto em:

<https://philsf-biostat.github.io/SAR-2023-028-AN/>

## 8.5 Dados utilizados

A tabela A3 mostra a estrutura da tabela de dados analíticos.

**Tabela A3** Estrutura da tabela de dados analíticos (continued below)

id	ck19	ep_cam	cd44	outcome	time	sexo	idade	idade_cat	tx	etiologia_p rincipal	etiologia_2aria	terapia_lo correçg	metodo	momento	recidiva	epoca_reci diva_eme ses	sorafenibe	numero_de lesoes	maior_tama nho	maior_tama nho_cat	classif_ma cro	cirrose	padrao_his tológico_p redom inante	padrao_his tológico_s ecundário
1																								
2																								
3																								
-																								
N																								

tipo_histologico	dif_edmonson_pred ominante	dif_edmonson_secun daria	dif_ons	dif_ons_secun daria	dif_ons_tec naria	inv_vascular_micr o	inv_vascular_ma cro	nargen	inf_inflam atorio_peri intra_tu moral	linfocitos	neutrofilos	esteatose associada	necrose_tu m	necrose_qt	prolif_ductu lar	mtx	local_mtx	epoca_mtx anos

Os dados utilizados neste relatório não podem ser publicados online por questões de sigilo.