

---

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

DOCUMENTO: SAP-2023-031-AN-v02

De: Felipe Figueiredo Para: Aline Nascimento

2023-12-01

## SUMÁRIO

1	LISTA DE ABREVIATURAS.....	2
2	CONTEXTO.....	2
2.1	Objetivos.....	2
2.2	Hipóteses.....	2
3	DADOS.....	3
3.1	Dados brutos.....	3
3.2	Tabela de dados analíticos.....	3
4	PARÂMETROS DO ESTUDO.....	4
4.1	Desenho do estudo.....	4
4.2	Critérios de inclusão e exclusão.....	4
4.3	Exposições.....	4
4.4	Desfechos.....	4
4.5	Covariáveis.....	4
5	MÉTODOS ESTATÍSTICOS.....	5
5.1	Análises estatísticas.....	5
5.1.1	Análise descritiva.....	5
5.1.2	Análise inferencial.....	5
5.1.3	Modelagem estatística.....	5
5.1.4	Dados faltantes.....	5
5.2	Significância e Intervalos de Confiança.....	5
5.3	Tamanho da amostra e Poder.....	5
5.4	Softwares utilizados.....	5
6	OBSERVAÇÕES E LIMITAÇÕES.....	6
7	REFERÊNCIAS.....	6
8	APÊNDICE.....	7
8.1	Análises associadas.....	7
8.2	Disponibilidade.....	7

---

---

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

## Histórico do documento

Versão	Alterações
01	Versão inicial
02	Novas variáveis incluídas na seção Exposições

## 1 LISTA DE ABREVIATURAS

- CHC: carcinoma hepatocelular
- IC: intervalo de confiança
- OR: razão de chances (*odds ratio*)

## 2 CONTEXTO

No diagnóstico e manejo de pacientes com carcinoma hepatocelular há o desafio de se determinar que exames auxiliam na confirmação do câncer. Esta análise assume uma relação entre três marcadores associados ao carcinoma hepatocelular e se propõe a identificar quais dentre estes ocorrem concomitantemente com fatores, para determinar uma bateria de testes diagnósticos adequada.

### 2.1 Objetivos

Avaliar associação entre diversos fatores diagnósticos e marcadores de carcinoma hepatocelular CK19, EpCAM e CD44.

### 2.2 Hipóteses

Os fatores diagnósticos (seção 4.3) estão associados aos marcadores de carcinoma hepatocelular.

### 3 DADOS

#### 3.1 Dados brutos

A base de dados original possuía 50 variáveis coletadas de 31 observações.

#### 3.2 Tabela de dados analíticos

Variáveis dicotômicas foram codificadas como indicadores com valores 0 e 1. Variáveis categóricas foram codificadas considerando todos os níveis observados. A unidade de análise é o participante e indivíduos em duplicata serão removidas para ter uma única observação por unidade.

Depois dos procedimentos de limpeza e seleção 15 variáveis foram incluídas na análise com 30 observações.

**Tabela 1** Estrutura da tabela de dados analíticos após seleção e limpeza das variáveis.

id	ck19	ep_can	cd44	etiologia_principal	dif_ons	recidiva	cirrose	inv_vascular_micro	terapia_locorreg	mtx	prolif_ductular	neutrofilos	idade_catt	maior_tamanho_catt
1														
2														
3														
...														
N														

Todas as variáveis da tabela de dados analíticos foram identificadas de acordo com as descrições das variáveis, e os valores foram identificados de acordo com o dicionário de dados providenciado. Estas identificações possibilitarão a criação de tabelas de resultados com qualidade de produção final.

## 4 PARÂMETROS DO ESTUDO

### 4.1 Desenho do estudo

Estudo caso-controle, aninhado à coorte retrospectiva de mortalidade por CHC. A mortalidade na coorte foi estimada em **SAR-2023-028-AN**.

### 4.2 Critérios de inclusão e exclusão

N/A

### 4.3 Exposições

1. Etiologia primária
2. Grau de diferenciação segundo OMS predominante
3. Recidiva
4. Cirrose
5. Invasão vascular microscópica
6. Terapia locorregional
7. Metástase
8. Proliferação ductular
9. Neutrófilos
10. Idade (Maior ou igual a 60 anos)
11. Tamanho da maior lesão (maior ou igual a 5cm)

### 4.4 Desfechos

**Especificação das medidas de desfecho** (Zarin, 2011):

1. (Domínio) Câncer
2. (Mensuração específica) Ocorrência dos marcadores CK19, EpCAM e CD44
3. (Métrica específica) Prevalência de participantes com cada marcador
4. (Método de agregação) Chance

#### Desfechos primários

- Chance do participante ser CK19 positivo
- Chance do participante ser EpCAM positivo
- Chance do participante ser CD44 positivo

### 4.5 Covariáveis

N/A

---

## 5 MÉTODOS ESTATÍSTICOS

### 5.1 Análises estatísticas

#### 5.1.1 Análise descritiva

A descrição das características epidemiológicas e clínicas dos participantes da coorte estão apresentadas na análise associada **SAR-2023-028-AN**. Para esta análise pretende-se descrever as ocorrências de interação entre os três marcadores. Como uma tabela de contingência de três variáveis exige uma estrutura tridimensional, optou-se por descrever recortes da tabela de contingência com secções nos valores de CD44. As frequências de ocorrência concomitante de múltiplos marcadores podem ser obtida a partir das secções.

#### 5.1.2 Análise inferencial

Todas as comparações entre os grupos serão avaliadas como análises univariadas. Diferenças nas distribuições de variáveis categóricas serão avaliadas com a OR e seu IC.

#### 5.1.3 Modelagem estatística

N/A

#### 5.1.4 Dados faltantes

Não serão realizadas imputações de dados faltantes. Todas as avaliações serão realizadas como análises de casos completos. As frequências e proporções de dados faltantes serão reportadas em tabelas.

### 5.2 Significância e Intervalos de Confiança

Todas as análises serão realizadas ao nível de significância de 5%. Todos os testes de hipóteses e intervalos de confiança calculados serão bicaudais.

### 5.3 Tamanho da amostra e Poder

N/A

### 5.4 Softwares utilizados

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

## 6 OBSERVAÇÕES E LIMITAÇÕES

### 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).

## 7 REFERÊNCIAS

- **SAR-2023-031-AN-v02** – Associação entre fatores diagnósticos e marcadores de CHC: estudo caso-controle
- **SAR-2023-028-AN** – Associação entre os marcadores CK19, EpCAM e CD44 e óbito por carcinoma hepatocelular: coorte retrospectiva
- Zarin DA, et al. The ClinicalTrials.gov results database – update and key issues. N Engl J Med 2011;364:852-60 (<https://doi.org/10.1056/NEJMsa1012065>).
- Gamble C, et al. Guidelines for the Content of Statistical Analysis Plans in Clinical Trials. JAMA. 2017;318(23):2337–2343 (<https://doi.org/10.1001/jama.2017.18556>).
- 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.ijsu.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

Este documento foi elaborado seguindo recomendações de estrutura para Planos de Análise Estatística (Gamble, 2017) para maior transparência e clareza.

### 8.1 Análises associadas

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

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

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

### 8.2 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-2023-031-AN/>