

# Desenhos Experimentais e Observacionais

## Desenhos Básicos, Viés e Confundimento, Exemplos

José Luiz Padilha – DEST/UFPR

26 de março, 2025

# Conteúdo

- 1 Introdução
- 2 Desenhos Básicos de Estudos Observacionais
- 3 Viés e Confundimento
- 4 Estratificação e Pareamento
- 5 Bibliografia

# Introdução

# Introdução aos Estudos Biomédicos

Busca-se determinar se, e em que grau, existe uma relação de *causa-efeito* entre as características ou tratamento e o evento de interesse.

Uma característica marcante é a presença da *variabilidade biológica*.

Tais estudos são muito diferentes de estudos determinísticos, nos quais há variabilidade insignificante ou nula (estudos em física, engenharia, etc.).

Alguns pilares metodológicos comuns são apresentados na sequência.

# Características Fundamentais

- Os dados, i.e. as medidas realizadas nos indivíduos, são interpretadas no contexto de um modelo estatístico-probabilístico.
- Os estudos devem ser bem planejados e detalhados antes do seu início e ser realizados com estrita aderência. Ocorrendo mudanças inevitáveis no *protocolo do estudo* é extremamente importante avaliar o impacto no modelo probabilístico usado para o estudo.
- Os dados devem ser analisados por métodos estatísticos apropriados, definidos no planejamento do estudo.
- Se tudo ocorrer como descrito, é possível estender os resultados obtidos na amostra para toda a população que tenha características similares à amostra estudada.

- É um documento que descreve todos os aspectos científicos de um estudo;
- Um manual de instruções, porque descreve em detalhes como o estudo deve ser realizado;
- Um documento legal, porque obriga os investigadores a seguirem os procedimentos descritos, oficialmente aprovados pelos autores e ou autoridades competentes (comitê de ética).

**É importante a colaboração entre o pesquisador e o estatístico nessa fase.**

- Pode ser trágica a participação do estatístico apenas na seção da análise: há um forte elo entre os objetivos, delineamento, *end-points*, modelo probabilístico, análise e interpretação dos resultados.
- Da mesma forma pode ser perigosa a ausência do pesquisador na análise estatística: muitos aspectos metodológicos da análise podem se beneficiar do conhecimento do especialista.

*“To call in the statistician after the experiment is done may be no more than asking him to perform a postmortem examination: he may be able to say what the experiment died of.”*

*Sir Ronald Aylmer Fisher. Indian Statistical Congress, Sankhya, 1938*



# Tipos de Estudos

Podemos dividir os estudos em dois grandes tipos: observacionais e experimentais.

- Estudos observacionais: o pesquisador simplesmente observa o fenômeno em estudo. Ex: caso-controle, coorte, transversal.
- Estudos experimentais: o pesquisador controla diretamente os fatores em estudo. Ex: ensaio clínico aleatorizado.

As medidas de efeito resultantes e a força das conclusões podem diferir bastante entre os diferentes tipos de estudo.

O objetivo é estudar a relação entre uma característica e um evento sem manipular as condições sob as quais o estudo é realizado.

O pesquisador é limitado a selecionar a amostra e “observar”.

Na presença de forte associação entre as características e o evento pode-se suspeitar de uma relação de *causa-efeito*.

A característica capaz de influenciar o surgimento ou progressão de uma doença de forma predizível e reprodutível é chamada de *fator prognóstico*.

Um fator prognóstico pode ter influência positiva ou negativa:

- *Negativa*  $\Rightarrow$  **fator de risco**. Ex.: fumar é fator de risco para câncer de pulmão. A probabilidade de desenvolver tal doença aumenta com o número de cigarros fumados por dia.
- *Positiva*  $\Rightarrow$  **fator protetor**. Ex.: nível de atividade física no histórico natural de hipertensão.

Num *estudo experimental* o pesquisador controla as condições sob as quais o estudo é realizado.

Permite um grau maior de certeza na avaliação da relação de causa-efeito entre uma característica e um evento.

A característica sob estudo é de fato “atribuída” aos indivíduos. A gama de objetivos possíveis de estudos experimentais em humanos é mais restrita que nos estudos observacionais.

A atribuição dos tratamentos pode ser feita por grupos de estudo ou individualmente, indivíduo por indivíduo.

Cada indivíduo recebe um único tratamento ou uma sequência de tratamentos (único tratamento pode ser uma combinação de drogas e/ou outras intervenções terapêuticas).

O pesquisador atribui indivíduos aos tratamentos por meio de um método chamado *aleatorização*.

Aleatorização está relacionada com mascaramento (*blinding*) que consiste em tornar os tratamentos do estudo indistinguíveis entre si tanto para o pesquisador quanto para os pacientes (estudos *duplo cego*) ou apenas aos pacientes (estudo *cego*).

A aleatorização não apenas permite *balanceamento na distribuição dos pacientes entre os tratamentos* do estudo, mas também assegura que os fatores sub-experimentais não explicitamente controlados pelo protocolo, ou mesmo desconhecidos, sejam distribuídos de forma homogênea entre os grupos sob comparação.

# Desenhos Básicos de Estudos Observacionais

Baseado no método amostral aplicado, estudos observacionais podem ser de três tipos:

- **Prospectivo:** seleção com base na presença ou ausência de uma característica (ex: fumante ou não fumante).
- **Retrospectivo:** seleção com base na presença ou ausência de um evento (ex: câncer ou não câncer).
- **Transversal:** seleção com base na presença ou ausência do evento ou características é feita simultaneamente.



**Concorrentes**  $\Rightarrow$  expostos e não expostos são selecionados no presente, i.e. início do estudo e são acompanhados no futuro até o fim do estudo.  
 $\rightarrow$  *coorte contemporânea*

O desfecho ainda não ocorreu.

**Não concorrentes**  $\Rightarrow$  o pesquisador *volta no tempo*, seleciona os expostos e não expostos e tenta encontrar toda a informação relevante desses indivíduos, até o tempo presente.  $\rightarrow$  *coorte histórica*

A exposição e o desfecho já ocorreram antes do início do estudo.

# Estudo Prospectivo Concorrente: Exemplo

Considere que o objetivo do pesquisador seja avaliar a associação entre o peso ao nascer e a mortalidade.

Para isso ele divide as crianças em dois grupos: um grupo com peso até 2.5kg e um grupo com peso igual ou superior a 2.5kg.

Os dois grupos são avaliados por um determinado período de tempo e verifica-se o desfecho (mortalidade).

# Estudo Prospectivo Concorrente

Após as coortes serem selecionadas, são seguidas por tempo suficientemente grande, durante o qual todos os eventos de interesse são registrados.

Finalmente, as incidências dos eventos de interesse nas duas coortes são comparadas.

Podemos calcular razões de risco.

⇒ É similar ao ensaio clínico experimental, mas a característica não é atribuída por aleatorização. A vantagem é que permite que os efeitos de fatores de risco *que não podem ser “atribuídos” aos indivíduos* sejam estudados.

# Estudo Prospectivo Não-concorrente: Exemplo

Considere agora que o objetivo do pesquisador seja investigar a associação entre o uso de cocaína pela gestante e prematuridade do bebê.

Para isso ele busca no passado informações sobre a exposição (uso de cocaína pela mãe) e avalia no presente a ocorrência do evento (prematuridade).

As características metodológicas são idênticas. Mas, na prática, há duas grandes diferenças.

- Estudos não concorrentes têm tempos de execução muito menores e custos muito mais baixos.
- Em estudos não concorrentes, pode ser mais difícil obter informações, especialmente sobre exposição, mas também a relacionada à ocorrência do evento.

Imagine que o interesse do pesquisador seja relacionar a obesidade com ocorrência de angina de esforço.

O pesquisador então seleciona dois grupos de indivíduos, um obeso e outro não obeso, e avalia-se no mesmo dia a frequência de angina de esforço submetendo todos os indivíduos a um teste de esforço (ex, esteira ou bicicleta).

A grande desvantagem é que o pesquisador não sabe ao certo se a característica foi adquirida antes ou após o evento.

São também chamados *estudos de prevalência* porque permitem o cálculo de taxas de prevalência do evento entre os expostos e os não expostos.

Não são adequados para calcular incidência, pois incidência se refere a intervalo de tempo, enquanto prevalência se refere a um único ponto no tempo.

# Estudo Retrospectivo

O pesquisador seleciona dois grupos de indivíduos, *um com e outro sem* o evento de interesse no tempo da seleção (“hoje”) e pesquisa no passado de cada indivíduo a informação quanto à exposição da característica sob estudo.

*Casos*  $\Rightarrow$  indivíduos afetados pelo evento;

*Controles*  $\Rightarrow$  indivíduos não afetados pelo evento (“não casos”).

Chamados de estudos *estudos caso controle*.



# Estudo Caso Controle: Exemplo

Um exemplo clássico de um estudo caso controle relaciona a ocorrência de câncer de pulmão com a exposição ao tabaco.

Os casos são os indivíduos com câncer, enquanto os controles são os indivíduos sem câncer de pulmão.

O pesquisador então busca no passado a informação entre os expostos e não expostos ao tabaco.

Em tais estudos, o risco relativo de expostos e não expostos de experimentar os eventos não pode ser estimado.

Mas uma medida de associação entre exposição e evento pode ser obtida calculando a *odds-ratio* (OR).

A OR estima bem o RR apenas se o evento/doença de interesse é raro!!!

## Exemplo: Kruttschnitt et al. (1986)

Um exemplo da Criminologia é o estudo sobre determinantes de comportamento violento<sup>1</sup>, realizada em Mineápolis, EUA.

Foram selecionados 100 homens presos que haviam sido condenados por delito que envolvesse o uso ou a ameaça do uso da força contra a vítima. Estes indivíduos foram comparados com um grupo controle de homens não violentos e que não haviam sido presos.

Se a amostra tivesse sido selecionada aleatoriamente, seria necessário entrevistar um número muito grande de homens, a maioria dos quais sem prisão por delito violento, para identificar 100 casos.

---

<sup>1</sup>Exposição à violência familiar, exposição à violência na TV, desempenho escolar, e outras experiências na adolescência.

## Exemplo: Kxuttschnitt et al. (1986)

Forma de seleção da amostra:

- Casos: homens entre 18 e 25 anos de idade do Condado de Hennepin que tinham sido condenados por um crime violento e estavam presos no momento do estudo em três estabelecimentos prisionais.
- Controles: grupo de voluntários (pareados por gênero, idade e setor censitário de residência dos 10 aos 14 anos) que não foram presos ou condenados por crime violento.

## Principais vantagens de um estudo de coorte:

- Os estudos de coorte são menos propensos aos vícios que podem ocorrer nos estudos caso controle;
- São os mais indicados para o estudo de riscos, quando a experimentação não é possível.

## Principais desvantagens de um estudo de coorte:

- É um estudo mais demorado e que envolve custos elevados pelos recursos necessários para acompanhar muitos indivíduos ao longo do tempo estabelecido;
- Não disponibiliza resultados a curto prazo;
- Não é viável para doenças raras.

## Principais vantagens de um estudo de caso controle:

- Custo;
- Tempo envolvido para obtenção da resposta.

Estes são relativamente pequenos quando comparados aos de outros estudos, como, por exemplo, o estudo de coorte.

## Principais desvantagens de um estudo de caso controle:

- Resultados propensos a vícios;
- Exposição ao fator de interesse ser medida utilizando-se de informações passadas.



# Viés e Confundimento

Estudos observacionais:

- Diferente dos experimentais, não têm a vantagem da aleatorização dos indivíduos quanto à exposição ao fator de risco de interesse;
- Nem a vantagem do controle sobre as condições experimentais.

É menos fácil obter grupos homogêneos em estudos observacionais que nos experimentais. Ao fim do estudo, qualquer diferença entre os grupos comparados pode ser devido ao acaso, mas também ao viés.

Diferenças entre os grupos levam a um fenômeno conhecido como *confundimento*.

# Na presença de associação...

Quando os resultados apontam associação entre exposição à característica e uma doença podemos nos perguntar:

- ❶ O resultado pode ser devido ao acaso?
- ❷ Pode ser devido ao viés?
- ❸ Pode ser devido ao confundimento?

# Questão 1 – Acaso e Análise Estatística

A análise estatística permite que a primeira questão seja respondida.

O objetivo de um teste estatístico é *quantificar a probabilidade de que um resultado seja devido ao acaso*.

Se essa probabilidade é pequena, se aceita que há uma diferença real entre os grupos.

## Questão 2 – Viés ou Distorção

É um erro sistemático que causa uma estimativa incorreta da associação entre exposição e evento.

Causado pelo pesquisador e/ou pelos indivíduos sob estudo e pode ocorrer em qualquer fase, desde a seleção (viés de seleção) até avaliação (viés de observação, determinado pela lembrança falha dos indivíduos sobre os eventos passados).

Viés não pode ser avaliado por meio da análise estatística.

## Questão 2 – Viés de Seleção

Um *viés de seleção* ocorre quando o erro sistemático é introduzido durante a seleção dos indivíduos.

- ❶ Mecanismo usado para seleção de indivíduos é diferente de um grupo para outro;
- ❷ Mecanismo pelo qual um indivíduo é atribuído a um grupo ou outro.

Um exemplo do primeiro tipo é quando os casos são selecionados de um hospital em uma região e os controles são selecionados de um hospital em outra região, as duas regiões tendo diferenças sócio-econômicas e culturais.

## Questão 2 – Viés de Seleção

O segundo tipo, também chamado viés de classificação, pode ocorrer quando num estudo *caso controle* o conhecimento da exposição influencia a atribuição de certos casos aos grupos de doentes ou não doentes.

Um *estudo de coorte* é afetado pelo viés de classificação quando a presença ou ausência do evento influencia a atribuição dos indivíduos aos grupos exposto ou não exposto.

## Questão 2 – Viés de Observação

- *Viés de memória*: a habilidade de lembrar é diferente entre indivíduos com e sem a doença, que pode influenciar a avaliação da exposição.
- *Viés do entrevistador*: questões perguntadas de modo diferente, dependendo do grupo ao qual o indivíduo pertence, possivelmente influenciando as respostas.
- *Perda de acompanhamento*: os indivíduos que se perdem do estudo diferem daqueles que permanecem.
- *Viés de resposta*: os indivíduos que respondem às questões diferem sistematicamente daqueles que não respondem.



## Questão 2 – Viés ou Distorção

As formas mais perigosas de viés são:

### **Estudos caso controle:**

- Viés de seleção;
- Viés de memória.

### **Estudos coorte:**

- Viés causado por perda de acompanhamento;
- Viés de seleção.

## Questão 3 – O Fenômeno do Confundimento

A raiz comum de todas as manifestações de confundimento é que *os efeitos de alguns fatores estão misturados*.

Confundimento se refere à contaminação do efeito de interesse, i.e. o efeito da característica/exposição sobre o evento, por fatores externos a ele.

Confundimento é considerado um problema que leva a conclusões erradas (estimativas viesadas ou mesmo com direção invertida).

Confundimento pode ser, pelo menos em parte, controlado pela análise estatística, o viés não.

## Questão 3 – O Fenômeno do Confundimento

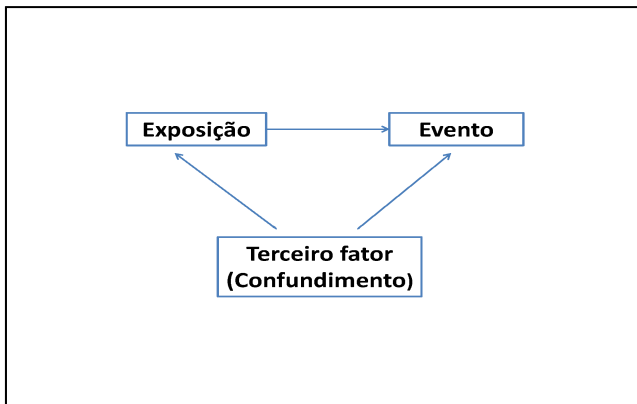


Figura 1: Fator potencialmente confundidor

# O Fenômeno do Confundimento: Exemplo 1

Considere o estudo da associação entre aumento na atividade física e redução de infarto do miocárdio.

Um fator potencialmente confundidor é a idade.

Pessoas jovens tendem a se exercitar mais e, independente do exercício, tendem a ter menor risco de infarto do miocárdio.

## O Fenômeno do Confundimento: Exemplo 2

Imagine que estejamos estudando a associação entre fumo e câncer de pâncreas.

Um fator potencialmente confundidor é álcool.

A propensão de fumar é maior entre as pessoas que bebem que entre as que não bebem; além disso o álcool, independente do fumo, é um fator de risco para câncer de pâncreas.

O viés pode ser evitado ou reduzido apenas através do delineamento cuidadoso e implementação correta do estudo.

O melhor método de controlar o confundimento é por meio da aleatorização à exposição, que não pode ser feita nos estudos observacionais. Então o confundimento é controlado pelas seguintes técnicas:

- Restrição (usada apenas na fase de planejamento).
- Estratificação e pareamento (usada tanto nas fases de planejamento e análise).
- Ajuste (usado apenas na fase de análise).

Os critérios de inclusão/exclusão no estudo são restritos de tal forma que apenas categorias específicas do fator potencialmente confundidor sejam incluídas.

Em alguns casos, o risco de confundimento pode ser completamente eliminado: ex, se o fator potencialmente confundidor é sexo pode-se admitir apenas homens ou apenas mulheres.

Em outros casos, pode ser minimizado. Ex, se o fator potencialmente confundidor é idade, pode-se admitir apenas indivíduos com uma faixa etária restrita.

Embora fácil, apresenta inconvenientes:

- ❶ Nem sempre é aplicável, pois frequentemente o fator potencial de confundimento é desconhecido.
- ❷ Reduz o número de indivíduos que podem participar do estudo.
- ❸ Reduz a habilidade de generalizar os resultados.



# Estratificação e Pareamento

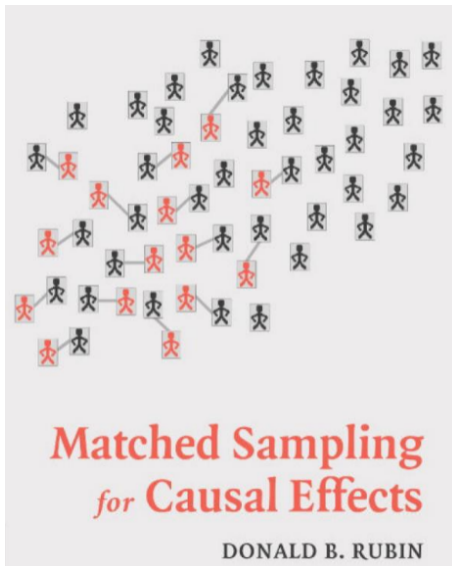
Técnicas usadas para reduzir o desbalanceamento em relação a potenciais fatores de confundimento.

**Estratificação:** divide-se casos e controles ou expostos e não expostos em subgrupos (estratos), baseados no nível do fator de interesse ou, no caso de múltiplos fatores, em combinações dos níveis.

Exemplo: Estratificação por sexo e presença de diabetes. Temos 4 subgrupos.

**Pareamento:** relaciona-se cada caso com um controle similar (ou múltiplos controles) com relação a um ou mais fatores de interesse.

Exemplo: mesmo sexo, raça, idade, nível sócio-econômico, etc.



# Exemplo 1

Para avaliar o efeito de um brinde nas vendas de determinado produto, planeja-se comparar as vendas em lojas que o vendem com brinde, com as vendas em lojas que não oferecem o brinde.

Para reduzir o efeito de variações devidas a outros fatores, as lojas foram agrupadas em pares de lojas, sendo que as lojas de um mesmo par sejam tão similares quanto possível, em termos de volume de vendas, localidade, identidade de preços, etc.

Em cada par de lojas, uma passou a oferecer o brinde, e a outra não.

# Exemplo 1

Par de lojas	Vendas	
	Sem brinde	Com brinde
1	33	43
2	43	39
3	26	33
4	19	32
5	37	43
6	27	43

- Desenhos alternativos?
- Correlação induzida pelo pareamento.
- Redução da variância...

Limitações do pareamento:

- Difícil de aplicar e requer tempo. Ex: com dois grupos, se dividimos idade em 5 categorias, sexo em 2 e raça em 3 há 30 ( $= 5 \times 2 \times 3$ ) combinações possíveis.
- Não é possível avaliar o efeito do fator potencialmente confundidor no resultado observado porque a distribuição deste fator foi forçada por delineamento a ser similar entre casos e controles.

Para serem eficientes, as duas técnicas devem ser consideradas tanto na fase de planejamento quanto de análise.

Quando o pareamento foi usado no delineamento, deve ser levado explicitamente em conta na análise estatística, isto é, similaridades forçadas entre indivíduos no mesmo bloco devem ser consideradas.

Isso é obtido por meio de métodos estatísticos específicos. Se não for considerado pareamento na análise, a verdadeira associação entre exposição e evento/doença é distorcida.

A estratificação *prospectiva* (definida no planejamento) é preferível à estratificação *a posteriori* pois:

- A segunda pode ser influenciada pelo conhecimento dos resultados.
- Com estratificação *a posteriori* pode-se deparar com problemas como dados ausentes ou inadequados em relação aos fatores potencialmente confundidores, ou estratos muito desbalanceados.

Essas questões, por delineamento, são evitadas pela estratificação prospectiva.



Obtém-se posteriormente estimativas gerais da associação (ajustadas) com base nas estimativas específicas.

Existem diferentes métodos para combinar as estimativas por estrato numa única estimativa específica (ex: Mantel-Haenszel).

São baseados no cálculo de médias ponderadas das estimativas específicas ao estrato.

Esses métodos variam dependendo do tipo de ponderação usado.

## Exemplo 2: Pequeno Exemplo Numérico

Imagine que desejamos avaliar o efeito de um tratamento binário ( $Tr$ ) em uma resposta contínua ( $Y$ ). Dados de 20 indivíduos estão disponíveis<sup>2</sup>.

Suponha que o efeito do tratamento possa ser confundido por três potenciais variáveis ( $Z1$ ,  $Z2$ , e  $Z3$ ) que também foram medidas.

```
library(tidyverse); load("dat.RData"); str(dat)
```

```
'data.frame': 20 obs. of 6 variables:
 $ id: int  1 2 3 4 5 6 7 8 9 10 ...
 $ Tr: int  1 1 1 1 0 0 0 0 0 0 ...
 $ Z1: int  1 1 1 0 1 1 1 0 0 0 ...
 $ Z2: num  0.25 1.95 2.24 3.54 0.26 1.95 2.26 3.5 3.56 2.58 ...
 $ Z3: num  1 2.03 1.44 0.67 0.99 2 1.47 0.7 0.77 -0.11 ...
 $ Y : num  3.5 5.2 6.59 7.56 1.55 ...
```

```
dat$Tr <- as.factor(dat$Tr)
```

---

<sup>2</sup>Os dados são hipotéticos para fins ilustrativos.

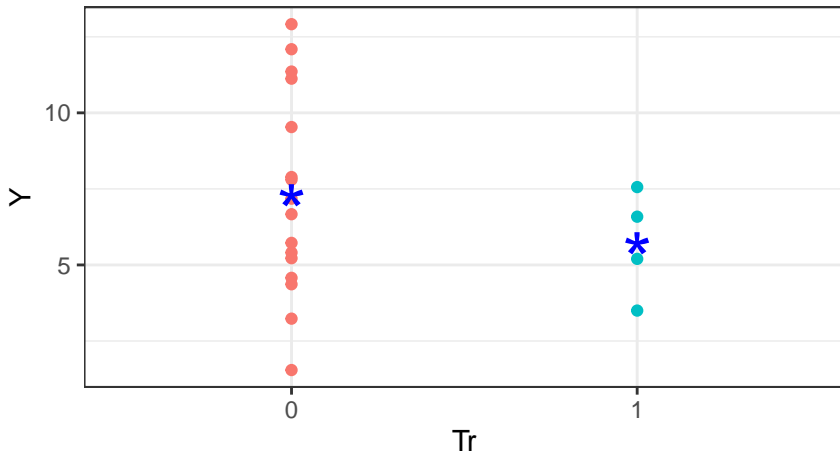
## Exemplo 2: Pequeno Exemplo Numérico

dat

	id	Tr	Z1	Z2	Z3	Y
1	1	1	1	0.25	1.00	3.500000
2	2	1	1	1.95	2.03	5.200673
3	3	1	1	2.24	1.44	6.586071
4	4	1	0	3.54	0.67	7.559968
5	5	0	1	0.26	0.99	1.545138
6	6	0	1	1.95	2.00	3.234772
7	7	0	1	2.26	1.47	4.576648
8	8	0	0	3.50	0.70	5.406685
9	9	0	0	3.56	0.77	5.221591
10	10	0	0	2.58	-0.11	9.533487
11	11	0	0	4.65	0.12	11.357201
12	12	0	0	4.23	-0.06	12.915181
13	13	0	0	4.46	0.90	5.724765
14	14	0	1	3.45	1.66	6.667957
15	15	0	1	4.20	1.87	7.889664
16	16	0	1	2.77	0.61	7.803914
17	17	0	1	3.66	0.24	12.091247
18	18	0	1	4.59	0.69	11.123853
19	19	0	0	3.10	0.90	4.364765
20	20	0	0	3.66	0.38	7.179193

## Exemplo 2: Pequeno Exemplo Numérico

```
ggplot(dat, aes(y = Y, x = Tr, color = Tr)) + geom_point() + theme_bw() +  
  guides(color = "none") + geom_point(data = dat %>% group_by(Tr) %>%  
    summarise(mean = mean(Y)), aes(y = mean, x = Tr), size = 10, shape = '*',  
    col = "blue")
```



## Exemplo 2: Pequeno Exemplo Numérico

```
t.test(Y ~ Tr, data = dat)
```

Welch Two Sample t-test

data: Y by Tr

t = 1.2988, df = 9.3065, p-value = 0.2253

alternative hypothesis: true difference in means between group 0 and

95 percent confidence interval:

-1.156737 4.312889

sample estimates:

mean in group 0 mean in group 1

7.289754 5.711678

```
#
```

```
with(dat, mean(Y[Tr == 1] - mean(Y[Tr == 0])))
```

[1] -1.578076

## Exemplo 2: Pequeno Exemplo Numérico

```
dat %>% group_by(Tr) %>% summarise(n = n(),  
                                   Y = mean(Y),  
                                   Z1 = mean(Z1),  
                                   Z2 = mean(Z2),  
                                   Z3 = mean(Z3))
```

```
# A tibble: 2 x 6
```

	Tr	n	Y	Z1	Z2	Z3
	<fct>	<int>	<dbl>	<dbl>	<dbl>	<dbl>
1	0	16	7.29	0.5	3.30	0.821
2	1	4	5.71	0.75	2.00	1.28

## Exemplo 2: Pequeno Exemplo Numérico

```
round(summary(glm(Tr ~ Z1 + Z2 + Z3, family = "binomial", data = dat))$coef, 3)
```

	Estimate	Std. Error	z value	Pr(> z )
(Intercept)	0.732	2.534	0.289	0.773
Z1	-2.018	2.802	-0.720	0.471
Z2	-1.072	0.756	-1.417	0.156
Z3	1.798	1.798	1.000	0.317

```
round(summary(glm(Y ~ Tr, family = "gaussian", data = dat))$coef, 3)
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	7.290	0.784	9.3	0.00
Tr1	-1.578	1.753	-0.9	0.38

```
round(summary(glm(Y ~ Tr + Z1 + Z2 + Z3, family = "gaussian", data = dat))$coef, 3)
```

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	1.789	1.173	1.526	0.148
Tr1	1.842	0.792	2.326	0.034
Z1	3.736	0.777	4.808	0.000
Z2	2.016	0.278	7.241	0.000
Z3	-3.691	0.578	-6.388	0.000

## Exemplo 2: Pequeno Exemplo Numérico

```
rr <- Matching::Match(Y = dat$Y, Tr = I(dat$Tr == "1"),
                      X = cbind(dat$Z1, dat$Z2, dat$Z3),
                      M = 1)
dat_par <- rbind(dat[rr$index.treated, ], dat[rr$index.control, ])
dat_par$par <- rep(1:length(rr$index.treated), 2)
dat_par <- dat_par %>% arrange(par) %>% print()
```

	id	Tr	Z1	Z2	Z3	Y	par
1	1	1	1	0.25	1.00	3.500000	1
2	5	0	1	0.26	0.99	1.545138	1
3	2	1	1	1.95	2.03	5.200673	2
4	6	0	1	1.95	2.00	3.234772	2
5	3	1	1	2.24	1.44	6.586071	3
6	7	0	1	2.26	1.47	4.576648	3
7	4	1	0	3.54	0.67	7.559968	4
8	8	0	0	3.50	0.70	5.406685	4

```
with(dat_par, mean(Y[Tr == 1] - mean(Y[Tr == 0])))
```

```
[1] 2.020867
```



## Exemplo 2: Pequeno Exemplo Numérico

```
t.test(dat_par$Y[dat_par$Tr == 1], dat_par$Y[dat_par$Tr == 0],  
       paired = FALSE)
```

Welch Two Sample t-test

```
data:  dat_par$Y[dat_par$Tr == 1] and dat_par$Y[dat_par$Tr == 0]  
t = 1.6558, df = 5.9882, p-value = 0.1489  
alternative hypothesis: true difference in means is not equal to 0  
95 percent confidence interval:  
 -0.9669087  5.0086428  
sample estimates:  
mean of x mean of y  
 5.711678  3.690811
```

## Exemplo 2: Pequeno Exemplo Numérico

```
t.test(dat_par$Y[dat_par$Tr == 1], dat_par$Y[dat_par$Tr == 0],  
       paired = TRUE)
```

Paired t-test

```
data:  dat_par$Y[dat_par$Tr == 1] and dat_par$Y[dat_par$Tr == 0]  
t = 44.237, df = 3, p-value = 2.543e-05  
alternative hypothesis: true mean difference is not equal to 0  
95 percent confidence interval:  
 1.875484 2.166250  
sample estimates:  
mean difference  
 2.020867
```

## Exemplo 2: Pequeno Exemplo Numérico

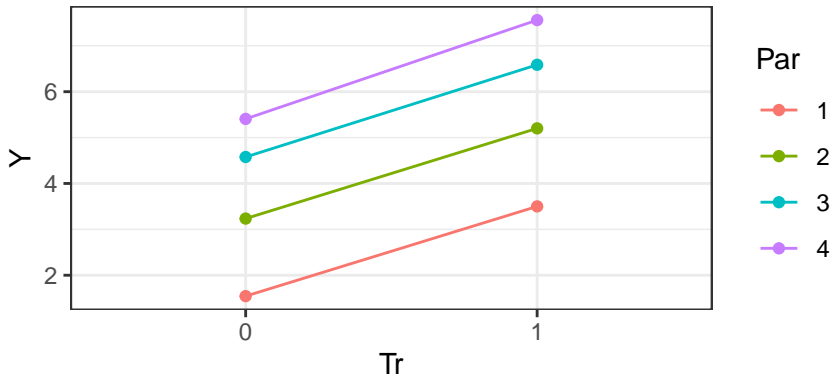
```
summary(rr)
```

```
Estimate... 2.0209  
AI SE..... 0.039562  
T-stat..... 51.08  
p.val..... < 2.22e-16
```

```
Original number of observations..... 20  
Original number of treated obs..... 4  
Matched number of observations..... 4  
Matched number of observations (unweighted). 4
```

## Exemplo 2: Pequeno Exemplo Numérico

```
ggplot(dat_par, aes(x = Tr, y = Y, color = factor(par))) + geom_point() +  
  geom_line(aes(group = par, color = factor(par))) + theme_bw() +  
  labs(color = "Par")
```



```
cor(dat_par$Y[dat_par$Tr == 1], dat_par$Y[dat_par$Tr == 0])
```

[1] 0.9995821

# Exemplo 3: Amiodarona e Doença de Chagas

International Journal of Cardiology 189 (2015) 182–184



Contents lists available at [ScienceDirect](#)

International Journal of Cardiology

journal homepage: [www.elsevier.com/locate/ijcard](http://www.elsevier.com/locate/ijcard)



Letter to the Editor

## Amiodarone and *Trypanosoma cruzi* parasitemia in patients with Chagas disease



Andre A.L. Carmo<sup>a</sup>, Manoel O.C. Rocha<sup>a</sup>, Jose L.P. Silva<sup>b</sup>, Barbara M. Ianni<sup>c</sup>, Fabio Fernandes<sup>c</sup>, Ester C. Sabino<sup>d</sup>, Antonio L.P. Ribeiro<sup>a,\*</sup>

<sup>a</sup> Hospital das Clínicas and Faculdade de Medicina, Universidade Federal de Minas Gerais, Minas Gerais, Brazil

<sup>b</sup> Department of Statistics, Universidade Federal de Minas Gerais, Minas Gerais, Brazil

<sup>c</sup> Cardiomyopathy Unit of the Heart Institute (InCor), da Faculdade de Medicina, da Universidade de São Paulo, São Paulo, Brazil

<sup>d</sup> Department of Infectious Diseases, Institute of Tropical Medicine, Universidade de São Paulo, São Paulo, Brazil

## Exemplo 3: Amiodarona e Doença de Chagas

Doença de Chagas (ChD) é uma das principais causas de cardiopatia na América Latina, causando milhares de mortes todo ano.

Amiodarona tem sido amplamente usada para prevenir arritmias ventriculares em pacientes com Doença de Chagas.

Há algumas evidências que a amiodarona tem também efeito anti-*Trypanosoma cruzi*.

O objetivo do estudo é verificar, em pacientes com cardiomiopatia chagásica, se aqueles usando amiodarona têm menor carga parasitária em comparação com quem não o usa.

# Exemplo 3: Amiodarona e Doença de Chagas

**Table 1**  
Patient characteristics before and after Genetic Matching.

	Unmatched			Genetic Matching		
	Amiodarone users N = 37	Non-amiodarone users N = 564	Std mean diff (%)	Amiodarone users N = 37	Non-amiodarone users N = 37	Std mean diff (%)
<i>General features</i>						
Age – years	49.7	48.2	4.1	49.7	49.7	0.0
Gender – % male	33.3	47.3	– 33.0	33.3	27.1	16.1
Years of exposure	17.6	17.7	– 0.6	17.6	17.4	1.7
NYHA Functional Class – % in class II	16.3	6.4	45.7	16.3	18.4	– 6.4
Cardiomyopathy – %	91.7	40.1	70.2	91.7	91.3	0.6
Beta-blockers – %	72.2	16.8	96.3	72.2	72.4	– 0.3
ACE inhibitors or angiotensin-receptor blockers – %	80.6	24.5	86.9	80.6	78.0	3.9
<i>Electrocardiogram</i>						
QRS duration – ms	138.6	102.1	35.8	138.6	134.5	3.1
Major ECG abnormalities – %	16.7	4.8	62.0	16.7	14.4	8.7
Minor ECG abnormalities – %	5.6	4.3	8.8	5.6	6.4	– 3.7
<i>Echocardiogram</i>						
Left ventricular ejection fraction – %	31.8	56.7	– 102.6	31.8	32.3	– 2.0
Left ventricular end-diastolic diameter – mm	63.1	47.9	33.7	63.1	61.8	2.9
Segmental wall-motion abnormality – %	85.8	21.1	96.3	85.8	85.9	– 0.1
Left ventricular septal E/e' ratio	11.8	7.9	43.9	11.8	12.4	– 6.4
Left ventricular lateral E/e' ratio	9.7	6.0	51.5	9.7	9.1	8.5
<i>Natriuretic peptides</i>						
NT <sub>pro</sub> BNP (pg/mL)	2191.9	371.5	86.1	2191.9	1888.3	13.7

Data expressed in mean values or proportions.

# Bibliografia



# Principais Referências Usadas

Kruttschnitt, C., et al. “Family violence, television viewing habits, and other adolescent experiences related to violent criminal behavior”. *Criminology*, 1986.

Carmo, A. A., et al. “Amiodarone and Trypanosoma cruzi parasitemia in patients with Chagas disease.” *International Journal of Cardiology*, 2015.

Bacchieri, A. and Della, Cioppa, G., “Fundamentals of Clinical Research, Bridging Medicine, Statistics and Operations”, *Springer-Verlag*, 2007.

Rosenbaum, Paul, R., “Design of Observational Studies”, *Springer-Verlag*, 2009.