

# **E**STUDOS DE CASO-CONTROLE

<b>INTRODUÇÃO</b> .....	<b>42</b>
<b>SELEÇÃO DE PARTICIPANTES</b>	
<b>Seleção de casos</b> .....	<b>42</b>
<b>Seleção de controles</b> .....	<b>42</b>
<b>TIPOS DE ESTUDOS</b>	
<b>Caso-controle de base populacional</b> .....	<b>42</b>
<b>Caso-controle aninhado</b> .....	<b>43</b>
<b>MEDIDA DE ASSOCIAÇÃO</b>	
<b>Análise univariada</b> .....	<b>43</b>
<b>METODOLOGIA</b>	
<b>Amostragem</b> .....	<b>44</b>
<b>Tamanho da amostra</b> .....	<b>44</b>
<b>Medida da exposição</b> .....	<b>45</b>
<b>Tipos de viés</b> .....	<b>45</b>
<b>ANÁLISE ESTRATIFICADA</b>	
<b>Confusão</b> .....	<b>46</b>
<b>Estratificação</b> .....	<b>46</b>
<b>Modificação de efeito</b> .....	<b>47</b>
<b>EMPARELHAMENTO</b> .....	<b>49</b>
<b>VANTAGENS E LIMITAÇÕES</b> .....	<b>50</b>
<b>ROTEIRO PARA DELINEAR UM ESTUDO DE CASO-CONTROLE</b> .....	<b>52</b>
<b>REFERÊNCIAS PARA LEITURA</b> .....	<b>53</b>
<b>EXERCÍCIOS</b> .....	<b>54</b>
<b>DICIONÁRIO DE BANCO DE DADOS</b> .....	<b>60</b>

## INTRODUÇÃO

É uma pesquisa, de modelo retrospectivo, na qual os participantes são selecionados entre indivíduos que já têm a doença (**casos**) e entre indivíduos que não a têm (**controles**); em cada um desses dois grupos, verifica-se o número de indivíduos expostos, a algum fator de risco. O objetivo é verificar a possível existência de associação causal entre a exposição aos fatores de risco e a doença em estudo. Se o fator está associado à doença, a proporção do fator entre os casos será maior que a mesma proporção entre os controles. Este tipo de estudo tem grande aplicação para as situações em que a doença é relativamente pouco freqüente e o tempo decorrido entre a exposição ao risco e a evidenciação do seu efeito, é longo. Os estudos caso-controle têm limitada implicação ética uma vez que não existe intervenção nem observação prospectiva de exposições de risco. Os estudos tipo caso-controle, inicialmente propostos para o estudo de doenças crônico-degenerativas, principalmente câncer, também têm aplicação no estudo de doenças infecciosas.

## SELEÇÃO DE PARTICIPANTES

**Seleção de casos** - O local de obtenção de casos e controles depende das características da doença em estudo. Casos podem ser identificados em hospitais, clínicas especializadas ou serviços de saúde (ex.: casos de hanseníase e de tuberculose). É possível fazer busca populacional de casos, mediante inquérito baseado em marcadores, como níveis de anticorpos.

**Seleção de controles** - A busca de controles deve seguir, como orientação geral, o princípio "*se o controle presente fosse um caso, ele seria encontrado onde os casos estão sendo encontrados*". Controles podem ser recrutados nos hospitais onde os casos foram obtidos, na vizinhança dos casos, nas mesmas escolas, entre amigos e colegas de trabalho dos casos, na população em geral sob esquema de amostra probabilística. Em qualquer situação haverá vantagens e desvantagens, sempre com possibilidade de resultados viesados. Controles obtidos por **sugestão dos próprios casos** podem ser muito semelhantes nos seus comportamentos e costumes, e, se o fator de risco estudado está relacionado a hábitos que podem ser comuns entre amigos, não será detectado. O custo e dificuldade operacional de obtenção de **controles populacionais** tornam esta abordagem pouco prática. No contexto das doenças infecciosas formas sub-clínicas e clínicas da doença podem ser detectadas. A estratégia a ser adotada para seleção do grupo controle depende do objetivo do estudo. Por exemplo, se a finalidade é avaliar fatores de risco para malária grave ou complicada (casos) o grupo controle deverá ser formado por indivíduos com parasitemia assintomática, ou formas leves de malária. Se o objetivo do estudo é determinar fatores de risco prognósticos para o desenvolvimento de cardiopatia chagásica, indivíduos soropositivos e com alterações ECG compatíveis com doença de Chagas serão selecionados como casos e dois tipos de controles, poderiam ser selecionados: a) indivíduos soronegativos e b) indivíduos soropositivos, ambos sem alterações ECG. Se indivíduos soronegativos são os controles (a), o estudo determinaria riscos de desenvolvimento de doença de Chagas enquanto para o segundo tipo de controles o estudo estaria avaliando fatores associados à evolução da infecção para doença.

## TIPOS DE ESTUDOS

. **Caso-controle de base populacional** - Neste tipo de delineamento casos e controles são selecionados da população; casos podem ser detectados através de triagem populacional, em área geográfica definida durante um determinado período de tempo. Pode-se utilizar registros de hospitais para se identificar todos os possíveis casos da área de estudo ou uma amostra aleatória deles. Os controles são selecionados através de uma amostra probabilística dos indivíduos sem a doença, pertencentes à mesma área geográfica dos casos.

. **Caso-controle aninhado** - É um delineamento no qual casos e controles são selecionados no decorrer de uma coorte pré-definida, na qual algumas informações sobre exposições e fatores de risco já encontram-se disponíveis. Para cada caso, controles são selecionados aleatoriamente de indivíduos que encontram-se sob risco no momento do diagnóstico do caso, o que significa um emparelhamento pelo efeito de confusão do tempo. Informações adicionais são coletadas e analisadas no momento da seleção de casos incidentes e controles.

## MEDIDA DE ASSOCIAÇÃO

### . Análise univariada

A estatística usada como medida de associação é o *odds ratio* (**OR**), que, para doenças com tempo longo de latência, é uma estimativa da razão das forças de morbidade entre expostos e não expostos ao fator de risco. O *odds ratio* é uma aproximação da razão de riscos (risco relativo), quando as incidências são baixas e uma aproximação da razão de prevalência, quando as prevalências são baixas. Quando casos e controles são selecionados da população geral, a proporção de expostos ao fator de risco no grupo controle pode ser utilizada como uma estimativa da proporção de expostos na população em geral. Esta vantagem propicia que se calcule o risco atribuível percentual (populacional), que expressa a proporção da doença na população de estudo que é atribuível à exposição ou fator e que poderia ser eliminada se o fator fosse removido. Esta é uma informação útil pois indica quais exposições ou fatores são os mais relevantes e devem prioritariamente ser eliminados como medidas de saúde pública.

O resultado amostral de um estudo caso-controle pode ser expresso em uma tabela 2x2:

EXPOSIÇÃO	CONDIÇÃO		TOTAL
	caso	controle	
Presente	a	b	a + b
Ausente	c	d	c + d
TOTAL	a + c	b + d	T

Onde:

- a + c** = número de casos
- a** = número de casos com o fator de risco presente
- c** = número de casos com o fator de risco ausente
- b + d** = número de controles
- b** = número de controles com o fator de risco presente
- d** = número de controles com o fator de risco ausente
- a + b** = número total de indivíduos que estiveram expostos ao fator de risco
- c + d** = número que não esteve exposto ao fator de risco
- T** = total do conjunto das amostras de casos e controles

*Odds* é uma medida de probabilidade, definido como o quociente de duas probabilidades, complementares entre si. Na tabela, o *odds* de presença de exposição ao fator, entre os casos, é  $a/c$ ; o *odds* de presença de exposição ao fator, entre os controles, é  $b/d$ . O quociente entre estes dois *odds* é o *odds ratio* (OR):  $(a/c) / (b/d) = (a.d) / (b.c)$ ; o *odds ratio*, sob a forma desta última expressão, é

também denominado “razão dos produtos cruzados”. Em português, *odds ratio* pode ser traduzido, literalmente, como "razão do quociente de probabilidades".

A associação entre o fator de risco e a doença pode ser especificada como positiva ou negativa, situação que leva a um teste estatístico monocausal; não sendo especificada, tem-se um teste bicaudal.

OR=1 indica que probabilidade de doença nos expostos ao fator de risco e nos não expostos são equivalentes. OR>1 indica que a exposição ao fator em estudo é de risco, podendo implicar em relação causa efeito. A decisão estatística é baseada na estatística  $\chi^2$  (qui quadrado), com um grau de liberdade. Se o nível de significância for  $\alpha = 5\%$  e o teste for bicaudal,  $\chi^2$  é comparado com 3,84; se maior indica uma associação estatisticamente significativa.

Como exemplo, em um surto epidêmico de diarreia em participantes de uma oficina de trabalho na Fundação Nacional de Saúde, um estudo de caso-controle foi conduzido para investigar o risco de intoxicação alimentar associado com ingestão de maionese. Casos e controles foram identificados através de entrevistas com indivíduos que comeram no buffet servido durante os dias da oficina. Os dados estão dispostos na tabela abaixo com os resultados do *odds ratio* e o intervalo da estimativa da *odds ratio*, com 95% de confiança.

INGESTÃO DE MAIONESE	DIARRÉIA		TOTAL
	Sim	Não	
Sim	75	152	227
Não	10	140	150
<b>TOTAL</b>	<b>85</b>	<b>292</b>	<b>377</b>

OR=6,91 (IC 95% 3,4-13,9)     $\chi^2=35,85$

## METODOLOGIA

### . Amostragem

Apesar de não se ter que utilizar formalmente nenhuma técnica de amostragem, deve-se verificar em que extensão o grupo escolhido representa todos os indivíduos com a doença, não só para interpretação dos resultados para os indivíduos incluídos no estudo (**validade interna do estudo**), como também para se poder fazer inferências, extrapolações para indivíduos não incluídos no estudo (**validade externa do estudo**)

### . Tamanho da amostra

O número de casos e controles a ser selecionado, depende do **tamanho da amostra** necessário para testar uma hipótese. De maneira geral, o tamanho da amostra é inversamente proporcional ao risco que se pretende detectar. Para se detectar um pequeno risco relativo (ex.: 1,2 ou seja, 20% de aumento no risco comparado ao grupo controle) é necessário incluir um grande número de

participantes. Estudos conduzidos com poucos casos apresentam baixo poder estatístico para detectar riscos.

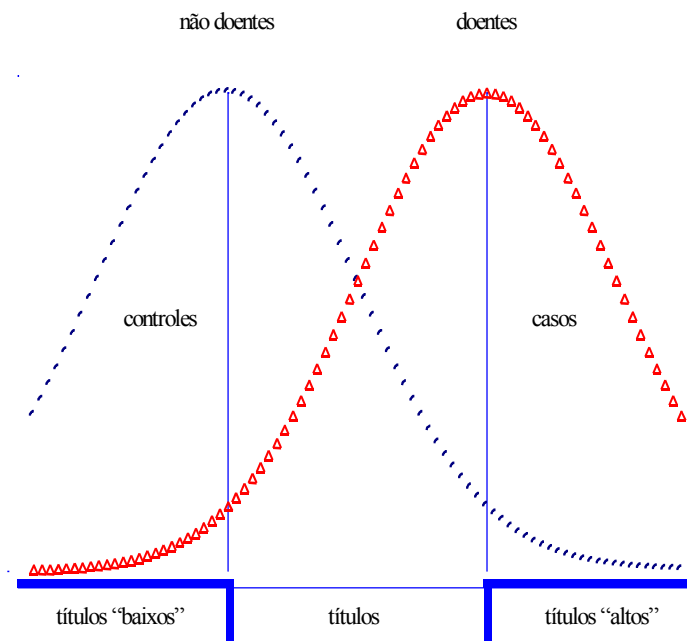
Para o cálculo do tamanho da amostra as seguintes informações são necessárias:

- (1) nível de significância do teste (geralmente  $\alpha=5\%$ )
- (2) o poder do teste (geralmente  $1-\beta=80\%$ )
- (3) a proporção de pessoas expostas ao fator de risco na população geral
- (4) o valor de *odds ratio* mínimo a ser detectado
- (5) razão entre o número de controles e casos

Por exemplo, para  $\alpha = 5\%$ , poder to teste =  $80\%$  ( $1-\beta$ ),  $10\%$  de expostos ao fator de risco na população geral e  $OR = 2$ , um total de 307 casos e 307 controles devem ser selecionados (ver exercícios em EPIINFO).

### . Medida da exposição

Para avaliar as medidas de exposição pode-se utilizar entrevistas, questionários padronizados, informações de parentes/vizinhos ou, marcadores biológicos. Os procedimentos devem ser iguais tanto para casos como para os controles. Os entrevistadores devem desconhecer o “status” caso-control do indivíduo para assegurar o caráter de mascaramento, minimizando-se o **viés do observador**.



### Tipos de viés

**Viés de classificação** - Estudos de caso-control estão sujeitos à diferentes fontes de bias. A classificação de indivíduos como casos, por terem a doença de estudo, e como controles, por não terem a doença deve seguir critérios bem definidos, a fim de evitar viés de classificação. Este é um

erro sistemático pelo qual doentes são selecionados como controles e indivíduos sem a doença são selecionados como caso. Exames laboratoriais, altamente sensíveis e específicos, são desejáveis, em complementação a diagnósticos clínicos; se o nível de anticorpos for usado, pode-se estabelecer dois pontos de corte na escala de títulos, o menor deles como limite superior para seleção de controles e o maior deles como limite inferior para seleção de casos. Este procedimento visa a minimizar a classificação de não doentes como casos (falsos casos) e a de doentes como controles (falsos controles); a figura acima mostra esquematicamente a situação.

**Viés de seleção** - Outra possibilidade de vícios em estudos caso-controle refere-se ao viés de seleção por incorreções ou limitações no delineamento do estudo, afetando a comparabilidade de casos e controles. Um dos motivos deve-se ao fato de que os princípios da comparabilidade na seleção de controles geralmente são influenciados pelo princípio da eficiência, ou seja, disponibilidade de recursos e tempo.

**Viés do observador** - Ao selecionar os controles deve-se assegurar que as observações sejam realizadas nos dois grupos sob as mesmas condições. O ideal seria que o investigador não tivesse o conhecimento de quem tem a doença e quem não tem, ou seja, quem é caso e quem é controle, para evitar tendenciosidades na coleta de informações (ou seja, o viés do observador).

**Casos prevalentes** - Outra fonte potencial de bias é a seleção de casos prevalentes ao invés de casos incidentes. A prevalência é afetada pela duração da doença que é influenciada pelo tratamento e cura e também pela mortalidade associada à doença. Quando se inclui casos prevalentes os fatores podem estar estatisticamente associados à doença devido a um "efeito de sobrevivência" e duração da doença e não devido à uma associação causal. Na infecção pelo *T. cruzi*, por exemplo, um percentual de óbitos ocorre na fase aguda de infecção, e portanto, fatores de risco avaliados para casos selecionados na fase crônica da infecção (casos prevalentes) estariam associados à sobreviventes da fase aguda.

## ANÁLISE ESTRATIFICADA

**Confusão** - Quando um fator está associado simultaneamente à exposição e à doença é chamado de variável de confusão. Confusão é uma distorção, causada por outra variável C, no resultado numérico que mede a associação entre uma variável E (exposição) e a condição D (doença), estando a variável C associada à variável E e à variável D. A confusão é um viés que deve ser controlado, o que pode ser feito na análise, mediante análise estratificada.

**Estratificação** - Se uma possível variável de confusão C tiver dois estratos, com C presente e C ausente, a associação da variável E com a condição D deverá ser examinada segundo estes estratos (análise estratificada), além da análise conduzida sem considerar C (análise bruta); o resultado está esquematizado nas tabelas a seguir:

### Análise bruta

EXPOSIÇÃO AO FATOR	CONDIÇÃO (D)		TOTAL
	caso	controle	
Presente	a	b	a + b
Ausente	c	d	c + d
TOTAL	a + c	b + d	T

**Análise Estratificada**

**ESTRATO1 (C presente)**

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso	controle	
Presente	a <sub>1</sub>	b <sub>1</sub>	a <sub>1</sub> + b <sub>1</sub>
Ausente	c <sub>1</sub>	d <sub>1</sub>	c <sub>1</sub> + d <sub>1</sub>
<b>TOTAL</b>	<b>a<sub>1</sub> + c<sub>1</sub></b>	<b>b<sub>1</sub> + d<sub>1</sub></b>	<b>T<sub>1</sub></b>

**ESTRATO 2 (C ausente)**

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso	controle	
Presente	a <sub>2</sub>	b <sub>2</sub>	a <sub>2</sub> + b <sub>2</sub>
Ausente	c <sub>2</sub>	d <sub>2</sub>	c <sub>2</sub> + d <sub>2</sub>
<b>TOTAL</b>	<b>a<sub>2</sub> + c<sub>2</sub></b>	<b>b<sub>2</sub> + d<sub>2</sub></b>	<b>T<sub>2</sub></b>

Na análise bruta  $OR_{bruto} = (a.d) / (b.c)$ . Na análise estratificada, o subscrito indica o estrato; o *odds ratio* do primeiro estrato é:

$OR_1 = (a_1.d_1) / (b_1.c_1)$  e o do segundo estrato é:

$OR_2 = (a_2.d_2) / (b_2.c_2)$

A medida de associação comum, *odds ratio* sintético, é:

$$OR_{MH} = \frac{((a_1.d_1)/T_1 + (a_2.d_2)/T_2)}{((b_1.c_1)/T_1 + (b_2.d_2)/T_2)}$$

Esta é a média ponderada dos *odds ratios* dos dois estratos, com pesos proporcionais aos logaritmos naturais das variâncias. O subscrito MH refere-se aos autores Mantel e Haenszel, que desenvolveram este estimador da associação comum. Se  $OR_{bruto} \neq OR_{MH}$  existe confusão, podendo, ainda, haver modificação de efeito. Se  $OR_{bruto} = OR_{MH}$  não há confusão, podendo, porém, haver modificação de efeito. A decisão sobre a diferença entre os *odds ratio* indicadora de confusão é arbitrária, não devendo ser feito teste de hipóteses para isto; uma conduta possível é estabelecer-se uma porcentagem limite para a diferença entre as medidas.

**Modificação de efeito** - Diz-se que há modificação de efeito quando as medidas de associação entre a variável E e a variável D nos estratos de C são diferentes, indicando processos desiguais de causalidade, segundo as características da variável C. Ao contrário da situação de confusão, no caso de modificação de efeito não se deve ter uma medida resumo de associação, mas sim, uma comparação das medidas de associação (OR) para cada estrato. A decisão se há ou não modificação

de efeito é tomada estatisticamente mediante o teste de heterogeneidade ou interação estatística. As hipóteses a serem testadas são:

$H_0$ : há homogeneidade entre os estratos

$H_A$ : há heterogeneidade entre os estratos

O teste de hipótese, na situação de análise estratificada, assumindo  $H_0$ : homogeneidade entre os estratos (ou seja, não haver modificação de efeito), é conduzido para o conjunto de todos os estratos, utilizando-se uma estatística com distribuição  $\chi^2$ , cujo grau de liberdade corresponde ao número de estratos menos 1.

### Exemplo

Continuando a análise do estudo da associação entre ingestão de maionese e diarreia, incluindo uma variável eventual de confusão, com duas possibilidades,  $< 30$  anos e  $\geq 30$  anos, teríamos:

#### ESTRATO1 (idade $< 30$ anos)

INGESTÃO DE MAIONESE	DIARRÉIA		TOTAL
	Sim (caso)	Não (controle)	
Sim	74	120	194
Não	5	54	59
<b>TOTAL</b>	<b>79</b>	<b>174</b>	<b>253</b>

OR=6,66 (exato IC95% 2,50-22,18)

#### ESTRATO 2 (idade $\geq 30$ anos)

INGESTÃO DE MAIONESE	DIARRÉIA		TOTAL
	Sim (caso)	Não (controle)	
Sim	1	32	33
Não	5	86	91
<b>TOTAL</b>	<b>6</b>	<b>118</b>	<b>124</b>

OR=0,54 (exato IC95% 0,01-5,10)

$OR_{MH} = 4,50$  (IC 95% 1,84-9,08)

$OR_{bruto} = 6,91$  (IC 95% 3,30-14,84)

Os resultados indicam associação entre consumo de maionese e diarreia. O teste de heterogeneidade foi significativo ( $p=0,00018$ ), o que significa que  $< 30$  anos e  $\geq 30$  anos devem ser trabalhados separadamente, pois eles tem riscos diferentes.



## EMPARELHAMENTO

Há estudos em que, para cada caso, é escolhido um controle (às vezes dois, ou três, etc), como um irmão, um vizinho, um companheiro de trabalho, com a finalidade de garantir uma maior comparabilidade entre os grupos de casos e de controles ou controlar fatores comuns não facilmente identificáveis. Idade e sexo são, em geral, consideradas variáveis que estão intimamente associadas à possibilidade de exposição e ao desenvolvimento da doença. Por este motivo, casos e controles são usualmente selecionados dentro da mesma faixa etária e sexo. A estratégia de levar em conta a variável de confusão já no delineamento do estudo, é denominada **emparelhamento**. Este emparelhamento dos casos e controles por sexo e idade torna os grupos mais comparáveis e minimiza potenciais distorções dos resultados na avaliação de riscos. O emparelhamento leva a ter-se um número de estratos igual ao número de casos, cada estrato composto de 1 caso e seu respectivo controle (ou controles). O princípio da análise é o mesmo, mas a variável que serviu como base do emparelhamento não pode ser analisada.

Para a situação 1 caso/1 controle, tem-se as seguintes possíveis configurações, para cada par caso/controle:

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso(1)	controle(0)	
Presente(1)	1	1	2
Ausente(0)	0	0	0
<b>TOTAL</b>	<b>1</b>	<b>1</b>	<b>2</b>

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso(1)	controle(0)	
Presente(1)	1	0	1
Ausente(0)	0	1	1
<b>TOTAL</b>	<b>1</b>	<b>1</b>	<b>2</b>

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso(1)	controle(0)	
Presente(1)	0	1	1
Ausente(0)	1	0	1
<b>TOTAL</b>	<b>1</b>	<b>1</b>	<b>2</b>

EXPOSIÇÃO AO FATOR	CONDIÇÃO		TOTAL
	caso(1)	controle(0)	
Presente(1)	0	0	0
Ausente(0)	1	1	2
<b>TOTAL</b>	<b>1</b>	<b>1</b>	<b>2</b>

O resultado para k pares pode ser expresso na seguinte tabela resumo:

CASO	CONTROLE	
	Exposto (1)	Não exposto (0)
Exposto(1)	$v_{11}$	$v_{10}$
Não Exposto(0)	$v_{01}$	$v_{00}$

onde:

$v_{11}$  = o número de pares em que o caso e o respectivo controle estiveram expostos ao fator de risco

$v_{10}$  = o número de pares em que o caso esteve exposto e o controle não

$v_{01}$  = o número de pares em que o caso não esteve exposto e o controle esteve

$v_{00}$  = o número de pares em que nem o caso nem o controle sofreram exposição ao fator de risco.

O OR é calculado como a razão entre os pares discordantes:

$$OR = \frac{v_{10}}{v_{01}}$$

A situação em que o número de controles varia para cada caso pode ser tratada mais facilmente como na situação de estratificação, embora o número de estratos naturalmente seja bem grande.

## VANTAGENS E LIMITAÇÕES

**.Vantagens** - Estudos de caso-controle são particularmente adequados para avaliação de doenças raras constituindo-se em método rápido, prático e de baixo custo para testar o efeito e interação de um grande número de fatores que se relacionam com o evento estudado. Ao contrário dos estudos de coorte, pelo fato de não necessitar de um grande número de participantes, pode-se empregar exames/testes mais dispendiosos e laboriosos. Também, evita perdas de seguimento de casos.

*Estudos de caso-controle e os programas de saúde pública* - Em saúde pública a metodologia caso-controle tem sido utilizada principalmente para avaliação de eficácia de vacinas, em investigações de epidemias, para avaliar testes de triagens populacionais, avaliação de impacto de medidas preventivas como programas de saúde pública, e eficácia de tratamentos. Na eficácia de vacinas

após a implantação em programas de rotinas a informação sobre o estado vacinal dos indivíduos é o fator de exposição a ser avaliado como associado ou não à doença. A eficácia da vacina (EV) pode ser estimada pela fórmula  $EV=1-OR$ .

**Limitações** - As principais limitações de estudos de caso-controle são: a) não adequados para investigações de exposições raras, a menos que o risco atribuído à exposição na população seja muito alto; b) não estima a incidência das doenças/infecções estudadas e, c) a informação sobre a exposição ou fator é obtida após a ocorrência da doença e portanto, não há como se distinguir uma cronologia nítida entre a exposição e o aparecimento da doença, diferente do que ocorre nos estudos de coorte onde a exposição é o ponto de partida. Por exemplo, para avaliação da associação entre estado nutricional e crianças infectadas por *T. cruzi*, se a soropositividade e o grau de desnutrição são medidos ao mesmo tempo, fica difícil determinar se a desnutrição levou a uma maior susceptibilidade à infecção pelo *T. cruzi*, ou se as crianças infectadas desenvolvem mais desnutrição.

## ROTEIRO PARA DELINEAR UM ESTUDO DE CASO-CONTROLE

### ① *Defina de forma clara a questão que o estudo pretende responder*

- . esclareça nos objetivos a doença/ evento, a população que será investigada e os fatores que serão testados como associados ao evento

### ② *Estabeleça a definição de caso e de controle*

- . faixa etária, sexo
- . testes laboratoriais, exames clínicos, parâmetros a serem avaliados, interpretação
- . critérios de inclusão (casos graves, leves) e exclusão
- . número de controles necessários por caso
- . necessidade de emparelhamento de casos e controles

### ③ *Especifique o critério de amostragem*

- . defina a população de referência
- . esclareça a fonte de seleção dos casos e dos controles: serviços de saúde, população
- . forneça informações sobre a validade interna e externa do estudo

### ④ *Defina as variáveis de exposição e as variáveis potenciais de confusão que serão controladas na análise*

- . descreva os métodos laboratoriais, parâmetros a serem avaliados, interpretação, exames clínicos, entrevistas
- . esclareça os procedimentos utilizados para minimizar os possíveis vieses do observador - mascaramento

### ⑤ *Calcule o tamanho da amostra necessário*

- . defina a proporção aproximada de exposição ao fator de risco na população geral (controles)
- . calcule o número mínimo e suficiente de casos e de controles necessário para estimar o risco relativo definido pelo estudo
- . defina o nível de significância e o poder estatístico do estudo com o número de participantes que serão estudados
- . certifique-se da factibilidade do estudo, em termos logístico e de tempo de recrutamento dos casos e controles

### ⑥ *Descreva as etapas da análise de dados*

- . indique os parâmetros (proporções, médias), métodos estatísticos e sub-grupos para avaliação das características básicas dos casos e controles e avaliação dos fatores de risco (odds ratio e IC95%)
- . descreva o tipo de análise - univariada, estratificada, multivariada

## REFERÊNCIAS PARA LEITURA

ARMENIAN HK (Ed.) Applications of the Case-Control Method. **Epidemiologic Review**,16(1). 1994.

BRESLOW, N.E. & DAY, N.E. *Statistical methods in cancer research*. Switzerland, International Agency for Research on Cancer, Lyon, V.1. Scientific publication N°. 32, 1980.

HENNECKENS, H.C. & BURING, J.E. *Epidemiology in Medicine*, 5<sup>th</sup> ed. Boston: Toronto, Ed. Little, Brown and Company, 1987.

KLEINBAUM, D.G., KUPPER, L.L. & MORGENSTERN, H. *Epidemiologic Research*. New York, Ed. Van Nostrand Reinhold, 1982.

KOPEC, J.A. & ESDAILE, J.M. Bias in case-control studies. A review. **Journal of Epidemiology and Community Health**,44:179-86, 1990.

MANTEL, N. & HAENSZEL, W. Statistical aspects of the analysis of data from retrospective studies of disease. **Journal Natl Cancer Inst**,22:719-48, 1959.

PEREIRA, M.G. *Epidemiologia - Teoria e Prática*. Guanabara Koogan, Rio de Janeiro, 1995.

RODRIGUES, L. & KIRKWOOD, B.R. Case-control designs in the study of common diseases: updates on the demise of the rare disease assumption and the choice of sampling scheme for controls. **International Journal of Epidemiology**,19:205-13, 1990.

SCHLESSELMAN, J.J. *Case-control studies - Design, conduct and analysis*, New York:Oxford University Press, 1982.

SMITH, P.G. & DAY, N.E. The design of case-control studies: The Influence of confounding and interaction effects. **Int. J. Epidemiol.**,13:356-65, 1984.

WACHOLDER, S., MCLAUGHLIN, J.K., SILVERMAN, D.T. & MANDEL, J.S. Selection of controls in case-control studies. **American Journal of Epidemiology**, 135(9):1019-50, 1992.

## EXERCÍCIOS

**Arquivos:**

1. **casoch.rec**
2. **casach.rec**
3. **casohan.rec**

1. **Fatores de risco para infecção chagásica** - um estudo de caso-controle de base populacional foi delineado para avaliar fatores de risco associados à infecção pelo *T.cruzi* na infância, em municípios do nordeste do Estado de Goiás/Brasil. Um mil novecentos e noventa (1990) crianças de 7-12 anos foram triadas sorologicamente para seleção de 149 crianças soropositivas (casos) e 298 soronegativas (controles). Casos e controles foram emparelhados por sexo, frequência de idade e localidade. Detalhes da metodologia estão na referência **Andrade et al, 1995**. A variável "NOPAR" indica o número do triplo, incluindo caso, controle 1 e controle 2. A análise adotada teve como objetivos: (1) verificar a comparabilidade entre casos e controles em relação às suas características básicas; (2) avaliar o risco associado à presença de moradia infestada por triatomíneos para infecção, e (3) avaliar a associação entre filho soropositivo e soropositividade dos pais. Utilize o arquivo **casoch.rec** para responder as questões abaixo.

**Questão 1.** **Construa uma tabela comparando as características básicas dos casos e controles, incluindo sexo, grupo etário "GRAGE" 7-9 e 10-12 anos, média de idade da criança, município de residência, antecedentes mórbidos e hospitalização. De acordo com os resultados da tabela o que pode ser concluído em relação a comparabilidade entre casos e controles?**

**Notas 1:**

```
READ CASOCH.REC (para abrir o arquivo)
TABLES SEX CACO
criar a variável GRAGE (grupo etário)
DEFINE GRAGE #
IF AGE >= 7 AND AGE <= 9 THEN GRAGE=1 ELSE
GRAGE=2
TABLES GRAGE CACO
MEANS AGE CACO /N
TABLES MUN CACO
para as variáveis MORB (antecedentes mórbidos) e HOSP
(hospitalização) excluir código 9 (Sem informação)
SELECT MORB <>9
TABLES MORB CACO
SELECT (para desativar a seleção)
SELECT HOSP <>9
TABLES HOSP CACO
SELECT
```

**Questão 2.** Calcule o *odds ratio* (OR) emparelhado e respectivo intervalo de 95% de confiança (IC 95%) para "moradia infestada por triatomíneo" ("BARB"), captura de triatomíneo no domicílio ("DCAP"), sorologia da mãe e sorologia do pai. Crie a variável "BARB" correspondente à "moradia infestada por triatomíneo". Utilize as variáveis "DCAP" e "DVTR" para obter "BARB"=1 (DCAP=1 OU DVTR=1) e "BARB"=2 (DCAP=2 E DVTR=2). Crie uma nova variável "GRNPES" ("numero de pessoas"). Sugestão: de 1-4 pessoas, de 5-8 e  $\geq 9$  pessoas por moradia. Para a variável "GRNPES" escolha o primeiro estrato (1-4 pessoas) como referência e calcule o OR (não emparelhado) e IC95% dos demais estratos comparando com a categoria de referência.

**Notas 2:**

```
criar a variável "BARB"
DEFINE BARB #
IF DCAP=1 OR DVTR=1 THEN BARB=1 ELSE BARB=2
comando MATCH para o cálculo do OR emparelhado.
MATCH NOPAR CACO BARB
MATCH NOPAR CACO DCAP
para as variáveis SORMAE (sorologia da mãe) e SORPAI (sorologia do pai)
excluir código 9 (Não realizada)
SELECT SORMAE <>9
MATCH NOPAR CACO SORMAE
SELECT
SELECT SORPAI <> 9
MATCH NOPAR CACO SORPAI
SELECT
criar variável GRNPES (categorias do número de pessoas na família)
LET GRNPES = NPES
RECODE GRNPES 1-4=1 5-8=2 9-HIGH=3
TABLES GRNPES CACO
anote os resultados
tecle F9 para ir ao DOS
digite STATCALC para cálculo do OR não emparelhado (selecione
Tables 2 x 2, 2 x n)
utilize os dados produzidos pela tabela anterior
pressione F10 para sair do STATCALC
digite EXIT para voltar ao ANALYSIS
CLOSE (feche o arquivo)
```

**Questão 3.** Quais as possíveis explicações para as diferenças dos OR obtidos para "captura de triatomíneo" (DCAP) e "moradia infestada" (BARB)? Utilize o IC 95% e a descrição das variáveis para responder.

**Questão 4.** Qual a relação entre número de pessoas na família e risco de infecção pelo *T.cruzi* em crianças? Relacione a significância estatística ao significado clínico.

**Questão 5.** A associação entre soropositividade da mãe e filho soropositivo pode ter ocorrido ao acaso? Quais as prováveis interpretações para esta associação? Qual a interpretação do resultado do OR da sorologia do pai e soropositividade do filho?

**2. Fatores de risco para infecção /infestação chagásica** - o arquivo **casach.rec** extraído da base de dados **casoch.rec** inclui somente uma criança (soropositiva-caso ou soronegativa-controle) por moradia, excluindo a possibilidade de casos e controles coabitando a mesma moradia. A unidade de estudo passa a ser a "moradia de criança soropositiva"/caso (n=89) e "moradia de criança soronegativa"/controle (n=178) e a variável caso/controle é CASAPOS. Detalhes da metodologia estão na referência **Andrade et al, 1995**.

**Questão 6.** Qual a relação entre aumento do número de cômodos e risco de infecção pelo *T.cruzi* em crianças? Construa categorias para a variável "número de cômodos" (NCOM), criando uma nova variável "GRNCOM". Sugestão: 1-3 cômodos, 4-6 e  $\geq 7$  cômodos. Escolha a categoria 1-3 como referência e calcule o OR não emparelhado para as demais categorias comparando com a categoria de referência.

**Notas 6:**

READ CASACH.REC (para abrir o novo arquivo)  
Para criar categorias para a variável "número de cômodos" ("GRNCOM") repita os comandos utilizados para categorias do "número de pessoas na família" (questão 2).  
TABLES GRNCOM CASAPOS  
anote os resultados  
tecle F9 para ir ao DOS  
Utilize o STATCALC para o cálculo do OR não emparelhado; siga os mesmos comandos da questão 2

**Questão 7.** Calcule a soroprevalência para as mães e para os pais no grupo controle. Que inferências podem ser feitas para a região de estudo em relação à endemicidade da infecção? Que implicações têm estes resultados do ponto de vista da possibilidade de transmissão congênita na área?



```
Notas 7:      SET PERCENTS=ON
              SELECT SORMAE <> 9
              TABLES SORMAE CASAPOS
              SELECT
              SELECT SORPAI <> 9
              TABLES SORPAI CASAPOS
              SELECT
              SET PERCENTS=OFF
```

**Questão 8.** Calcule para cada base de dados (*casoch.rec* e *casach.rec*) o poder do estudo de acordo com o aumento do número de controles por caso. Considere  $\alpha=0,05$ , número de casos=149 e 89, OR 2,0 e 2,2 respectivamente e prevalência da variável de exposição=15%. Qual a relação entre número de controles por caso e o poder do estudo? Discuta o custo-benefício e aspectos operacionais referentes a seleção de mais de um controle por caso, levando em conta a estratégia para recrutamento de casos e controles adotada nesta investigação. Qual a relação entre o número de indivíduos (tamanho do estudo) e o poder do estudo? (responda comparando o número de participantes dos dois bancos de dados).

```
Notas 8:      tecla F9 para ir ao DOS
              digite EPITABLE para cálculo do poder estatístico (comando
              SAMPLE / POWER CALCULATION / CASE-CONTROL
              STUDY)
              pressione F10 para sair do EPITABLE
              digite EXIT para voltar ao ANALYSIS
              CLOSE (para fechar o arquivo)
```

**3. Efetividade da vacina BCG na Hanseníase** - um estudo de caso-controle foi delineado para avaliar a efetividade da vacina BCG na hanseníase. Detalhes do estudo estão na referência **Rodrigues et al.,1992**. 62 crianças/escolares menores de 16 anos recém-diagnosticadas com hanseníase pelo ambulatório de referência da cidade, no período de 1989-1990, foram incluídas como **Casos**. **Controles** (3:1) foram selecionados de escolas localizadas na mesma área geográfica de procedência dos casos. Casos e Controles foram emparelhados por frequência de sexo e grupo etário. A exposição ao BCG foi avaliada pela presença de cicatriz vacinal de BCG no antebraço direito. A análise de dados foi planejada tendo como objetivos: (1) avaliar a comparabilidade entre casos e controles em relação às suas características básicas; (2) mensurar a proteção (efetividade) do BCG em relação a hanseníase e (3) determinar a associação entre BCG e formas clínicas de hanseníase. Utilize o arquivo **casohan.rec** para responder as questões abaixo.

**Questão 9.** Faça uma tabela incluindo as características gerais dos casos e controles - sexo e idade. O que pode ser concluído em relação à comparabilidade entre os grupos?

**Notas 9:**

```
READ CASOHAN.REC (para abrir o arquivo)
criar a variável GRAGE (grupo etário)
LET GRAGE = AGE
RECODE GRAGE low- 5=1 6-10=2 11-15=3
SET PERCENTS=ON
TABLES GRAGE TIPO
TABLES SEX TIPO
SET PERCENTS=OFF
MEANS AGE TIPO /N
```

**Questão 10.** Analise a média de idades entre indivíduos vacinados e não vacinados com BCG. Quais as prováveis interpretações destes resultados? Responda considerando que a vacina BCG foi implementada pelo Programa Nacional de Imunização a partir de 1975.

**Notas 10:** MEANS AGE CIC /N

**Questão 11.** Calcule o *odds ratio* e IC 95% da vacina BCG. Qual a efetividade da vacina (IC 95%) na proteção da hanseníase?

**Notas 11:**

```
TABLES CIC TIPO
anote a tabela produzida
tecle F9 para ir ao DOS
digite EPITABLE para cálculo da eficácia (comando
STUDY /VACCINE EFFICACY / CASE-CONTROL
STUDY)
utilize os dados da tabela anterior
pressione F10 para sair do EPITABLE
digite EXIT para voltar ao ANALYSIS
```

**Questão 12:** Avalie a associação entre forma clínica de hanseníase e exposição à vacina BCG. Esses resultados poderiam ser explicados pelo acaso? O que você conclui sobre a implementação do BCG como uma das estratégias de Eliminação da Hanseníase? Responda levando em conta a cobertura atual da vacina BCG.

**Notas 12:**

```
SELECT TIPO=1 (para selecionar apenas os casos)
SET PERCENTS=ON
TABLES CIC FC2
SELECT
SET PERCENTS=OFF
CLOSE (para fechar o arquivo)
```

## REFERÊNCIAS BIBLIOGRÁFICAS

ANDRADE, A.L.S.S., ZICKER, F., SILVA, I.G., SOUZA, J.M.P. & MARTELLI, C.M.T. Risk factors for *Trypanosoma cruzi* infection among children in Central Brazil: a case-control study in vector control settings. **American Journal of Tropical Medicine and Hygiene**, **52(2)**:183-187, 1995.

RODRIGUES, M.L.O., SILVA, S.A., NETO, J.C.A., ANDRADE, A.L.S.S., MARTELLI, C.M.T. & ZICKER, F. Protective effect of intradermal BCG against leprosy: a case-control study in Central Brazil. **International Journal of Leprosy and Other Mycobacterium Disease**, **60(3)**:335-339, 1992.

Arquivo: **casoch.rec**

Nome	Descrição	Código	Descrição do Código
<b>NOID</b>	Número de identificação	1 a 463	
<b>NOPAR</b>	Identificador dos triplos	1 a 154	
<b>CACO</b>	Caso/Controle	1 2	Caso Controle
<b>AGE</b>	Idade em anos completos	7 a 12	
<b>SEX</b>	Sexo	1 2	Masculino Feminino
<b>MUN</b>	Município	1 2 3	Posse Guarani de Goiás Simolândia
<b>NCOM</b>	Número de cômodos na moradia atual	1 a 13	
<b>NPES</b>	Número de moradores	2 a 14	
<b>CP</b>	Casa própria	1 2	Sim Não
<b>DCAP</b>	Captura triatomíneo no domicílio	1 2	Sim Não
<b>DVTR</b>	Vestígios triatomíneo no domicílio	1 2	Sim Não
<b>MORB</b>	Antecedentes morbidos	1 2 9	Sim Não Sem informação
<b>HOSP</b>	Antecedentes de hospitalização	1 2 9	Sim Não Sem informação
<b>SORPAI</b>	Sorologia do pai	1 2 9	Positivo Negativo Não realizada
<b>SORMAE</b>	Sorologia da mãe	1 2 9	Positivo Negativo Não realizada

Arquivo: **casach.rec**

Variável	Descrição	Código	Descrição do Código
<b>NOID</b>	Número de identificação	1 a 463	
<b>CASAPOS</b>	“Caso”/ “Controle”	1 2	Casa caso Casa controle
<b>NCOM</b>	Número de cômodos na moradia atual	1 a 13	
<b>SORPAI</b>	Sorologia do Pai	1 2 9	Positivo Negativo Não realizada
<b>SORMAE</b>	Sorologia da Mãe	1 2 9	Positivo Negativo Não realizada
<b>NOPAR2</b>	Identificador dos triplos	1 a 89	

Arquivo: **casohan.rec**

Variável	Descrição	Código	Descrição do Código
<b>NO</b>	Número de identificação	1 a 301	
<b>AGE</b>	Idade em anos completos	2 a 15	
<b>SEX</b>	Sexo	1 2	Masculino Feminino
<b>CIC</b>	Presença de cicatriz vacinal - BCG	1 2	Sim Não
<b>TIPO</b>	Identificação caso/controle	1 2	Caso Controle
<b>FC2</b>	Forma Clínica	P M	Paucibacilar Multibacilar