



Universidade de São Paulo
Faculdade de Saúde Pública
Departamento de Epidemiologia

Validade em Estudos Epidemiológicos - Viés

ALESSANDRA C GOULART
ANA PAULA SAYURI SATO

O que o Epidemiologista mede?

I - Medidas de Ocorrência

- Proporção de incidência
- Taxas
- Prevalências

II - Medidas de Associação

- Risco Relativo
- Odds ratio
- Hazard ratio
- Razão de Prevalência

.... Ainda que sejam somente *estimativas* do
“verdadeiro” valor.....

Validade

Delineamento → condução do estudo → análise

- Produto final: ESTIMATIVA EPIDEMIOLÓGICA

Acurácia (pouco erro)

Validade da estimação

- Pesquisa epidemiológica → “Exercício de mensuração”
- Estimativa que seja generalizável à populações-alvo relevantes



Devo acreditar na mensuração?

Zika



Microcefalia

OR = 73,1

**Associação Estatística
(verdadeira)**

causal

não causal

Ao acaso?

Confusão?

Viés?

Possíveis explicações para uma associação entre uma exposição e uma doença

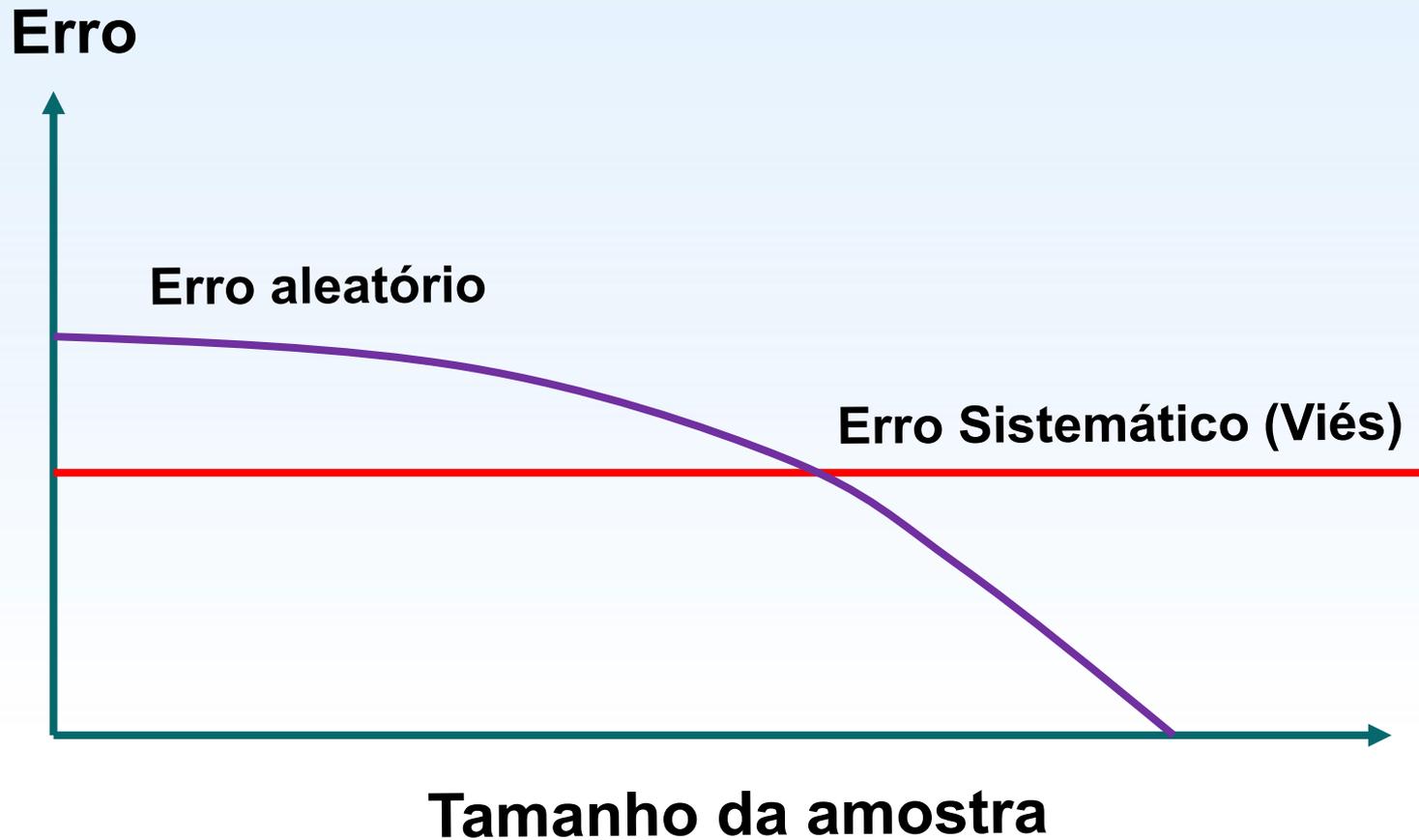
1. Causal \implies Critérios de causalidade

2. Variação aleatória \implies Inferência Estatística

3. Viés \implies Julgamento epidemiológico

4. Fator de confusão \implies Julgamento epidemiológico

Erros em Estudos Epidemiológicos



Validade - Definição

- Ausência de viés
- Se o desenho e procedimentos do estudo não apresentarem vieses, o estudo é considerado válido, pois os resultados tenderão a estar corretos.

Validade Interna

- Até que ponto estão corretas as conclusões do investigador sobre o que realmente ocorreu no estudo

Validade Interna

Validade Externa

- Capacidade de generalização
- O quanto essas conclusões se aplicam a pessoas e eventos externos ao estudo

Validade Externa

Formulando
conclusões

Verdade
no
universo

←
Inferência

Verdade
no
estudo

←
Inferência

Achados
no
estudo

Delineando e
implementando

Pergunta
da
pesquisa

→
Delinea
mento

Plano de
pesquisa

→
Implementa
ção

Pesquisa

Validade
Externa

Validade
interna

Viés - Definição

- “ Erro sistemático no delineamento, condução ou análise de um estudo em que os resultados apresentam uma estimativa equivocada do efeito de uma exposição no risco para uma doença” (Gordis, 2009)
- “Erro sistemático no delineamento ou condução do estudo”
 - Resultados observados tendem a ser diferentes dos verdadeiros resultados (Szklo, 2019)
 - Processo

Viés – Definição Epidemiológica

Porta M: A Dictionary of epidemiology, ed., 6th Edition, IEA 2014

“Desvio dos resultados ou inferências da verdade, ou processos que levam a tal desvio. Qualquer tendência na coleta, análise, interpretação, publicação ou revisão de dados que possa levar à conclusões que são sistematicamente diferentes da verdade”

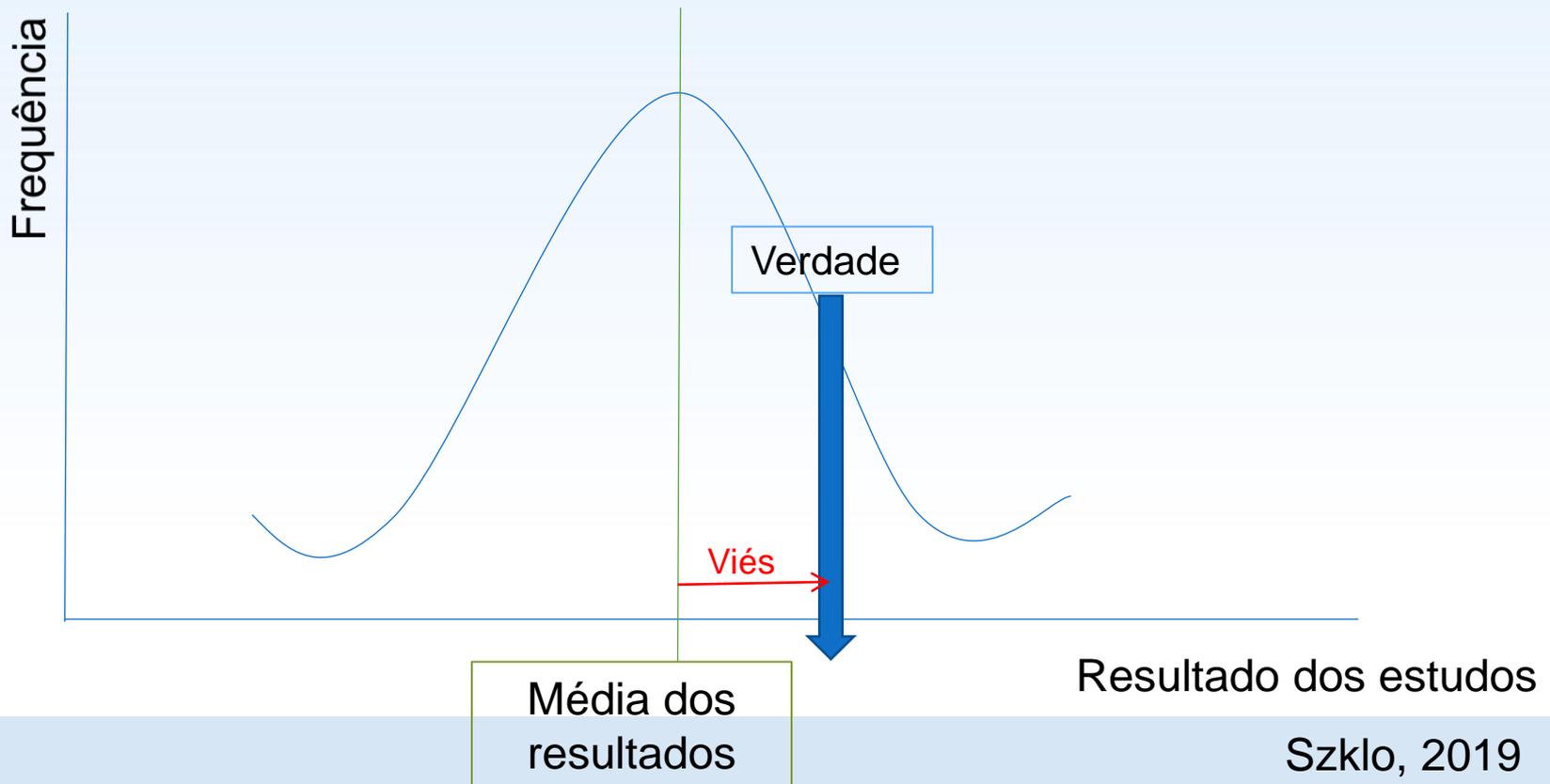
Viés - Definição

- **Processo:** Desenho, métodos e procedimentos
 - Desenho de estudo é apropriado para responder às hipóteses do estudo
 - Procedimentos padronizados de monitoramento da coleta de dados → Validade e reprodutibilidade
 - Procedimentos de análise de dados apropriados

Viés – Definição Estatística

Ocorre quando a média de resultados obtida de um número infinito de estudos não é o verdadeiro valor

Distribuição de número infinito de estudos



VIES

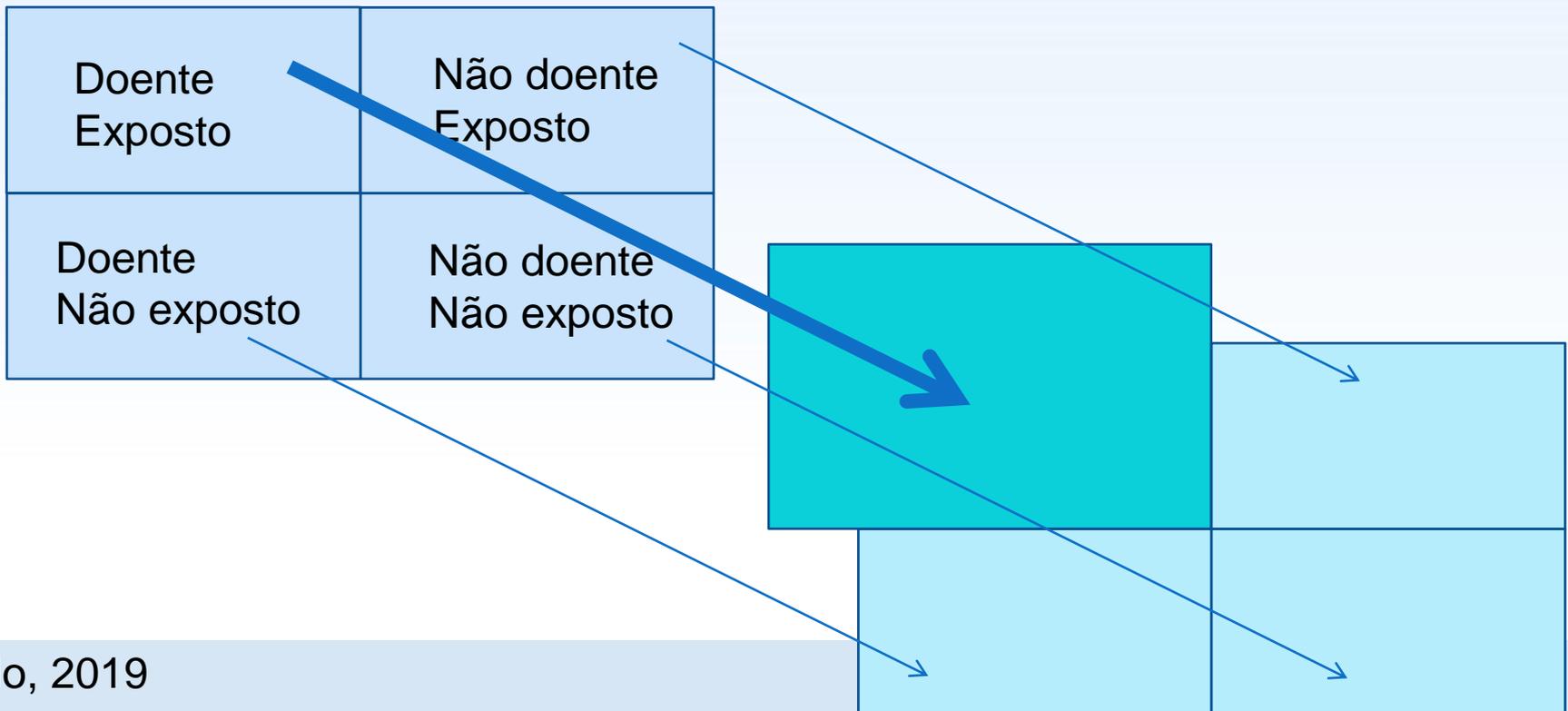
1. Seleção: está relacionado ao delineamento do estudo, quando a probabilidade de seleção é dependente do *status* de exposição (estudos caso-control) ou do desfecho (estudos de coorte)

2. Informação: pode ocorrer, por exemplo, por erros no recordatório ou do entrevistador, resultando na classificação incorreta da exposição ou do “efeito”

Tipos de Viés

- **Viés de seleção**

“Quando os indivíduos tem probabilidades diferentes de serem incluídos na amostra do estudo, de acordo com suas exposições ou desfechos”



Um Estudo Caso-Controle de Alcoolismo e Pneumonia

Casos e Controles selecionados de pacientes hospitalizados

Alcoólatras com pneumonia têm maior probabilidade de serem admitidos do que não alcoólatras com pneumonia

O risco de pneumonia associada ao alcoolismo apresenta em viés com superestimação

Pneumonia e Alcoolismo - Populacional

		Pneumonia	
		<i>Sim</i>	<i>Não</i>
Alcoolismo	<i>Sim</i>	10	10
	<i>Não</i>	90	90
		100	100

$OR = \frac{10 / 90}{10 / 90}$

$OR = 1.0$

Pneumonia e Alcoolismo em Hospital

		Pneumonia	
		<i>Sim</i>	<i>Não</i>
Alcoolismo	<i>Sim</i>	20 ↑	10
	<i>Não</i>	80	90
		100	100

$OR = \frac{20 / 80}{10 / 90}$

$OR = 2,25$

Quão representativos são os casos de pneumonia hospitalizados dos casos ocorridos na comunidade?

Viés de seleção

- **Viés de Berkson (1946):**

- Tanto a exposição quanto a doença afetam a seleção
- Estudos caso-controle, hospitalizados
- Associação espúria encontrada, que resulta de taxas de admissão no hospital diferentes entre as pessoas com a doença, sem a doença e com a característica em estudo.
- Ex: Estrogênio exógeno e câncer de endométrio

Table 1. Association in the general population as reported in Berkson²

	Exposed	Unexposed	Total
Cases	3000	97 000	100 000
Non-cases	297 000	9 603 000	9 900 000
Total	300 000	9 700 000	10 000 000

Odds ratio = 1.0.

Table 2. Association using hospitalized cases and general population controls

	Exposed	Unexposed	Total
Cases	590	5311	5901
Controls	29 700	960 300	990 000
Total	30 290	965 611	995 901

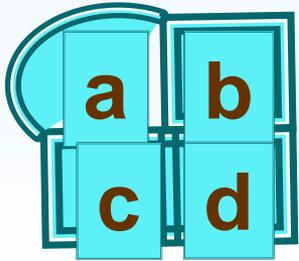
Odds ratio = 3.59.

Viés de Diagnóstico

Viés de Diagnóstico por se conhecer o status de exposição

Exemplo: Contraceptivo oral e câncer uterino

Contraceptivo Oral	Casos câncer uterino	Controles
SIM	a 	b
NÃO	c	d



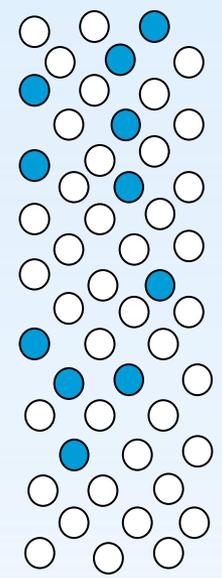
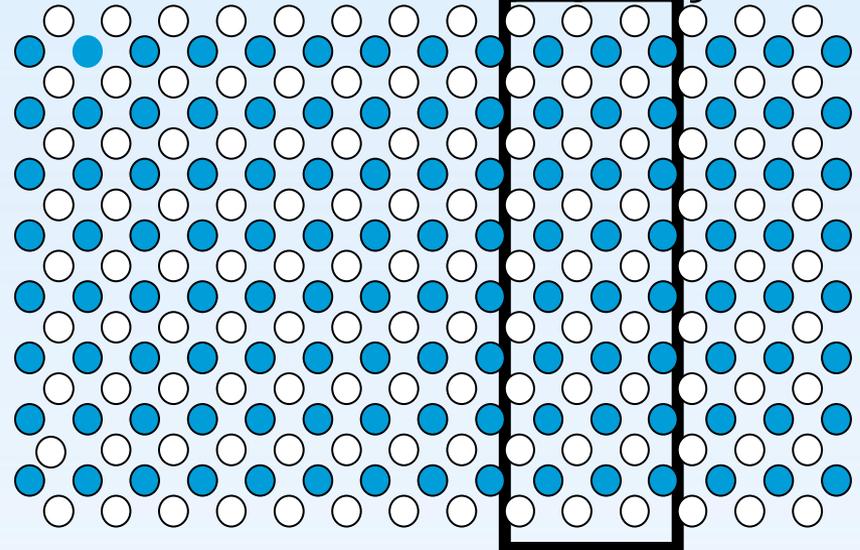
Uso de Contraceptivo oral \Rightarrow Aumenta a probabilidade de diagnóstico de câncer uterino. Portanto, aumenta a proporção de expostos entre os casos e superestima o OR

Viés de Seleção

1) Estudo hipotético tipo caso-controle incluindo todos os casos e não casos de uma população definida. Assumindo que não exista erro de classificação, nem fator de confusão ou erro aleatório (Será o “padrão ouro” para os exemplos a seguir)

Fator de Risco A	População Total	
	Casos	Não casos (controles)
Presente	50	180
Ausente	50	720
Total	100	900
Odds da Exposição	50:50=1:1	180:720=1:4
Odds Ratio	$(50/50)/(180/720)=4,0$	
Conclusões	Odds de exposição sem viés em casos e controles Odds Ratio sem viés	

Fonte da população

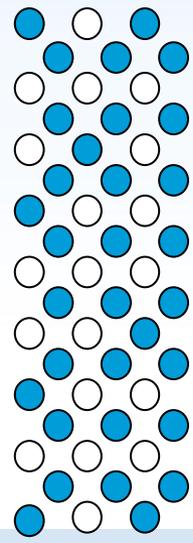
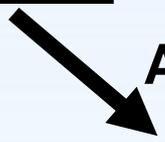


Odds de Exposição entre os casos

Casos

- Expostos
- Não expostos

Amostra



Odds Exposição entre os controles

Controles

Viés de Seleção

- 2) Um hipotético estudo tipo caso-controle incluindo uma amostra sem viés da população de referência, com 50% dos casos e 10% dos controles

Fator de Risco A	População Total	
	Casos	Não casos (controles)
Presente	25	18
Ausente	25	72
Total	50	90
Odds da Exposição	25:25=1:1	18:72=1:4
Odds Ratio	$(5/5)/(18/72)=4,0$	
Conclusões		

Viés de Seleção

3. Um exemplo de viés de seleção na seleção dos casos num hipotético estudo tipo caso-controle incluindo 50% da amostra de casos e 10% da amostra de não casos originários da população de referência

Fator de Risco A	População Total	
	Casos	Não casos (controles)
Presente	$50 \times 0,60 = 30$	$180 \times 0,10 = 18$
Ausente	$50 \times 0,40 = 20$	$720 \times 0,10 = 72$
Total	$100 \times 0,50 = 50$	$900 \times 0,10 = 90$
Odds da Exposição	$30:20=1,5:1$	$18:72=1:4$
Odds Ratio	$(30/20)/(18/72)=6,0$	
Conclusões		

Viés de seleção

Viés de auto seleção:

- **Efeito do trabalhador saudável**
- Pessoas empregadas são pessoas relativamente saudáveis, enquanto pessoas desempregadas, aposentadas, incapacitadas, ou que por qualquer outro motivo estejam fora do mercado de trabalho constituem um grupo menos saudável;
- Provavelmente menos pessoas empregadas se candidatariam para um estudo.

Efeito de Trabalhadores Saudáveis

	Trabalhadores expostos	População geral
Óbitos	50	7.000
Pessoas- tempo (em anos)	1.000	100.000
Mortalidade (casos/ano)	0,05	0,07

$$RR = 0,05 / 0,07 = 0,7$$

Efeito de Trabalhadores Saudáveis

	Trabalhadores expostos	População Geral		
		Trabalhadores	Não Trabalhadores	Total
Óbitos	50	4.500	2.500	7.000
Pessoas-tempo	1.000	90.000	10.000	100.000
Mortalidade (casos/ano)	0,05	0,05	0,25	0,07

Viés de seleção

Viés de auto seleção:

- **Viés de publicidade**

- Pessoas que se autorreferem para pesquisadores,
- após um anúncio publicitário de um estudo;
- Autorreferência de indivíduos é considerada uma ameaça à validade, pois as razões para a autorreferência podem estar associadas com o desfecho em estudo;

Viés de seleção

- **Viés de voluntário:**

- Voluntários apresentam características diferentes dos que se recusam a participar.
- Voluntários que aderem ao tratamento são diferentes dos que não aderem.
- A gravidade da doença pode variar nos voluntários que aparecem no início ou final do estudo.

Viés de seleção

Viés de perda de seguimento

- Quando há uma perda diferencial de seguimento que está relacionada com a exposição;
- O planejamento e a implementação do estudo deveriam tentar minimizar este tipo de perda, e deveria ser tentado um seguimento completo e igualmente rigoroso pra todos os grupos;
- o cegamento único ou duplo é recurso para se garantir que todos os grupos serão igualmente seguidos.

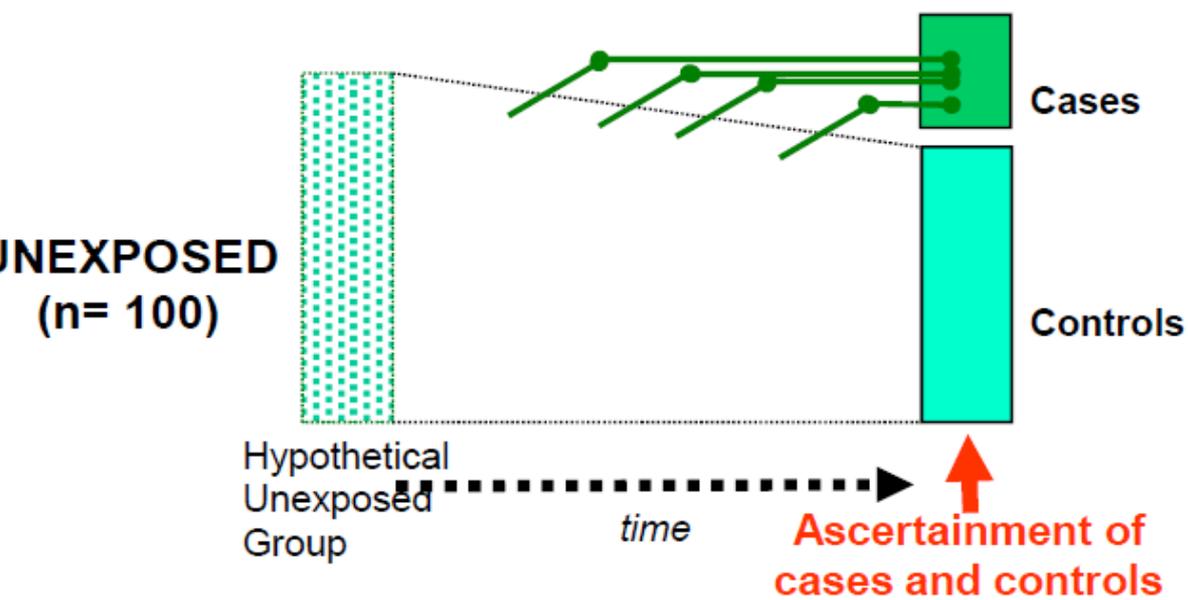
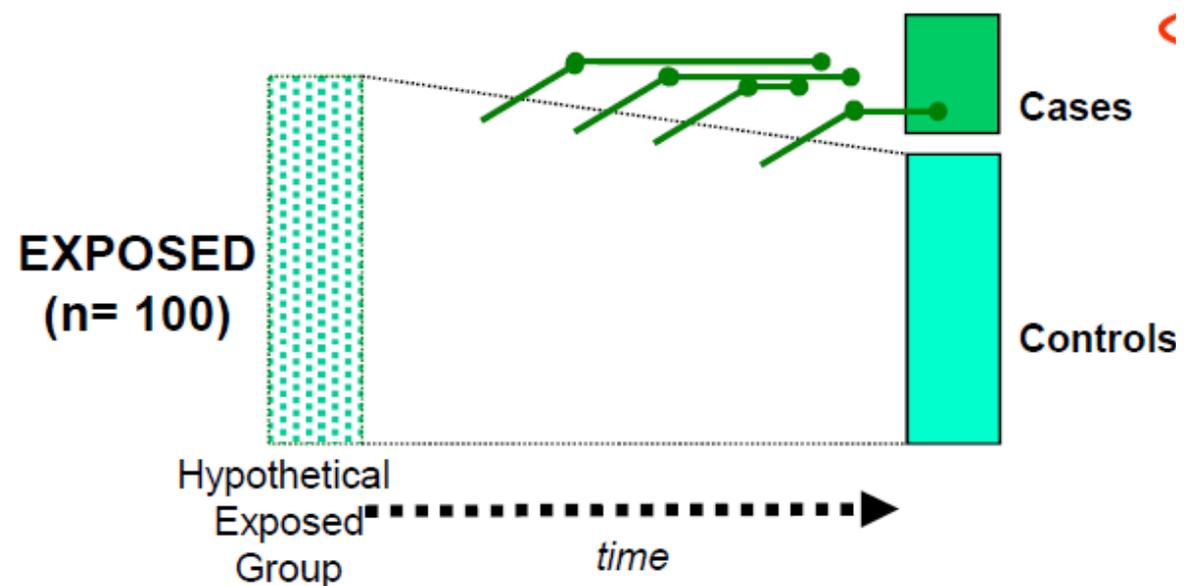
Viés de seleção

- **viés de sobrevida relativa/**

incidência-prevalência (viés de Neyman):

- Dependendo do tempo após a exposição que as pessoas são investigadas, pode-se perder as que já morreram ou que se recuperaram.
- Casos prevalentes são usados para se estudar a associação exposição-doença e está relacionado a dois fenômenos:
 - Uma vez que a pessoa recebe um diagnóstico, ela pode trocar o hábito que está associado à doença;
 - Casos prevalentes representam sobreviventes da doença em questão, e como sobreviventes podem ser atípicos com relação à situação de exposição, que pode estar enviesada.

ASSUME NO LOSSES IN EXPOSED AND UNEXPOSED



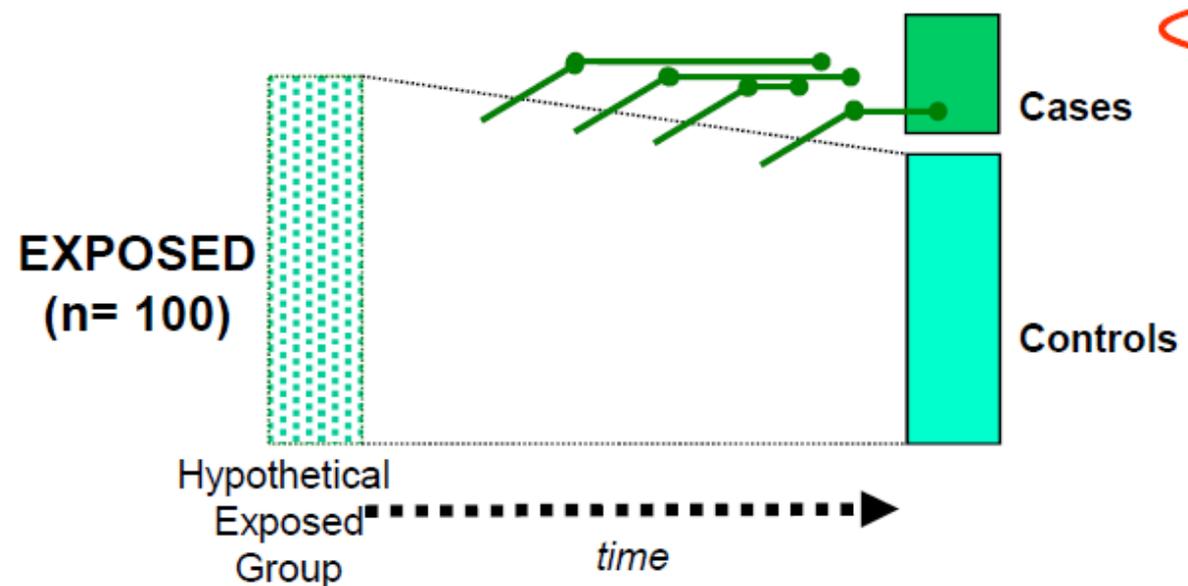
CROSS-SECTIONAL CASE-CONTROL STUDY INCLUDING ONLY POINT PREVALENT CASES, BUT ALL NON-CASES

Exposed?	Cases	Controls*
Yes	1	96
No	4	96
	5	192

$$OR = \frac{1}{4} \div \frac{96}{96} = 0.25$$

SELECTION/SURVIVAL BIAS (ALSO KNOWN AS PREVALENCE-INCIDENCE BIAS)

ASSUME NO LOSSES IN EXPOSED AND UNEXPOSED

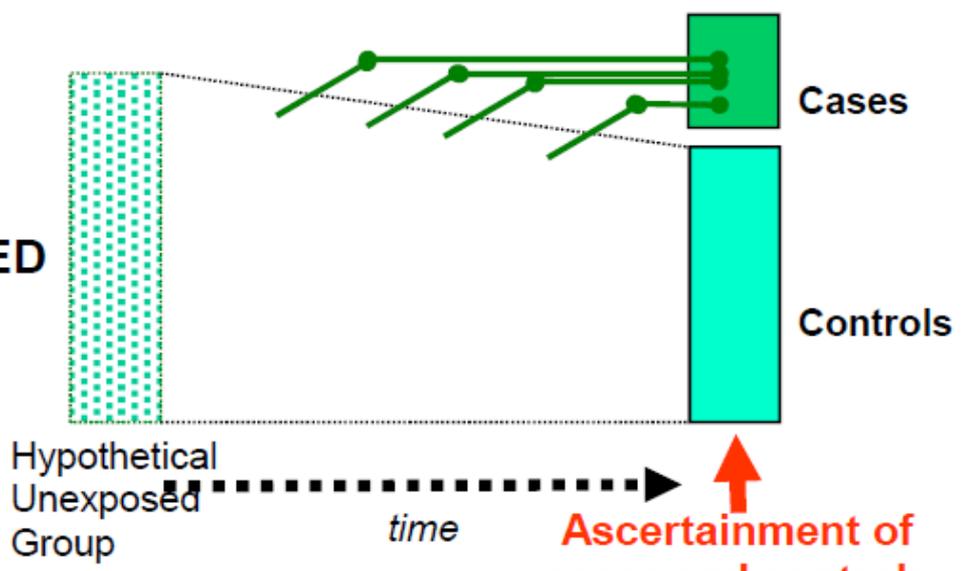


CASE-CONTROL STUDY INCLUDING ALL INCIDENT CASES AND NON-CASES*

Exposed?	Cases	Controls*
Yes	4	96
No	4	96
	8	192

$$OR = \frac{4/4}{96/96} = 1.0$$

UNEXPOSED (n= 100)



CROSS-SECTIONAL CASE-CONTROL STUDY INCLUDING ONLY POINT PREVALENT CASES, BUT ALL NON-CASES

Exposed?	Cases	Controls*
Yes	1	96
No	4	96
	5	192

$$OR = \frac{1/4}{96/96} = 0.25$$

Ascertainment of cases and controls

SELECTION/SURVIVAL BIAS (ALSO KNOWN AS PREVALENCE-INCIDENCE BIAS)

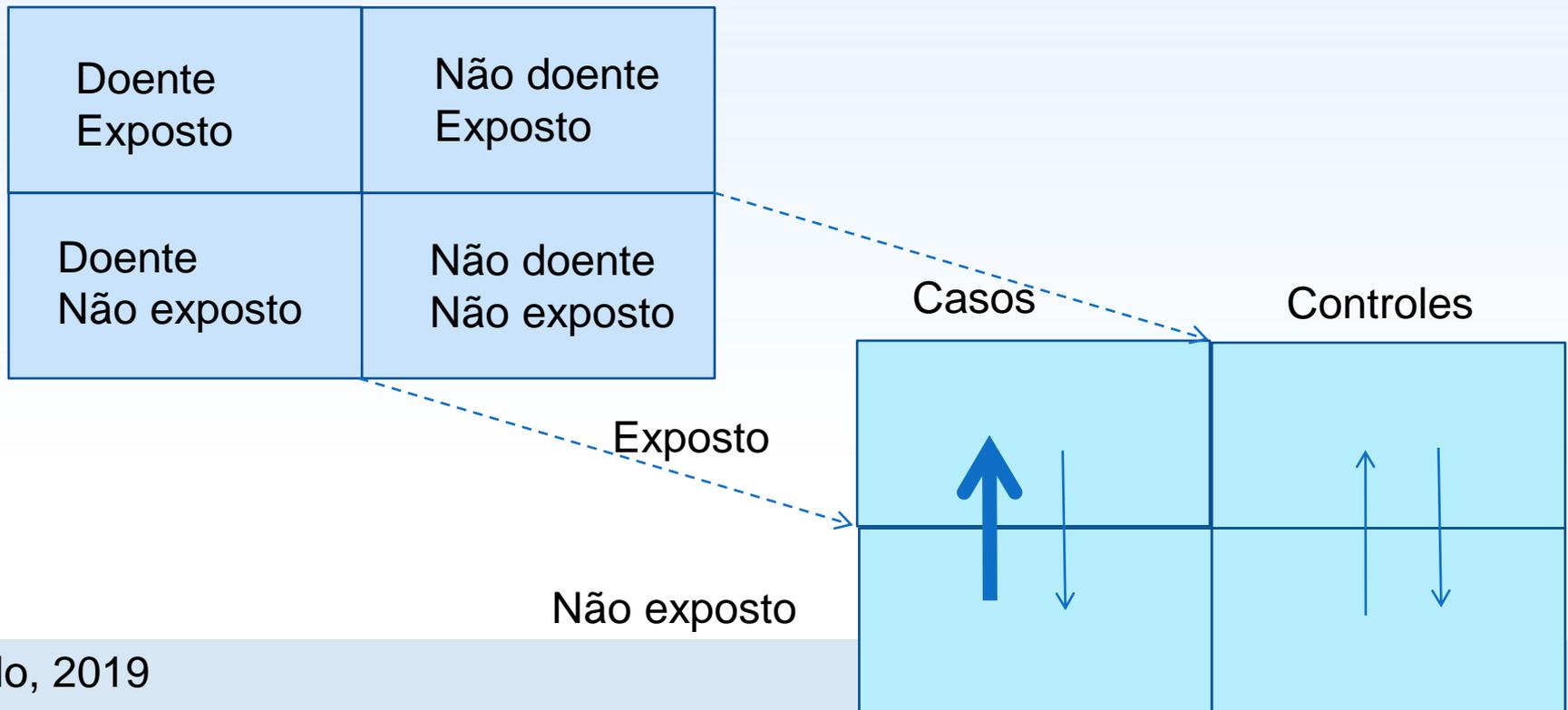
Viés de seleção

- Resultam de distorções dos procedimentos usados para selecionar sujeitos e de fatores que influenciam a participação no estudo.
- A relação entre exposição e doença é diferente para aqueles que participam e para aqueles que, teoricamente, seriam elegíveis para o estudo.
 - **Diferenças seletivas entre grupos de comparação que impactam a associação entre exposição e desfecho**
- **Usualmente resulta de grupos de comparação que não provêm da mesma população base do estudo,**
 - **E portanto não são representativos**

Tipos de Viés

- **Viés de informação**

“Resulta da tendência sistemática dos indivíduos selecionados serem classificados equivocadamente em relação as suas categorias de exposição/desfecho”



Viés de Informação

Conceito

- Viés de Informação (erro de classificação)
- Existe certo grau de erro de classificação com respeito a informações de exposição nos casos e controles
- Tendem- a errar a informação relativa a exposição em maior extensão entre os casos não expostos do que entre os controles

Viés de informação

Erro de classificação

- Não diferencial – aleatórios: classifica de forma equivocada tanto casos como controles, ou seja, a sensibilidade e especificidade não variam segundo status de exposição ou doença.
- Diferencial – sistemáticos: classifica de forma equivocada casos em uma direção e controles na outra

Tipos de Vieses de Informação

Não-Diferencial

Diferencial

- Estudo de coorte: pelo status de exposição
- Estudo caso-controle: pelo status caso/controlado

Viés de Informação

1.Exemplo hipotético do efeito de um erro de classificação **não diferencial** de exposição com 30% de erro de classificação entre casos e controles expostos

Sem erro de classificação			30% de erro de classificação na exposição em casos e controles		
Exposição	Casos	Controles	Exposição	Casos	Controles
Sim	50	20	Sim	50-15 = 35	20-6 = 14
Não	50	80	Não	50+15=65	80+6= 86
OR = (50/50)/ (20/80) = 4,0			OR = (35/65)/(14/86) = 3,3		

-30%

Viés de Informação

2. Exemplo hipotético do efeito de um erro de classificação **diferencial** de exposição com 30% dos casos não expostos e 10% dos controles não expostos classificados como expostos.

Sem erro de classificação			30% de erro de classificação dos casos não expostos e 10% entre controles não expostos como expostos		
Exposição	Casos	Controles	Exposição	Casos	Controles
Sim	50	20	Sim	50+15 = 65	20+8 = 28
Não	50	80	Não	50-15=35	80-8= 72
OR = (50/50) / (20/80) = 4,0			OR = (65/35) / (28/72) = 4,8		

-10%
-30%

Viés de Informação

- ✓ **Sensibilidade:** habilidade de *teste* identificar corretamente quem **tem a doença / exposição**
- ✓ **Especificidade:** habilidade de *teste* identificar corretamente quem **não tem a doença / exposição**

Viés de Informação

Resultado do teste

Doente

Sadio

Positivo

Possui a doença e o teste foi positivo
= VERDADEIROS POSITIVOS

VP

Não possuem a doença mas o teste foi positivo
= FALSOS POSITIVOS

FP

Negativo

Possui a doença mas o teste foi negativo
= FALSOS NEGATIVOS

FN

Não possuem a doença e o teste foi negativo
= VERDADEIROS NEGATIVOS

VN

Sensibilidade = $VP / VP + FN$

Especificidade = $VN / VN + FP$

Sensibilidade e Especificidade

	D	ND
T+	a	b
T-	c	d

$$\text{Sensibilidade} = \frac{a}{a + c}$$

$$\text{Especificidade} = \frac{d}{b + d}$$

Viés de Informação

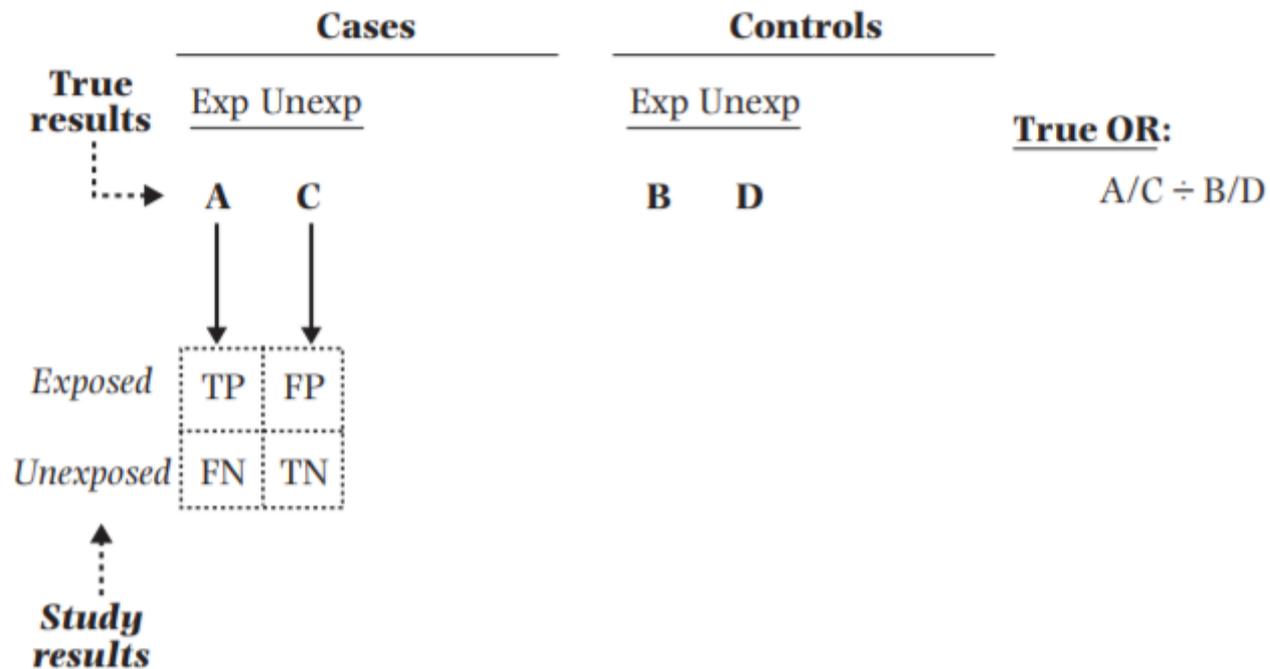
FIGURE 4-4 Application of sensitivity/specificity concepts in misclassification of exposure: schematic representation of true and misclassified relative odds. Sensitivity of exposure ascertainment = $TP \div (TP + FN)$; specificity of exposure ascertainment = $TN \div (TN + FP)$.

	<u>Cases</u>	<u>Controls</u>	
True results ⋮	<u>Exp Unexp</u>	<u>Exp Unexp</u>	True OR: $A/C \div B/D$

Note: Exp: Exposure; TP: True positive; FP: False positive; FN: False negative; TN: True negative

Viés de Informação

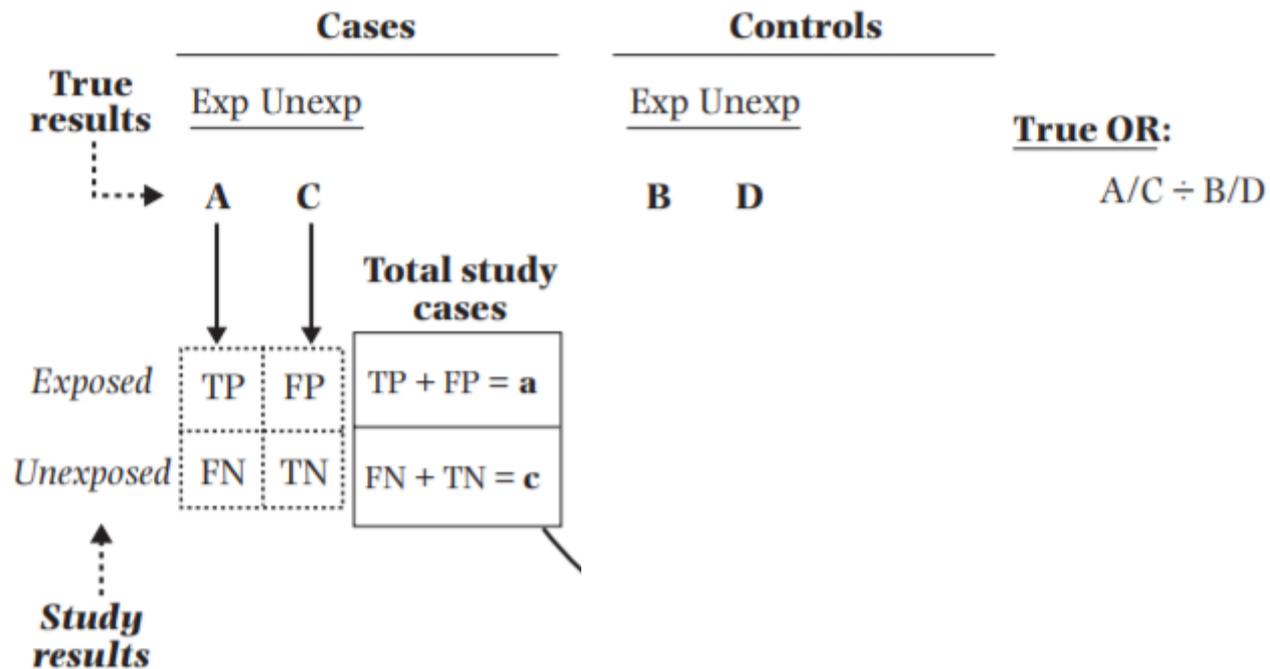
FIGURE 4-4 Application of sensitivity/specificity concepts in misclassification of exposure: schematic representation of true and misclassified relative odds. Sensitivity of exposure ascertainment = $TP \div (TP + FN)$; specificity of exposure ascertainment = $TN \div (TN + FP)$.



Note: Exp: Exposure; TP: True positive; FP: False positive; FN: False negative; TN: True negative

Viés de Informação

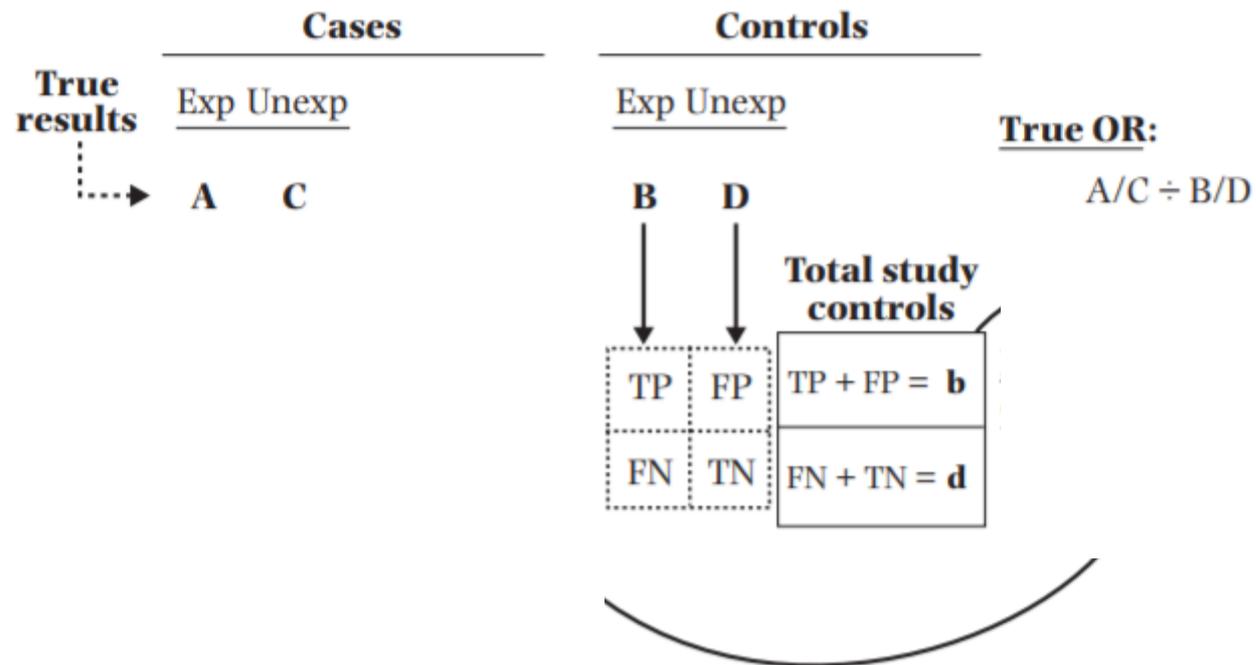
FIGURE 4-4 Application of sensitivity/specificity concepts in misclassification of exposure: schematic representation of true and misclassified relative odds. Sensitivity of exposure ascertainment = $TP \div (TP + FN)$; specificity of exposure ascertainment = $TN \div (TN + FP)$.



Note: Exp: Exposure; TP: True positive; FP: False positive; FN: False negative; TN: True negative

Viés de Informação

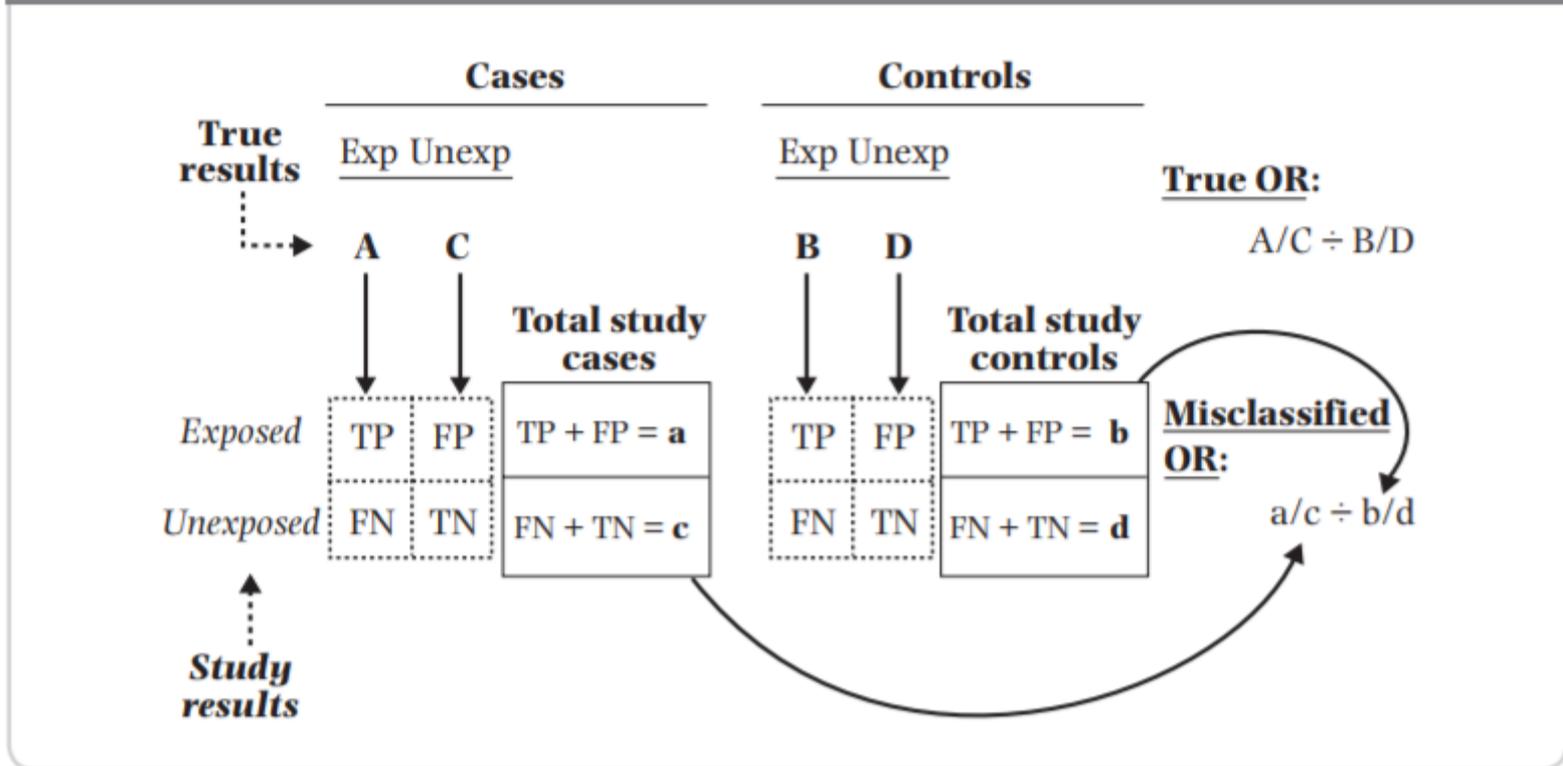
FIGURE 4-4 Application of sensitivity/specificity concepts in misclassification of exposure: schematic representation of true and misclassified relative odds. Sensitivity of exposure ascertainment = $TP \div (TP + FN)$; specificity of exposure ascertainment = $TN \div (TN + FP)$.



Note: Exp: Exposure; TP: True positive; FP: False positive; FN: False negative; TN: True negative

Viés de Informação

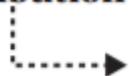
FIGURE 4-4 Application of sensitivity/specificity concepts in misclassification of exposure: schematic representation of true and misclassified relative odds. Sensitivity of exposure ascertainment = $TP \div (TP + FN)$; specificity of exposure ascertainment = $TN \div (TN + FP)$.



Note: Exp: Exposure; TP: True positive; FP: False positive; FN: False negative; TN: True negative

Viés de Informação

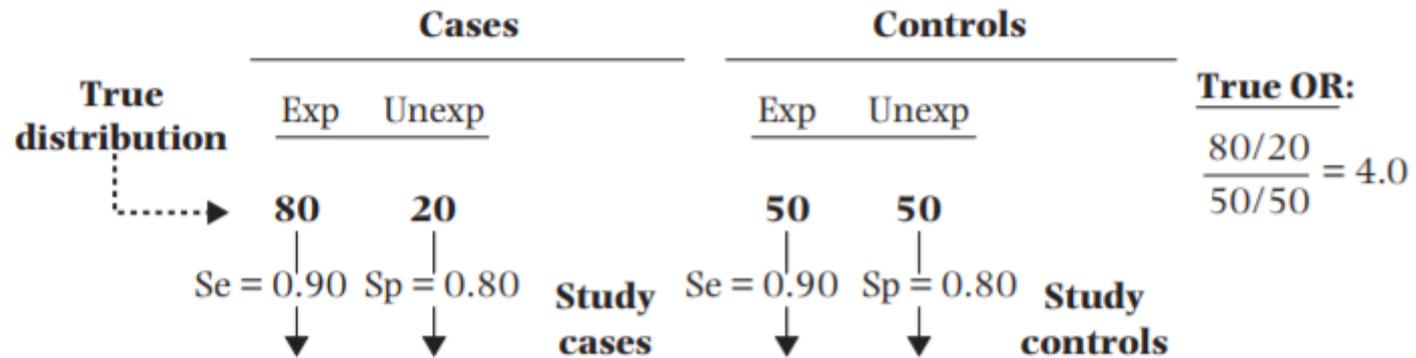
EXHIBIT 4-3 Effects of nondifferential misclassification on the odds ratio (sensitivity = 0.90; specificity = 0.80).

True distribution	Cases		Controls		True OR: $\frac{80/20}{50/50} = 4.0$
	<u>Exp</u>	<u>Unexp</u>	<u>Exp</u>	<u>Unexp</u>	
	80	20	50	50	

Note: Exp: Exposed; Unexp: Unexposed; Se: Sensitivity; Sp: Specificity; OR: Odds ratio

Viés de Informação

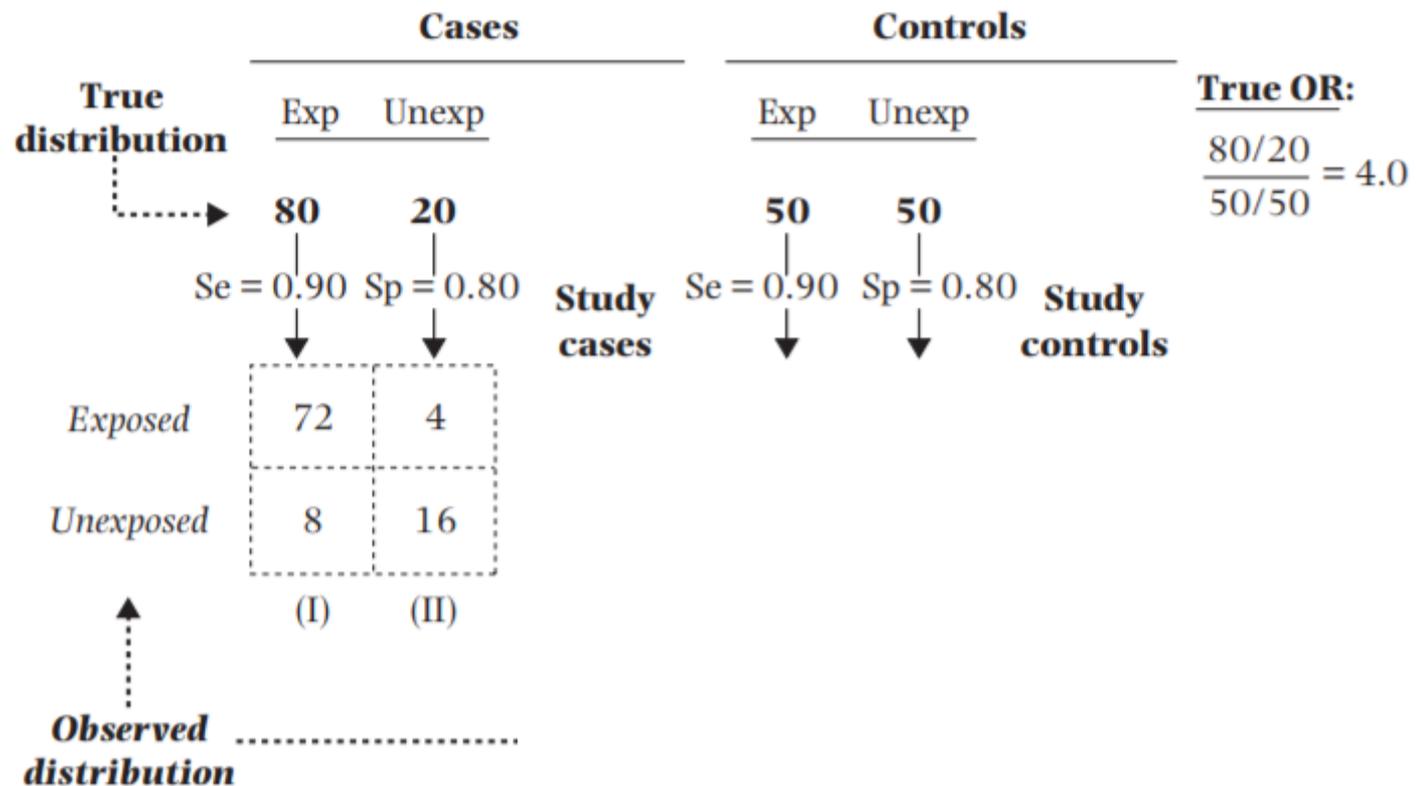
EXHIBIT 4-3 Effects of nondifferential misclassification on the odds ratio (sensitivity = 0.90; specificity = 0.80).



Note: Exp: Exposed; Unexp: Unexposed; Se: Sensitivity; Sp: Specificity; OR: Odds ratio

Viés de Informação

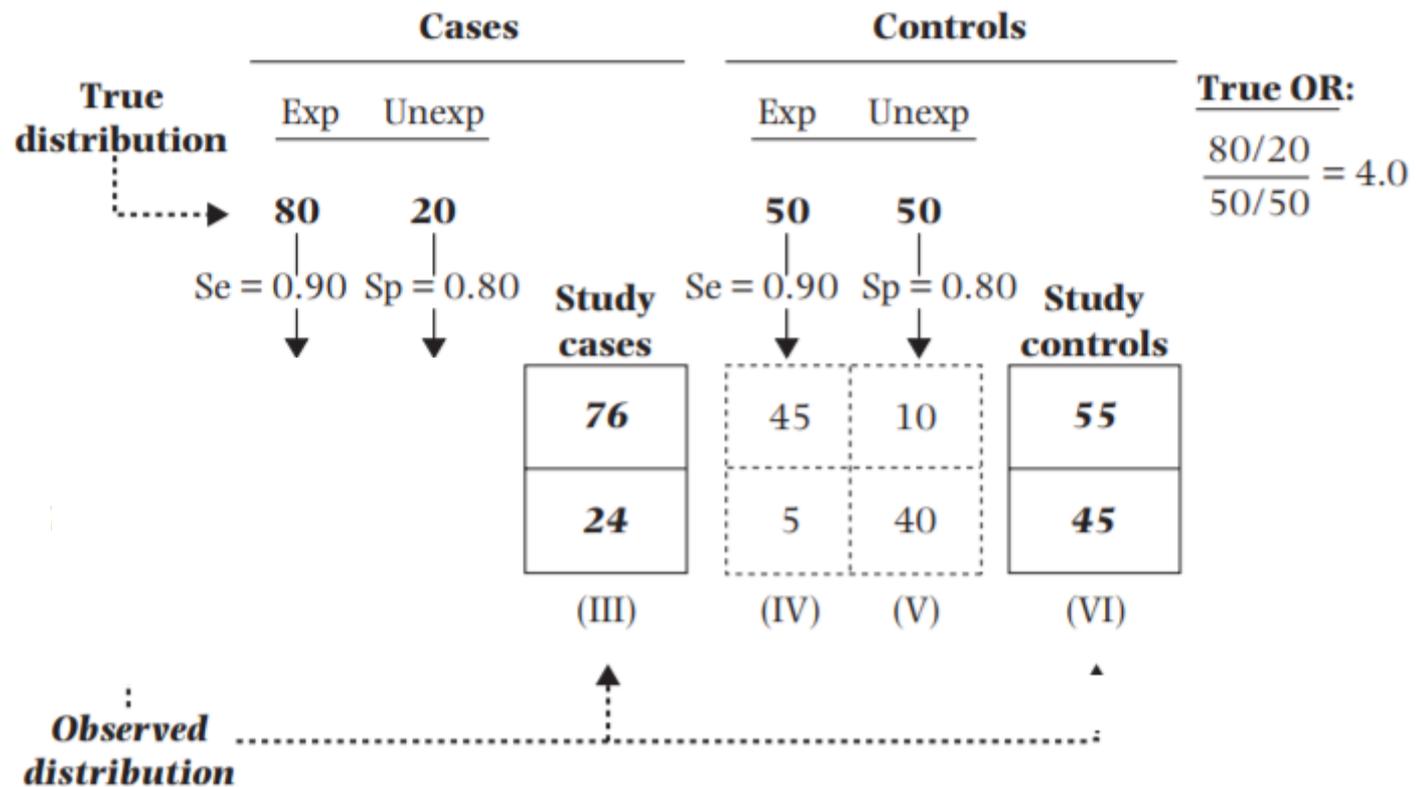
EXHIBIT 4-3 Effects of nondifferential misclassification on the odds ratio (sensitivity = 0.90; specificity = 0.80).



Note: Exp: Exposed; Unexp: Unexposed; Se: Sensitivity; Sp: Specificity; OR: Odds ratio

Viés de Informação

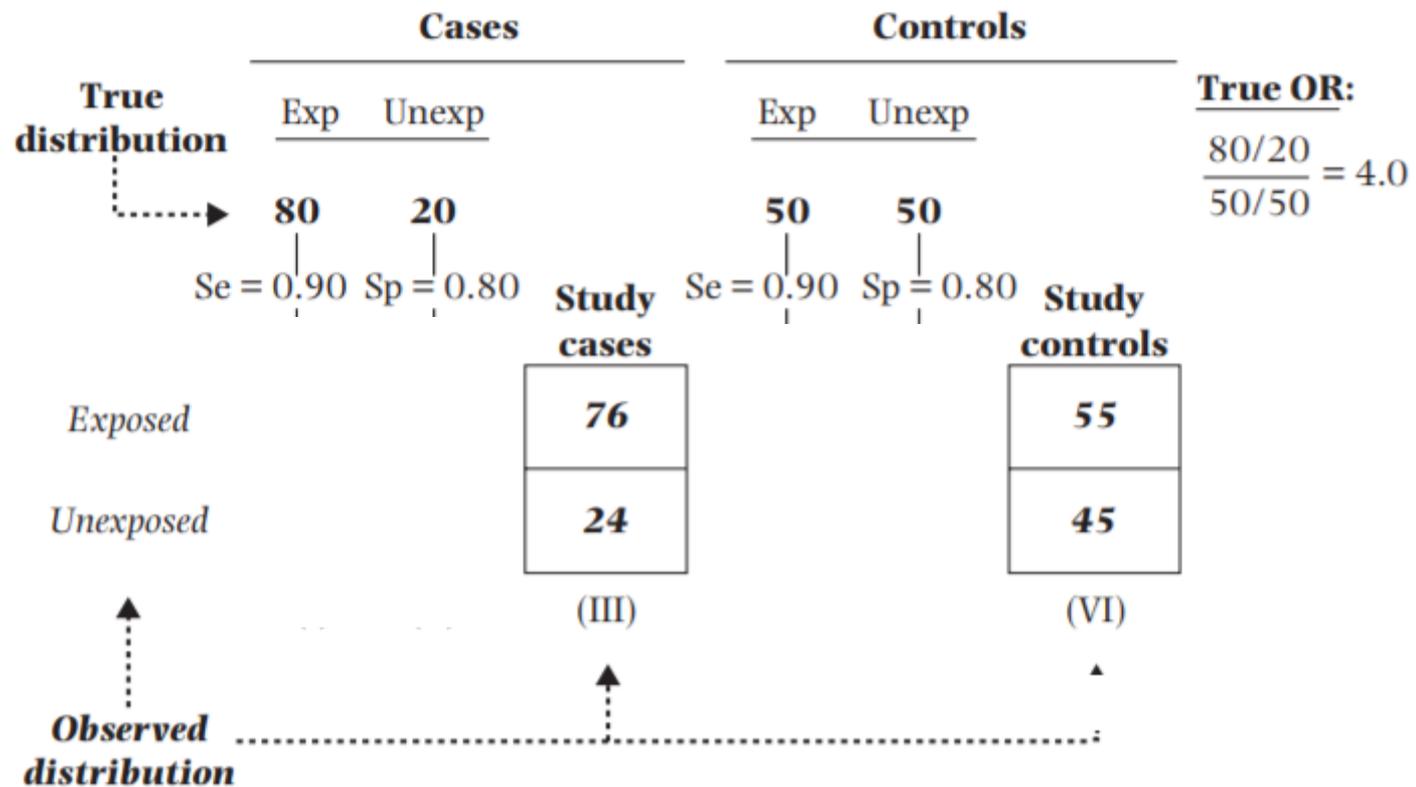
EXHIBIT 4-3 Effects of nondifferential misclassification on the odds ratio (sensitivity = 0.90; specificity = 0.80).



Note: Exp: Exposed; Unexp: Unexposed; Se: Sensitivity; Sp: Specificity; OR: Odds ratio

Viés de Informação

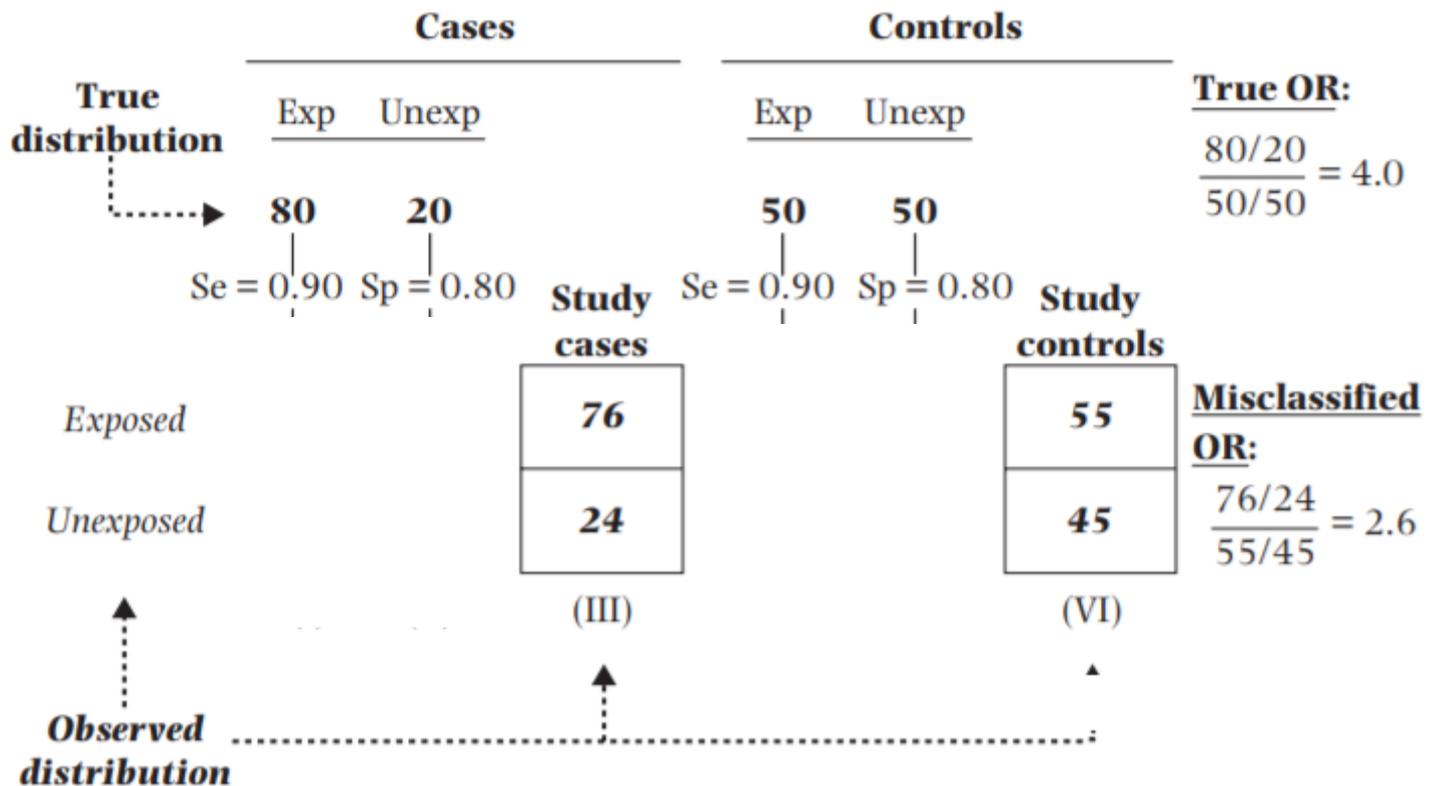
EXHIBIT 4-3 Effects of nondifferential misclassification on the odds ratio (sensitivity = 0.90; specificity = 0.80).



Note: Exp: Exposed; Unexp: Unexposed; Se: Sensitivity; Sp: Specificity; OR: Odds ratio

Viés de Informação

EXHIBIT 4-3 Effects of nondifferential misclassification on the odds ratio (sensitivity = 0.90; specificity = 0.80).



Note: Exp: Exposed; Unexp: Unexposed; Se: Sensitivity; Sp: Specificity; OR: Odds ratio

Viés de Informação

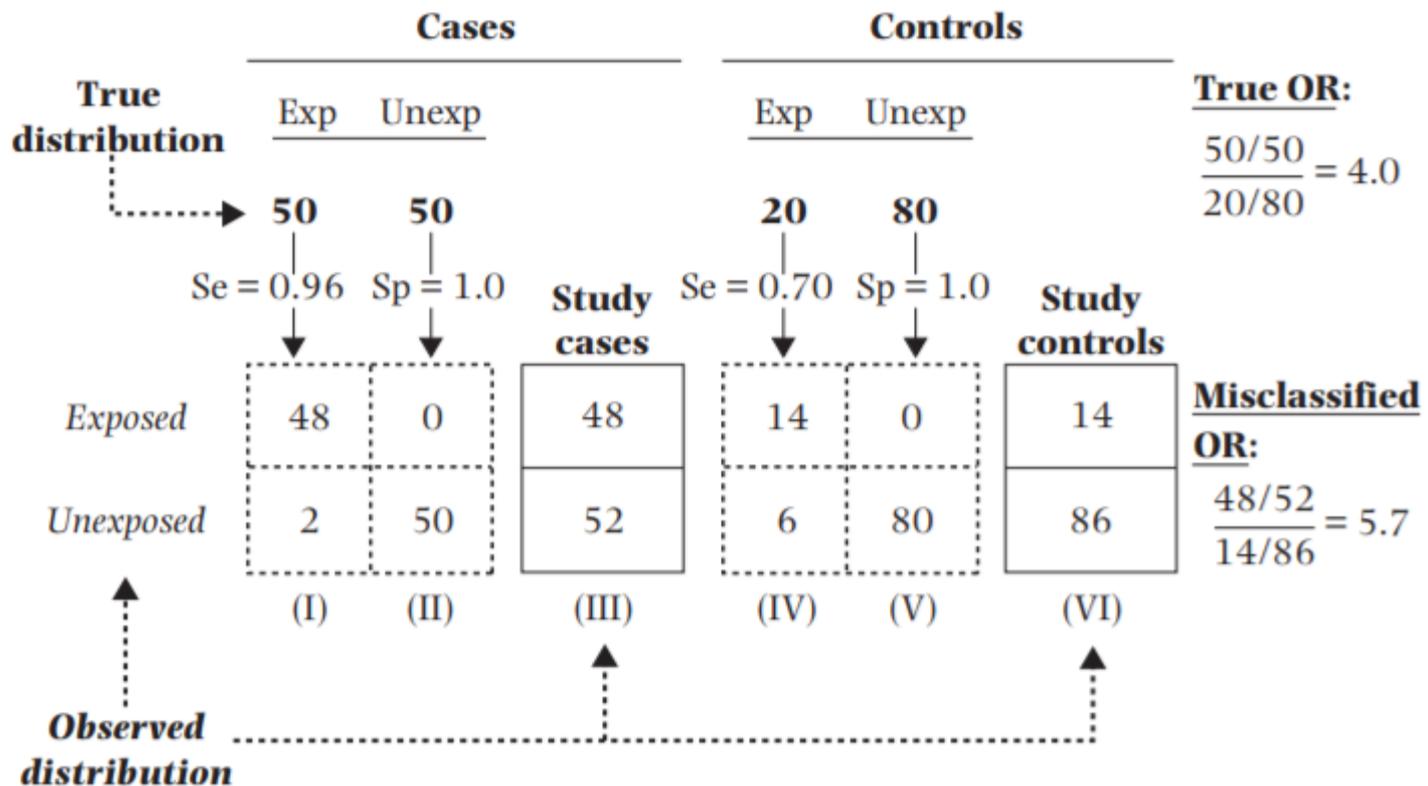
EXHIBIT 4-5 Hypothetical example of the effects of differential misclassification on the odds ratio, in which, for sensitivity, cases > controls, and for specificity, cases = controls.

True distribution	Cases		Controls		True OR:
	Exp	Unexp	Exp	Unexp	
→	50	50	20	80	$\frac{50/50}{20/80} = 4.0$
	Se = 0.96	Sp = 1.0	Se = 0.70	Sp = 1.0	
		Study		Study	

Note: Exp: Exposed; Unexp: Unexposed; Se: Sensitivity; Sp: Specificity; OR: Odds ratio

Viés de Informação

EXHIBIT 4-5 Hypothetical example of the effects of differential misclassification on the odds ratio, in which, for sensitivity, cases > controls, and for specificity, cases = controls.



Note: Exp: Exposed; Unexp: Unexposed; Se: Sensitivity; Sp: Specificity; OR: Odds ratio

Viés de informação

- Erros na mensuração das informações necessárias:
 - Viés na identificação da exposição
 - Viés de memória (entrevistado)
 - Viés do entrevistador / observador
 - Viés de registro
 - Viés no status caso-controle
 - Viés no critério diagnóstico

Viés de informação

Viés de memória:

- Os casos têm maior sensibilidade de se lembrar da exposição
 - estudos caso-controle
 - quando a história de exposição é obtida retrospectivamente os casos podem lembrar-se melhor de sua história de exposição;
 - controles, sem a carga da doença, tendem a se esquecer mais facilmente de sua história passada;

Exemplo: Mães de bebês com malformações

Viés de informação

- **Viés do observador:** processo pelo qual a coleta de informação pelo investigador difere entre casos e controles. Ex: observador inquiri mais longamente os casos do que controles em relação a exposição.

Idealmente, entrevistador desconhecer a hipótese do estudo e quem é caso e controle.

Informações de forma objetiva, instrumentos padronizados

Viés de informação

Viés de registro:

Indivíduos com doenças mais graves tendem a ter registros mais completos sobre exposições e portanto pode-se encontrar uma maior associação nestes casos

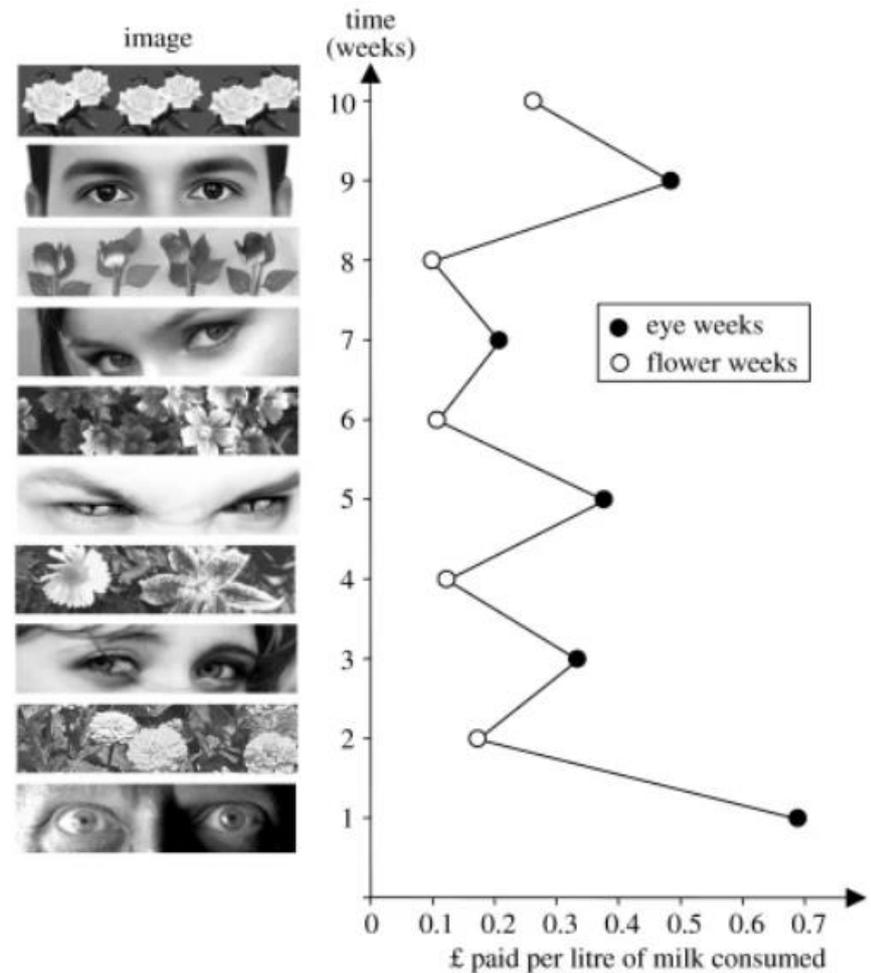
Quando um efeito é evidenciado apenas devido à atenção que os pacientes recebem dos pesquisadores.

Efeito Hawthorne

Indivíduos que sabem que são participantes de um estudo comportam-se de forma diferente

Viés de informação

Efeito Hawthorne (1926): segundo o qual as pessoas se comportam de forma diferente quando sabem que estão sendo observadas. Ou seja, apenas observar um indivíduo já promove uma intervenção em seu comportamento. Experimento realizado numa fábrica.



Valor pago por litro de leite consumido, em função da semana e do tipo de imagem. Bolinhas pretas representam semanas em que olhos foram exibidos e bolinhas brancas representam semanas em que flores eram mostradas.

Fonte: Bateson (2006)

Controlando Vieses de Informação

- Padronização do protocolo para coleta de dados
- As fontes e métodos de coleta de dados devem ser similares para todos os grupos de estudo.
- Manter, quando possível, entrevistadores e pessoal em desconhecimento da condição exposição/doença.
- Adaptar a estratégia para avaliar potenciais vieses de informação

Explicações	Estratégias de avaliação
Variação aleatória	Estimação de precisão (IC95%) e nível descritivo (p-valor)
Confusão	Desenho experimental; ajuste
Viés	Controle e garantia de qualidade
Relação causal	Eliminar demais explicações; Critérios de Hill

Referências

- ✓ Gordis L. Epidemiology. W.B. Saunders Company, Philadelphia, 2009.
- ✓ Hulley SB et al. Delineando a Pesquisa Clínica: uma abordagem epidemiológica. Trad. Duncan MS. Porto Alegre: Artmed, 2008.
- ✓ Szklo M; Javier Nieto F. Epidemiology: Beyond the basics. Fourth Edition Burlington, Massachusetts: Jones & Bartlett Learning; 2019.
- ✓ Rothman KJ. Epidemiology. An introduction. New York. Oxford University Press. 2002.