



Estudos caso-controllo

Planeamento e análise

Avaliação da associação entre factores de risco & doença

Estudos ...

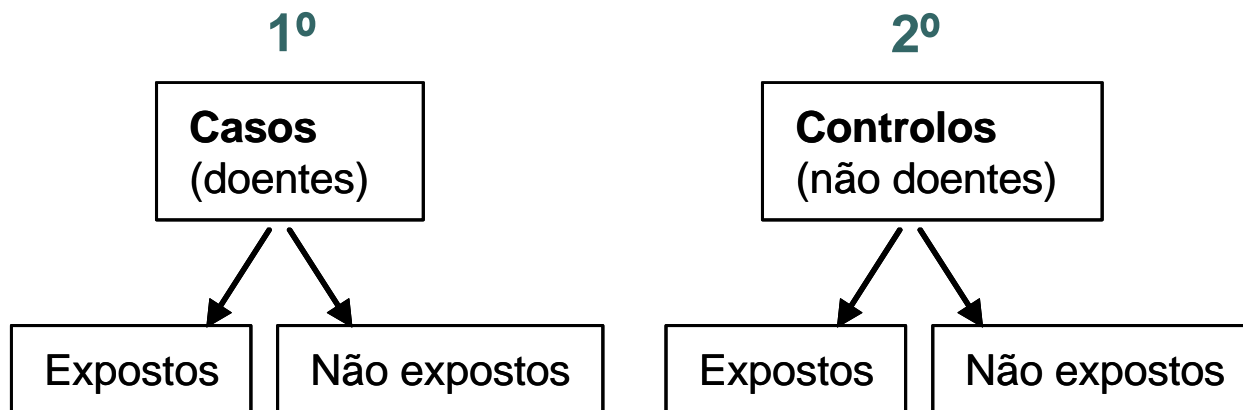
- Transversais
- Caso-controlo
- de Coortes
- Interventivos

Módulo	Metodologia de Investigação em CS
0	Apresentação
2	Epidemiologia descritiva
3	Estudos transversais
4	Estudos caso-controlo
5	Estudos de coortes
17	Ensaio clínico
6	Testes de diagnóstico

Estudos caso-control

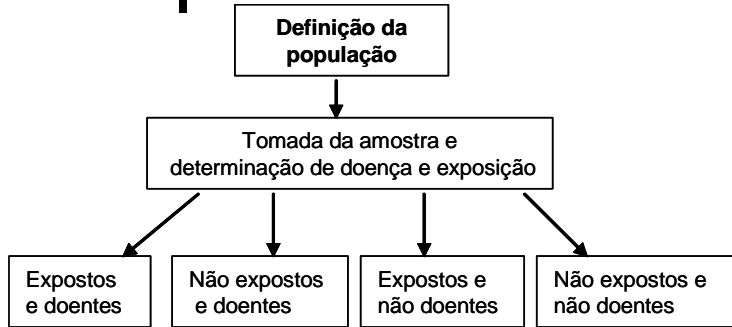
1. Toma-se uma amostra de indivíduos c/ a doença (os **casos**)
2. Toma-se outra amostra de indivíduos do “mesmo tipo” mas sem a doença (**controles**)

Investiga-se casos e controles quanto à exposição ao factor de risco



Se houver associação doença-factor risco, espera-se maior proporção de expostos entre os casos que entre os controles

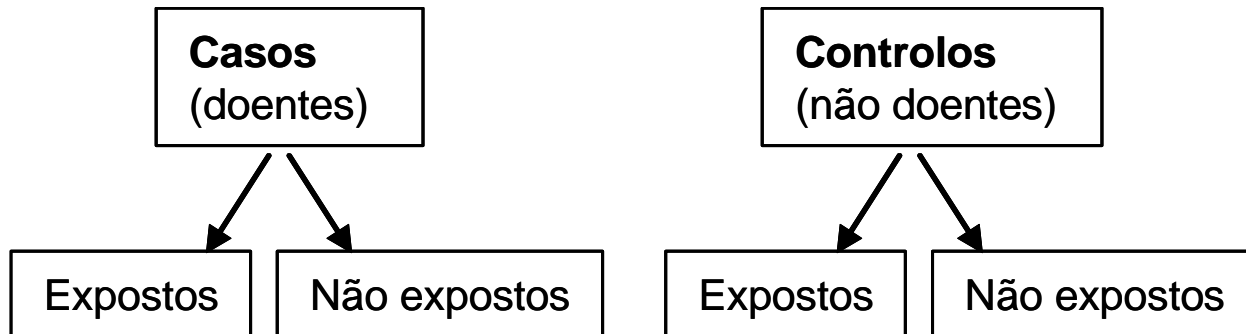
Transversais vs. Caso-controlo



Transversais (cross-sectional)

Amostra única (n) tomada de uma vez, em pouco tempo
Mede-se simultaneamente exposição e prevalência da doença

Caso-controlo





Escolha dos casos

- Casos seleccionados após definição prévia de caso
Existe tendência para alargar a definição a meio do processo, a fim de incluir mais casos
- Usar todos os encontrados (doenças raras) ou uma amostra de casos
- Se for amostra, deve ser representativa
- Podem ser localizados em hospitais, clínicas, ficheiros, etc.
Os casos podem provir de fontes múltiplas (e.g. vários hospitais)

DOENÇA MENINGOCÓCICA - DEFINIÇÃO DE CASO*

Descrição Clínica

Quadro clínico compatível com doença meningocócica, por exemplo meningite e/ou meningococcemia que pode degenerar rapidamente em púrpura fulminante, choque e morte. São possíveis outras manifestações.

Critérios laboratoriais para o diagnóstico

- Isolamento de *Neisseria meningitidis* de um local normalmente estéril (por exemplo, sangue ou líquido céfalo-raquidiano (LCR), ou, menos habitualmente, do líquido articular, pleural ou pericárdico).
- Detecção de ácido nucleico da *N. Meningitidis* a partir de um local normalmente estéril.
- Detecção do Antígeno da *N. meningitidis* a partir de um local normalmente estéril.
- Demonstração por microscópio de diplococos gram-negativos a partir de um local normalmente estéril.

Para um caso provável:

- Apenas um título elevado de anticorpos meningocócicos no soro convalescente.

Classificação do caso

Possível: Não aplicável.

Provável: Um quadro clínico compatível com doença meningocócica invasiva sem qualquer confirmação laboratorial, ou com identificação de *N. meningitidis* a partir de local não estéril, ou com níveis elevados de anticorpos meningocócicos em soro convalescente.

Confirmado: Um caso clinicamente compatível confirmado laboratorialmente.

É de salientar que os portadores assintomáticos não deverão ser notificados.

* Definição de caso publicada no Jornal Oficial das Comunidades Europeias a 03-04-2002.

Meningococo
(*Neisseria meningitidis*)

Copyright © 2001 Dennis Kunkel Microscopy, Inc. / Dennis Kunkel



Escolha dos controlos

- Em tudo idênticos aos casos... excepto que não têm a doença. Características idênticas às dos casos: mesma distribuição de idades, sexo, origem socio-económica, etc. (**comparabilidade**)
- Provir da mesma população em risco de doença que os casos e ser representativos dos não-doentes nessa população (**representatividade**)
- A sua selecção envolve muitas vezes uma fase não-aleatória
Em geral, os casos não são amostra aleatória da população e os controlos espelham isso.

Na prática, a **comparabilidade** entre controlos e casos é difícil de conseguir. A formação do grupo de controlo é em geral a parte difícil deste tipo de estudos.
A **comparabilidade** é mais importante que a representatividade



Explo – fontes de casos e controlos

Casos	Controlos
Todos os casos encontrados na população	Não-casos, numa amostra aleatória da população
Todos os casos num hospital	Amostra de pacientes no mesmo hospital que não têm a doença
Casos encontrados oportunisticamente em população difícil de definir	Esposos, vizinhos, familiares, amigos dos casos

Em qualquer das situações, os controlos são seleccionados de forma a ter:

- Mesma distribuição de idades que os casos
- Mesmo sex-ratio
- Mesma origem social
- Mesmos hábitos
- ...

Não são doentes mas podiam ser



Avaliação da exposição ao factor de risco

Exposição ao factor de risco é avaliada retrospectivamente

- É investigada retrospectivamente (arquivos, entrevistas etc)
- Assume-se que exposição teve lugar antes e/ou durante o desenvolvimento da doença
- Frequentemente sujeita a erros de memória e/ou enviesamento do entrevistador

Situação simples: 1 só factor de risco

	doentes	Não-doentes	
Factor de risco	CASOS	CONTROLOS	Total
Exposto	a	b	$a+b$
não exposto	c	d	$c+d$
Total	$a+c$	$b+d$	n

Calcular RR como de costume ?
NÃO !

n já não é uma amostra aleatória da população

A proporção de doentes $(a+c)/n$ já NÃO representa essa mesma proporção na população

Por que razão não se pode calcular RR

Sejam A, B, C, D os valores absolutos na população

Seja f_1 e f_2 as proporções de doentes e controlos amostradas

Factor de risco	(a) POPULAÇÃO			(b) Valores esperados na amostra		
	Doentes	Não doentes	Total	Casos	Controlos	Total
Exposto	A	B	A+B	$A f_1$	$B f_2$	$A f_1 + B f_2$
não exposto	C	D	C+D	$C f_1$	$D f_2$	$C f_1 + D f_2$
Total	A+C	B+D	N	$(A+C) f_1$	$(B+D) f_2$	n

RR verdadeiro:

$$RR = \frac{A(C + D)}{C(A + B)}$$

RR no conjunto das duas amostras:

$$\frac{A f_1 (C f_1 + D f_2)}{C f_1 (A f_1 + B f_2)} \neq \frac{A(C + D)}{C(A + B)}$$

Pode-se calcular OR ?

Sejam A, B, C, D os valores absolutos na população
Seja f_1 e f_2 as proporções de doentes e controlos amostradas

Factor de risco	(a) POPULAÇÃO			(b) Valores esperados na amostra		
	Doentes	Não doentes	Total	Casos	Controlos	Total
Exposto	A	B	A+B	$A f_1$	$B f_2$	$A f_1 + B f_2$
não exposto	C	D	C+D	$C f_1$	$D f_2$	$C f_1 + D f_2$
Total	A+C	B+D	N	$(A+C) f_1$	$(B+D) f_2$	n

OR verdadeiro:

$$OR = \frac{AD}{BC}$$

OR no conjunto das duas amostras:

$$OR = \frac{A f_1 D f_2}{B f_2 C f_1} = \frac{AD}{BC} \quad \text{Pode !}$$

Múltiplos factores de risco

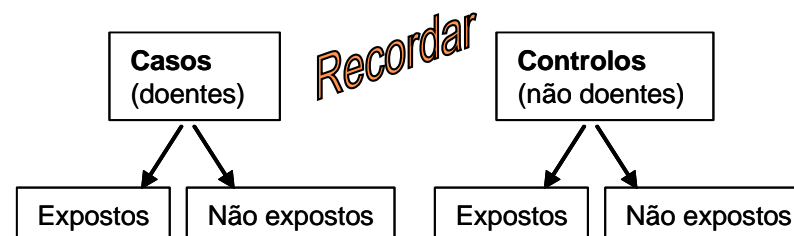
Expo: Surto de 37 casos de gastroenterite num self-service
Seleccionaram-se aleatoriamente 58 comensais para controlos
Não havia certeza sobre o dia em que se deu a infecção (15 ou 16 Ago ?)

	Gastroenterite		Sem gastroenterite	
	Comeu	Não comeu	Comeu	Não comeu
dia 15 Agosto	6	31	9	48
dia 16 Agosto	18	19	14	43
salada	12	24	5	52
sandwiches	16	21	14	44
ovos	4	33	4	54

Soma não dá 37

Soma não dá 58

Houve respostas “não me lembro”



A tabela original é decomposta para a análise de cada factor

	Gastroenterite		Sem gastroenterite	
	Comeu	Não comeu	Comeu	Não comeu
dia 15 Agosto	6	31	9	48
dia 16 Agosto	18	19	14	43
salada	12	24	5	52
sandwiches	16	21	14	44
ovos	4	33	4	54

<i>Factor de risco</i>	<i>Gastroenterite</i>	<i>não doente</i>	<i>Total</i>
comeu em 15 Ago	6	9	15
não comeu	31	48	79
Total	37	57	94

$$31 = 37 - 6$$

$$48 = 57 - 9$$

$$OR = (6/9) / (31/48) = 1,03$$



Intervalo de confiança para o OR constrói-se como habitualmente

O erro padrão do $\ln(OR)$ é estimado por $\hat{e}(\ln OR) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$

Um IC a 95% para o OR é dado por,

$$LI = e^{\ln OR - 1.96 \hat{e}(\ln OR)}$$

$$LS = e^{\ln OR + 1.96 \hat{e}(\ln OR)}$$

Se o IC **não** incluir 1, há 95% probabilidade de o factor ser mesmo de risco (ou de protecção)



Estudos caso-control: vantagens

- Bons para doenças raras – permitem reunir numa amostra os poucos casos encontrados
- Muito mais rápidos e baratos que os estudos de coortes (estes acompanham indivíduos ao longo dos anos). São por isso bons para doenças de longa latência
- Podem-se estudar vários factores de risco em simultâneo. Basta interrogar casos e controlos sobre mais de um factor

... e desvantagens



- Várias fontes de enviesamento:
 - Os controlos não são totalmente equivalentes aos casos
 - A 'memória' de exposição dos casos é em geral muito melhor que a dos controlos. Às vezes é "boa" demais.
- O odds ratio é apenas uma aproximação ao RR
- Ser um caso pode ter mais a ver com ser "sobrevivente" que ser um doente
(em doenças que matam, podemos não conseguir incluir no estudo os doentes mais expostos ao factor de risco porque já morreram)
- É difícil estabelecer uma sequência temporal 1º exposição, 2º doença



Alguns estudos históricos do tipo caso-controlo

1950's

- Tabaco e cancro do pulmão

1970's

- Diethyl stilbestrol (DSE) e adenocarcinoma vaginal
- Estrogéneos post-menopausa e cancro do endométrio

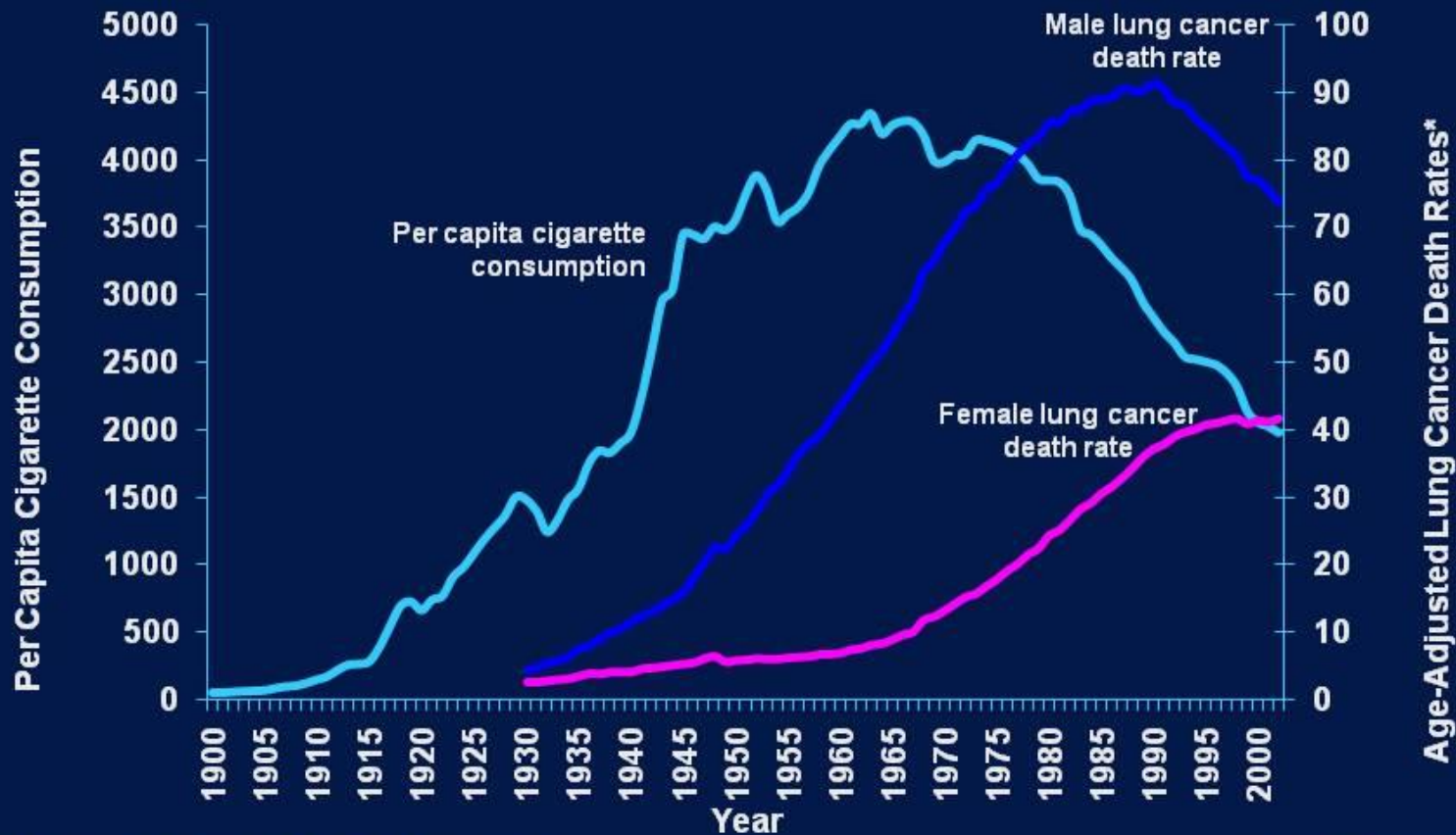
1980's

- Aspirina e síndrome de Reyes
- Certo tipo de tampões menstruais e Síndrome do Choque Tóxico (TSS)
- Práticas sexuais que provocam rompimento das mucosas e o VIH/SIDA

1990's

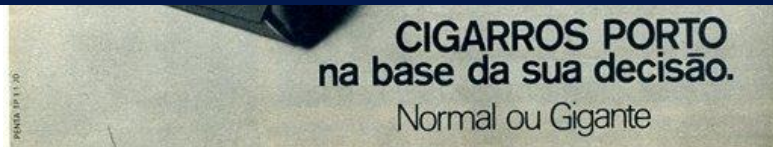
- Eficácia vacinal
- Dieta e cancro
- *Helicobacter pylori* e cancro do estômago

Tobacco Use in the U.S., 1900-2002



*Age-adjusted to 2000 US standard population.

Source: Death rates: US Mortality Public Use Tapes, 1960-2002, US Mortality Volumes, 1930-1959, National Center for Health Statistics, Centers for Disease Control and Prevention, 2005. Cigarette consumption: US Department of Agriculture, 1900-2002.





Alguns estudos históricos do tipo caso-controlo

1950's

- Tabaco e cancro do pulmão

1970's

- Dietil stilbestrol (DSE) e adenocarcinoma vaginal
- Estrogéneos post-menopausa e cancro do endométrio

1980's

- Aspirina (salicilatos) e síndrome de Reyes
- Certo tipo de tampões menstruais e Síndrome do Choque Tóxico (TSS)
- Práticas sexuais que provocam rompimento das mucosas e o VIH/SIDA

1990's

- Eficácia vacinal
- Dieta e cancro
- *Helicobacter pylori* e cancro do estômago



Houve também enganos históricos (1982)

Casos – 20 homossexuais com sarcoma de Kaposi

Controlos – 40 homossexuais sem sarcoma

Factor de risco – uso/não uso de “poppers” (estimulantes sexuais)

OR= 10 !!

Durante anos discutiu-se a possibilidade dos estimulantes serem mutagénicos !

Confundimento !

A toma de estimulantes estava associada a práticas sexuais de risco



Risco Atribuível

A proporção de casos de doença na população que se devem à exposição ao factor de risco

$$RA = \frac{p_E (RR - 1)}{1 + p_E (RR - 1)}$$

p_E = prevalência do factor de risco na população
 RR = risco relativo

RR não é calculado mas o risco atribuível ainda pode ser calculado se:

- 1) O OR for uma boa aproximação ao RR (doença relativamente rara).
- 2) A proporção de expostos ao factor de risco no grupo de controlo, é uma boa estimativa dessa proporção na população.

A prevalência dos expostos é então estimada por $b/(b+d)$.



Risco Atribuível

Substituindo RR por ad/bc
 p_E por $b/(b+d)$



$$RA = \frac{p_E(RR - 1)}{1 + p_E(RR - 1)}$$



$$RA = \frac{ad - bc}{d(a + c)}$$

Estima o Risco Atribuível



Intervalo de confiança para o risco atribuível

O verdadeiro valor do Risco Atribuível está no IC definido por

$$\left(1 + \frac{1 - RA}{RA} e^{\pm u} \right)^{-1}$$

Para um IC de 95%,

$$u = \frac{1.96 d(b+d)}{a(b+d) - b(a+c)} \sqrt{\frac{a}{c(a+c)} + \frac{b}{d(b+d)}}$$



Escolha de controlos (revisão)

Preferível:

Controlos escolhidos na mesma comunidade cultural e sócio-económica de onde os casos são oriundos

Frequente: emparelhamento

Familiares, amigos ou vizinhos escolhidos a dedo por terem a mesma idade, sexo, hábitos etc.

Não confundir: emparelhamento com comparabilidade

Comparabilidade: tomada de amostra aleatória com idêntica distribuição de idades, sexos, etc.

Emparelhamento: escolha individualizada, não aleatória, de pares para cada caso (em geral 2 ou 3 emparelhamentos por caso)
(O emparelhamento garante sempre comparabilidade).

Com emparelhamento, a análise estatística é um pouco diferente



Quantos controlos ?

Número de casos – em geral limitados pelos que há disponíveis

Número de controlos – em geral existe possibilidade de escolha

Recordar

$$\hat{e}(\ln OR) = \sqrt{\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d}}$$

$$LI = e^{\ln OR - 1.96 \hat{e}(\ln OR)}$$

$$LS = e^{\ln OR + 1.96 \hat{e}(\ln OR)}$$

Número de controlos demasiado pequeno

Gera estimativas de OR com um IC demasiado grande

Número de controlos demasiado grande

Significa gasto de recursos sem contrapartidas em termos de precisão



Noções fundamentais sobre... ... o método científico !

Suponhamos que se suspeita que o factor de risco X está associado à doença.

Formalmente o procedimento estatístico evolui da seguinte forma:

- (1) Coloca-se a hipótese de que X NÃO está associado à doença
É a chamada hipótese nula, ou H_0
- (2) Concebe-se um teste para rejeitar H_0
por ex: contruir um IC 95% em torno do OR e comparar com 1
- (3) O teste pode rejeitar ou não a H_0
- (4) - Se rejeitar, dizemos “os dados fornecem evidência de que H_0 é errada”
Temporariamente aceitamos H_1

- Se não rejeitar dizemos “não existe evidência suficiente para rejeitar H_0 ”



Dois erros possíveis no teste

Qualquer estudo científico deve evitar dois erros:

(1) Tipo I - Concluir que a exposição está associada à doença quando de facto não está

A probabilidade de cometer este erro chama-se “nível de significância”, α

o nosso teste controla α , ao construir o IC a 95%, estabelece $\alpha=0.05$

(2) Tipo II - Concluir que exposição não está associada à doença, quando de facto está

A probabilidade de cometer este erro denomina-se β ; e $1-\beta$ chama-se a “potência” do estudo

Se, na realidade, a exposição tem um RR diferente de 1, a potência é a probabilidade de que o OR seja significativamente diferente de 1.

O aumento da potência consegue-se com o aumento da amostra (casos e/ou controlos)

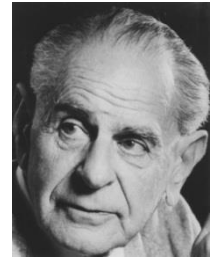
Hipótese nula: Não há associação

		Teste de H_0	
		Não rejeita H_0	Rejeita H_0
<i>Mundo real</i>	H_0 é verdadeira	$1-\alpha$	α
	H_0 é falsa	β	$1-\beta$ (potência)

Se o teste rejeita, sabemos que a probabilidade de errar é α

Se o teste não rejeita, em geral não conhecemos β

Raciocínio “Popperiano”



Sir Karl Popper, 1980

Racionalismo crítico

Para ser genuinamente científica, uma hipótese (ou teoria) tem de ser “falsificável”

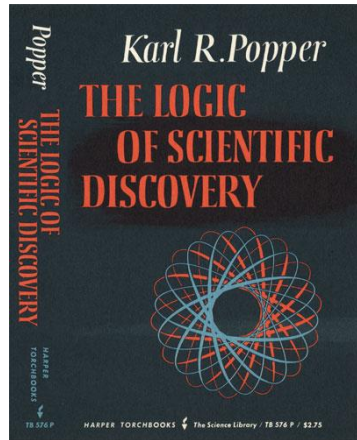
i.e. tem de ser possível conceber um teste (observacional ou experimental) que ponha a hipótese à prova.

Um teste não exclui a possibilidade de uma teoria não rejeitada vir a ser rejeitada mais tarde por um teste mais apurado.

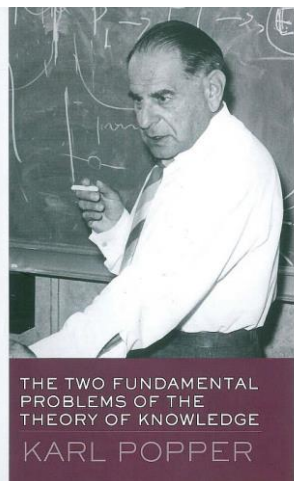
Os nossos testes não podem “demonstrar” as teorias, apenas as podem rejeitar.



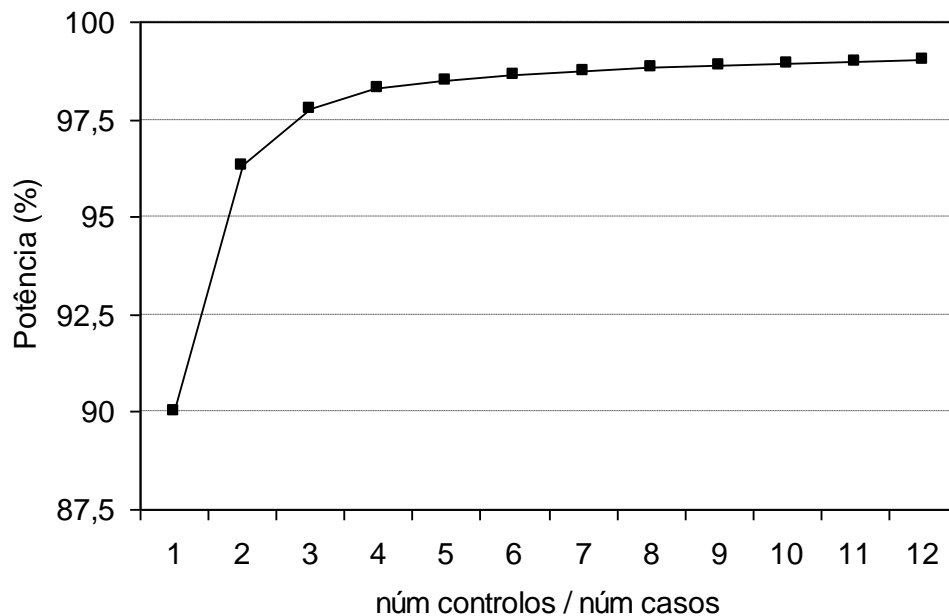
Sir Karl Popper, 1990



(c) 1934



Aumento da potência com o número de controlos



Potência

Probabilidade de que a hipótese nula (H_0 : não existe associação factor-doença) seja rejeitada quando é falsa

Em geral **2 a 3 controlos por caso** é o ideal

Raramente vale a pena ter mais de 4 controlos por caso