



Estudos de coortes

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

A coorte

Coorte de nascimento = Classe anual

Conjunto de todos os indivíduos nascidos no mesmo intervalo de tempo ao longo de toda a sua vida.

Em geral o intervalo de tempo considerado é 1 ano

Idades	1996	1997	1998	1999	2000	2001
0	1000					
1		800				
2			600			
3				450		
4					300	
5						150

Coortes e população



Coorte de 1996

Coorte de 1995

1996	1997	1998	1999	2000	2001
1000	1100	980	1020	975	1151
750	800	735	811	820	740
620	625	600	583	605	570
435	455	425	450	470	440
315	280	326	276	300	305
128	155	162	130	155	150

Coorte (*lato sensu*)

Coorte

Conjunto de indivíduos **definidos segundo determinado critério**, ao longo de um período de tempo.

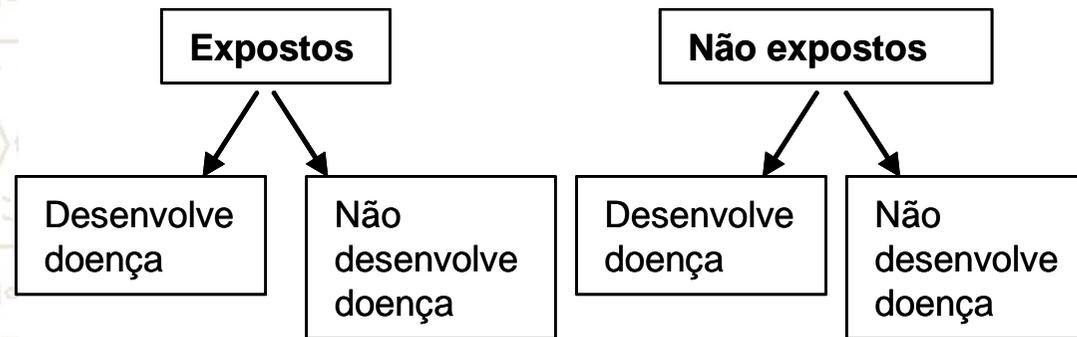
(Em geral o período para o qual o critério é válido)

Expos

- Os alunos que dão entrada na FE-UAAlg e usam a cantina
- Os alunos que dão entrada na FE-UAAlg e não usam a cantina
- Os alunos que entraram no MGUS em 2011 ao longo do curso
- Os homens fumadores com 35 a 60 anos de idade de uma vila algarvia
- Os trabalhadores da secção de pinturas de uma fábrica de automóveis

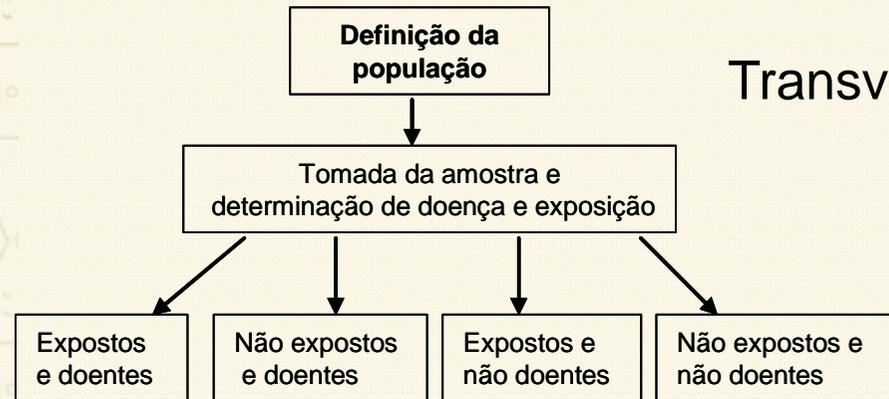
Estudos de coortes

1. Seleccionam-se 2 grupos de indivíduos não doentes. Um grupo está exposto ao factor de risco. Outro não está.
2. Acompanham-se os dois grupos ao longo do tempo e, no fim, a incidência da doença é comparada entre os dois grupos

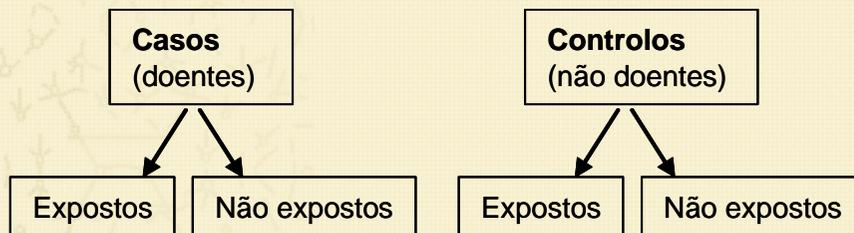


Se houver associação factor-doença, espera-se maior proporção de doentes no grupo dos expostos que nos não-expostos

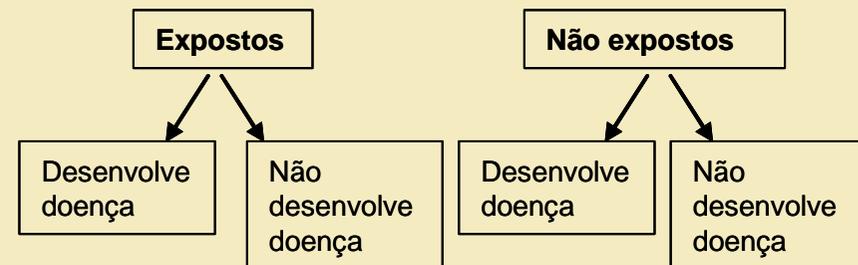
Transversais, Caso-controlo, Coortes



Transversais (cross-sectional)



Caso-controlo
(retrospectivos)

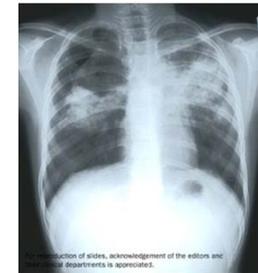


Coortes
(longitudinais)

Exemplo

513 toxicod dependentes foram testados para HIV e seguidos durante 2 anos
Registando-se se desenvolveram TB-doença

Estado inicial	Desenvolveu TB ?		total
	sim	não	
HIV +	8	207	215
HIV -	0	298	298

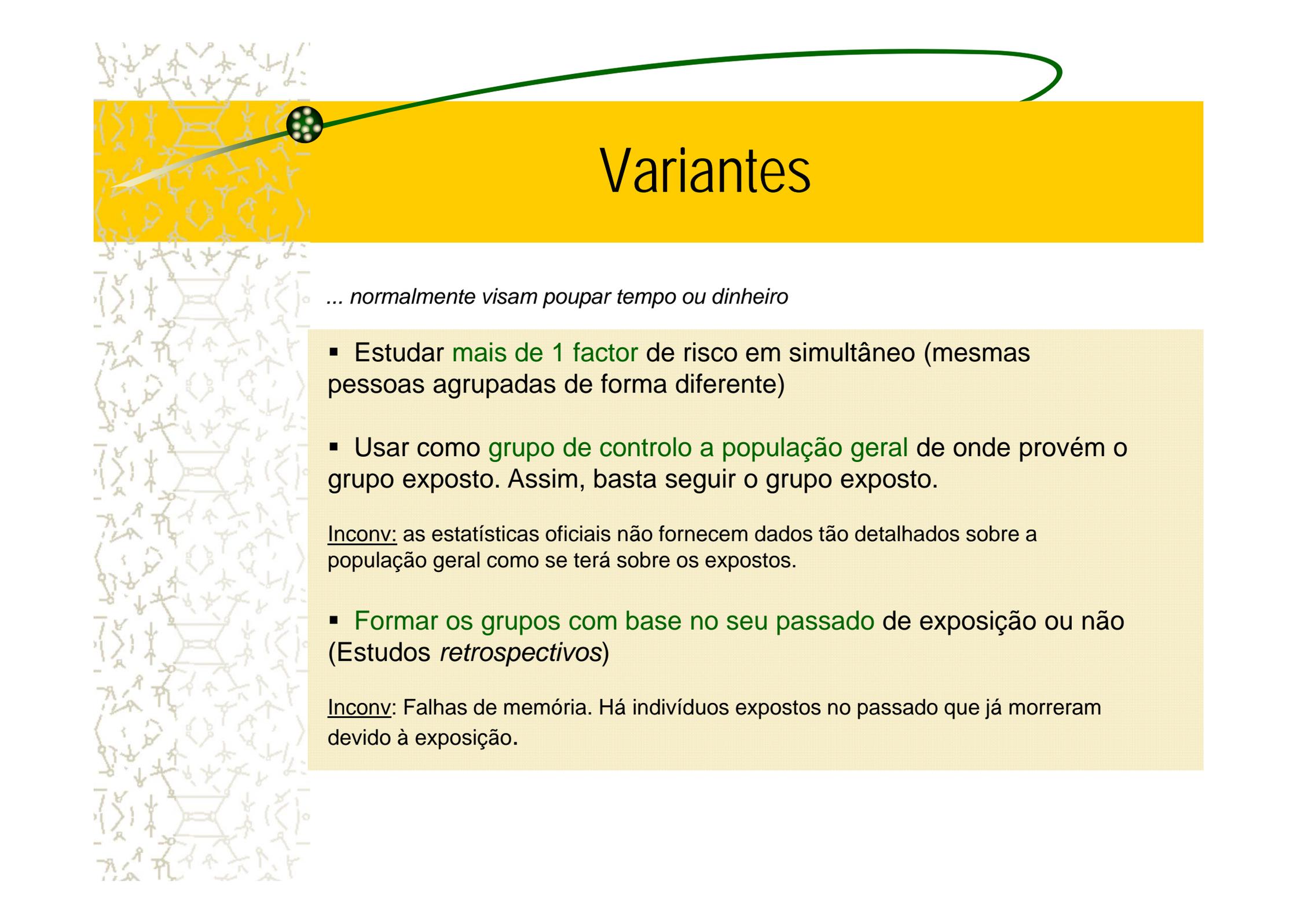


No início, os indivíduos devem ser testados para TB-doença e excluídos caso tenham sintomas.

(Fundamental para se poder sequer falar em causalidade)

Neste exemplo, a exposição (ou não) ao factor não é óbvia, tem de ser investigada





Variantes

... normalmente visam poupar tempo ou dinheiro

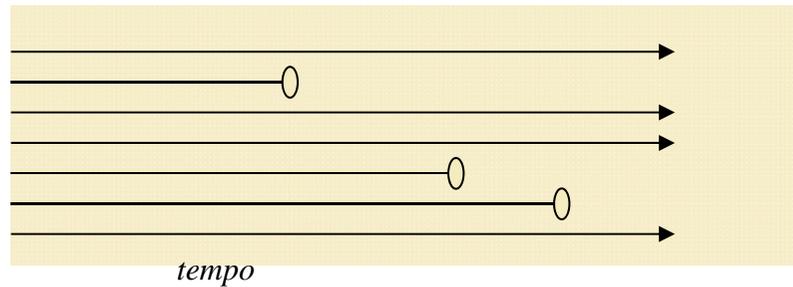
- Estudar **mais de 1 factor** de risco em simultâneo (mesmas pessoas agrupadas de forma diferente)
- Usar como **grupo de controlo a população geral** de onde provém o grupo exposto. Assim, basta seguir o grupo exposto.

Inconv: as estatísticas oficiais não fornecem dados tão detalhados sobre a população geral como se terá sobre os expostos.

- **Formar os grupos com base no seu passado** de exposição ou não (Estudos *retrospectivos*)

Inconv: Falhas de memória. Há indivíduos expostos no passado que já morreram devido à exposição.

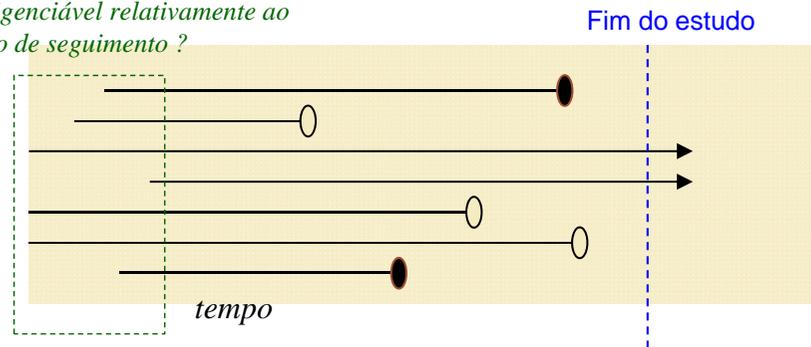
Coorte fixa vs coorte variável



○ doença

Coorte fixa – Todos começam ao mesmo tempo e são seguidos durante o mesmo tempo (ou até que ocorra doença)

Negligenciável relativamente ao tempo de seguimento?



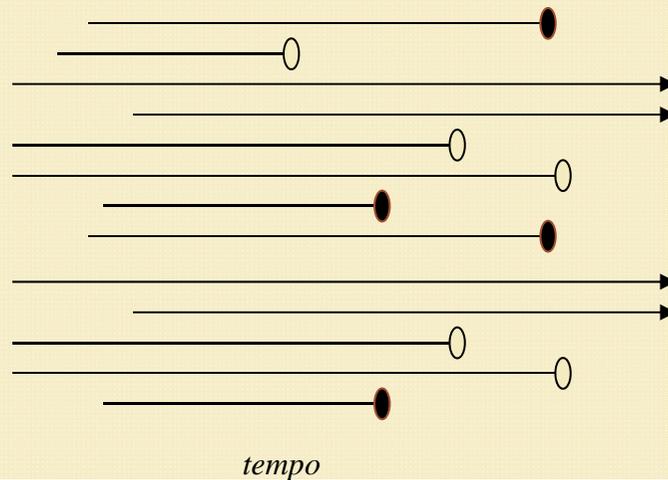
○ doença

● abandono

→ censurados

Coorte variável – Nem todos começam ao mesmo tempo. Nem todos chegam ao fim, apesar de não terem adoecido (“abandonos”) e alguns são “censurados”

Amostras grandes: riscos em competição



- doença
- Abandono por razões alheias ao estudo
- censurados

Coortes de amostras grandes Expto: ensaios clínicos de medicamentos

Os “Riscos em competição” tornam-se muito relevantes
A probabilidade de aparecimento de abandonos por razões alheias à doença aumenta.

Cálculos para coorte fixa

	<i>Doente</i>	<i>não Doente</i>	<i>Total</i>
Exposto	<i>a</i>	<i>b</i>	<i>a+b</i>
não Exposto	<i>c</i>	<i>d</i>	<i>c+d</i>
Total	<i>a+c</i>	<i>b+d</i>	<i>n</i>

Risco para os expostos = $a/(a+b)$
Risco para os não expostos = $c/(c+d)$

Risco relativo

$$RR = \frac{a/(a+b)}{c/(c+d)} = \frac{a(c+d)}{c(a+b)}$$

RR pode ser calculado porque se define *n* à partida, quando ainda ninguém é doente

Intervalos de confiança calculam-se como habitualmente

Pessoas-tempo

Coorte variável - Os indivíduos entram e saem do estudo em instantes diferentes
- Os riscos em competição provocam abandonos
Como contar os indivíduos em estudo ?

Ideia : usar o somatório de tempos de estadia no estudo (em vez do número de indivíduos)

1. Para cada indivíduo: medir tempo de estadia até adoecer ou abandonar o estudo
2. Calcular Z = soma de “pessoas-tempo” para cada grupo (expostos e não-expostos)

Exemplo

8 seropositivos para a hepatite B, foram seguidos em alturas diferentes, registando-se se adoeciam com cirrose.

Registo do seguimento:

Seguimento				
Indivíduo	Início	Fim	Tempo de seguimento	Doença ?
1	05-Out-56	01-Dez-93	446	sim
2	10-Out-69	31-Dez-97	339	sim
3	10-Jun-79	31-Dez-97	223	não
4	30-Ago-84	28-Set-94	121	não
5	08-Mai-62	08-Jul-91	350	sim
6	01-Nov-66	10-Mai-79	150	sim
7	21-Mar-54	30-Jun-91	447	não
8	08-Jun-61	29-Jul-65	50	sim
total:			2126	sim= 5
			meses	não=3

Z = 2126 pessoas-mês

Taxa de incidência, X/Z

Taxa de Incidência (TI):

$$\frac{X}{Z}$$

Nº casos de doença

Pessoas-tempo

casos de doença por pessoa-tempo

Exposto	Adoeceu ?		Total (pessoas-tempo)
	sim	não	
sim	a		Z dos expostos
não	c		Z dos não-expostos

Totais marginais têm natureza diferente. Não são a soma das linhas da tabela.

TI dos expostos: a/Z exp

TI dos não expostos: c/Z não-exp

Características da taxa de incidência

$\frac{X}{Z}$ → N° casos doença
→ Pessoas-tempo

Não é uma proporção

X não é limitado superiormente

X/Z aumenta quando a unidade de Z alarga:

$$\frac{X}{\text{pessoas} \times 365 \text{ dias}} \quad , \quad \frac{X}{\text{pessoas} \times 12 \text{ meses}} \quad , \quad \frac{X}{\text{pessoas} \times 1 \text{ ano}}$$

Unidades do denominador

São arbitrárias: - pessoa-ano
- pessoa-mês
- 1000 pessoas-ano

Taxa instantânea

Pode ser interpretada como uma velocidade

60 Km/hora dá a velocidade instantânea, embora expressa para um período não-instantâneo (1 h)

Unidades da TI

Tempo⁻¹

IC para a taxa instantânea, com X grande

$$TI = \frac{X}{Z}$$

Se os casos de doença ocorrem aleatoriamente ao longo do tempo, devem ser bem descritos por um processo de Poisson. Se X for grande ($X > 20$), um intervalo de confiança a 95% é dado por,

$$\left(\frac{X}{Z} \right) \pm 1.96 \sqrt{\frac{X}{Z^2}}$$

IC para a taxa instantânea, com X pequeno

$$TI = \frac{X}{Z}$$

Se os casos de doença ocorrem aleatoriamente ao longo do tempo, devem ser bem descritos por um processo de Poisson. Se X for pequeno ($X < 20$), um intervalo de confiança devia ser construído por métodos exactos. Uma boa aproximação, contudo, é dada pela fórmula de Byar,

$$\frac{(X + 0.5) \left(1 - \frac{1}{9(X + 0.5)} \pm \frac{1.96}{3} \sqrt{\frac{1}{X + 0.5}} \right)^3}{Z}$$

(1.96 na fórmula é para um IC de 95%)

Da TI ao risco

Para converter uma TI em risco:

$$r = 1 - e^{-TI \times t}$$

Se $TI = 0$, obtém-se risco = 0

Se $TI = +\infty$, obtém-se risco = 1

Assume-se que

$$\frac{dN}{dt} = -TI N$$

(EDO de 1ª ordem)

Ver apontamentos teóricos

RR com taxas instantâneas

$$RR = \frac{r_1}{r_2} = \frac{1 - e^{-TI_1 t_1}}{1 - e^{-TI_2 t_2}}$$

*Risco Relativo de doença
nos expostos ao fim do tempo t*

RR :
Quantas vezes número de casos
ocorridos em uma unidade de tempo é
superior entre expostos relativamente a
não expostos”.

Se o expoente ($TI t$) for pequeno, verifica-se que

$$1 - e^{-TI t} \approx TI t$$

$TI t$	$1 - e^{-TI t}$
0.01	0.01
0.05	0.05
0.10	0.10
0.15	0.14
0.20	0.18

Logo,

$$RR = \frac{r_1}{r_2} \approx \frac{TI_1}{TI_2}$$

NOTA – Se t for grande, esta simplificação não deve ser usada

IC para o RR com taxas instantâneas

$$RR = \frac{r_1}{r_2} \approx \frac{X_1/Z_1}{X_2/Z_2}$$

X_1/Z_1 é a TI nos expostos

X_2/Z_2 é a TI nos não expostos

Intervalo de confiança aproxim a 95%

$$RR_{inf,sup} = e^{Ln RR \pm 1.96 \sqrt{1/X_1 + 1/X_2}}$$

Retomando este exemplo...

8 seropositivos para a hepatite B, foram seguidos em alturas diferentes, registando-se se adoeciam com cirrose.

Registo do seguimento:

Indivíduo	Seguimento		Tempo de seguimento	Doença ?
	Início	Fim		
1	05-Out-56	01-Dez-93	446	sim
2	10-Out-69	31-Dez-97	339	sim
3	10-Jun-79	31-Dez-97	223	não
4	30-Ago-84	28-Set-94	121	não
5	08-Mai-62	08-Jul-91	350	sim
6	01-Nov-66	10-Mai-79	150	sim
7	21-Mar-54	30-Jun-91	447	não
8	08-Jun-61	29-Jul-65	50	sim
total:			2126	sim= 5
			meses	não=3

Z = 2126 pessoas-mês

Exemplo (conti.)

Outro grupo de 14 indivíduos seronegativos, foi também seguido durante um total de 4725 pessoas-mês, tendo 3 adoecido e 11 não adoecido.

Construindo a tabela de contingência,

	Cirrose	Não cirrose	total	TI	IC a 95% para as TI	
					L inf	L sup
HBsAg +	5		2126	0,00235	0,00029	0,00441
HBs Ag -	3		4725	0,00063	-0,00008	0,00135
	8		6851	0,00117	0,00036	0,00198

X (casos doença)

Z (pessoas tempo)

$$\left(\frac{X}{Z} \right) \pm 1.96 \hat{e} \left(\frac{X}{Z} \right)$$

$$RR = 0.00235/0.00063 = 3.7$$

Exemplo (conti.)

	Cirrose	não cirrose	total
HBsAg +	5		2126
HBs Ag -	3		4725
	8		6851

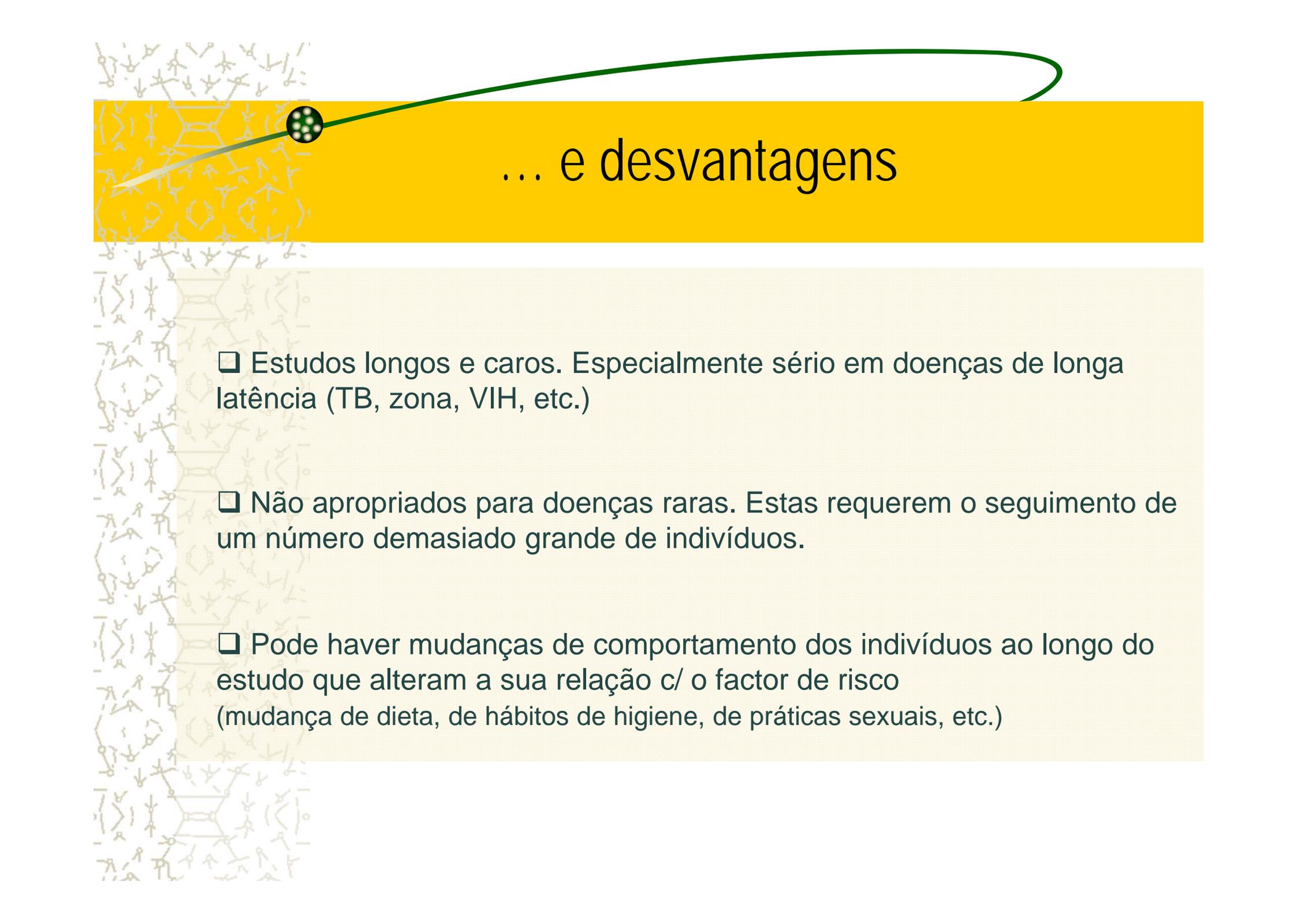
Os IC's para o RR calculam-se por: $RR_{inf, sup} = e^{Ln RR \pm 1.96 \sqrt{1/X_1 + 1/X_2}}$

RR	Ln RR	RaizQ(1/X ₁ + 1/X ₂)	LI Ln(RR)	LS (RR)	LI RR	LS RR
3.70	1.309	0.730	-0.122	2.741	0.88521	15.500



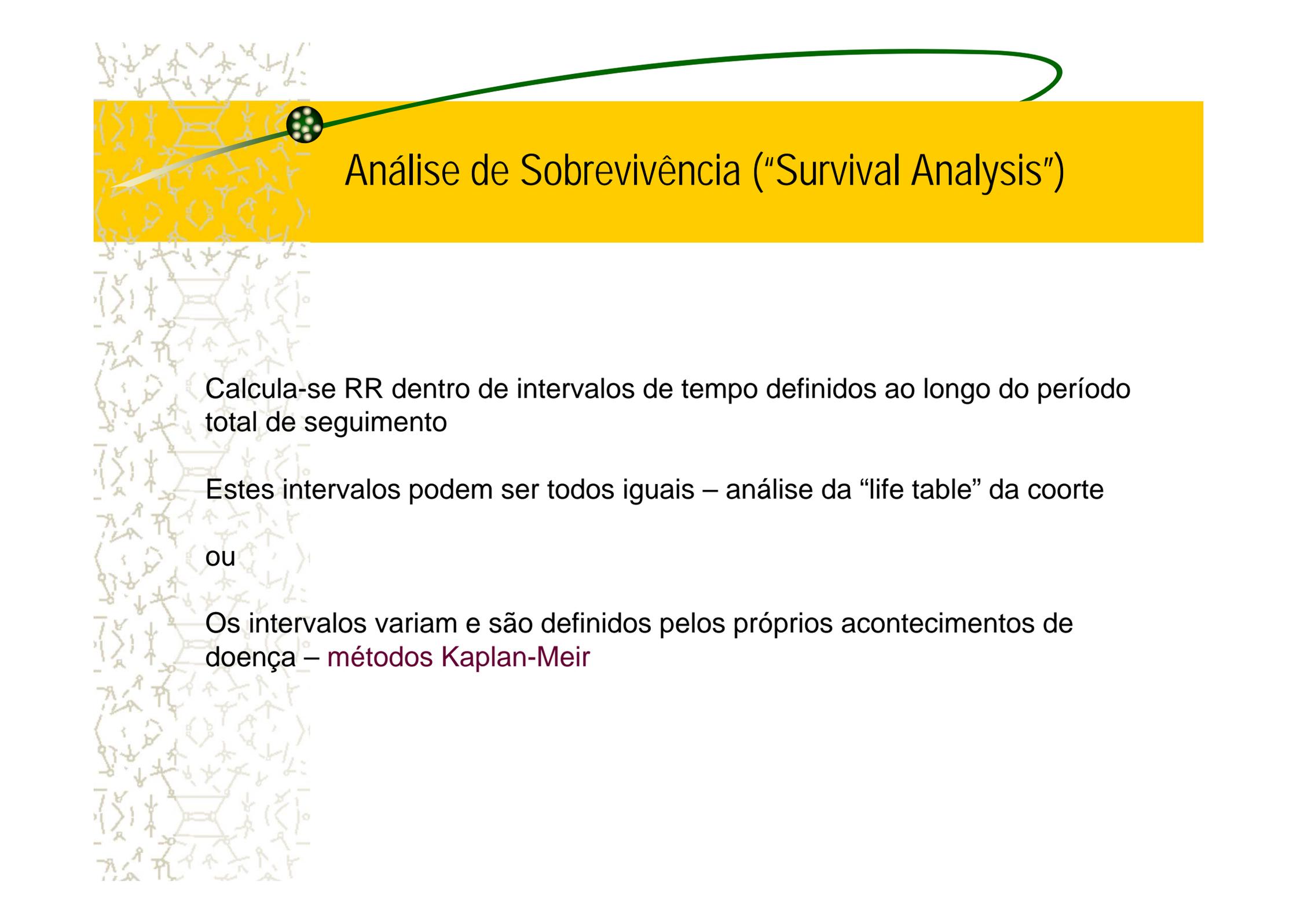
Estudos de coortes: vantagens

- Acompanham a sequência de acontecimentos – permitem o estabelecimento de causalidade
(nos estudos transversais e caso-controlo, já há doentes à partida)
- Possível estudar várias doenças ao mesmo tempo. Basta registar a sua incidência ao longo do tempo.



... e desvantagens

- ❑ Estudos longos e caros. Especialmente sério em doenças de longa latência (TB, zona, VIH, etc.)
- ❑ Não apropriados para doenças raras. Estas requerem o seguimento de um número demasiado grande de indivíduos.
- ❑ Pode haver mudanças de comportamento dos indivíduos ao longo do estudo que alteram a sua relação c/ o factor de risco (mudança de dieta, de hábitos de higiene, de práticas sexuais, etc.)



Análise de Sobrevivência ("Survival Analysis")

Calcula-se RR dentro de intervalos de tempo definidos ao longo do período total de seguimento

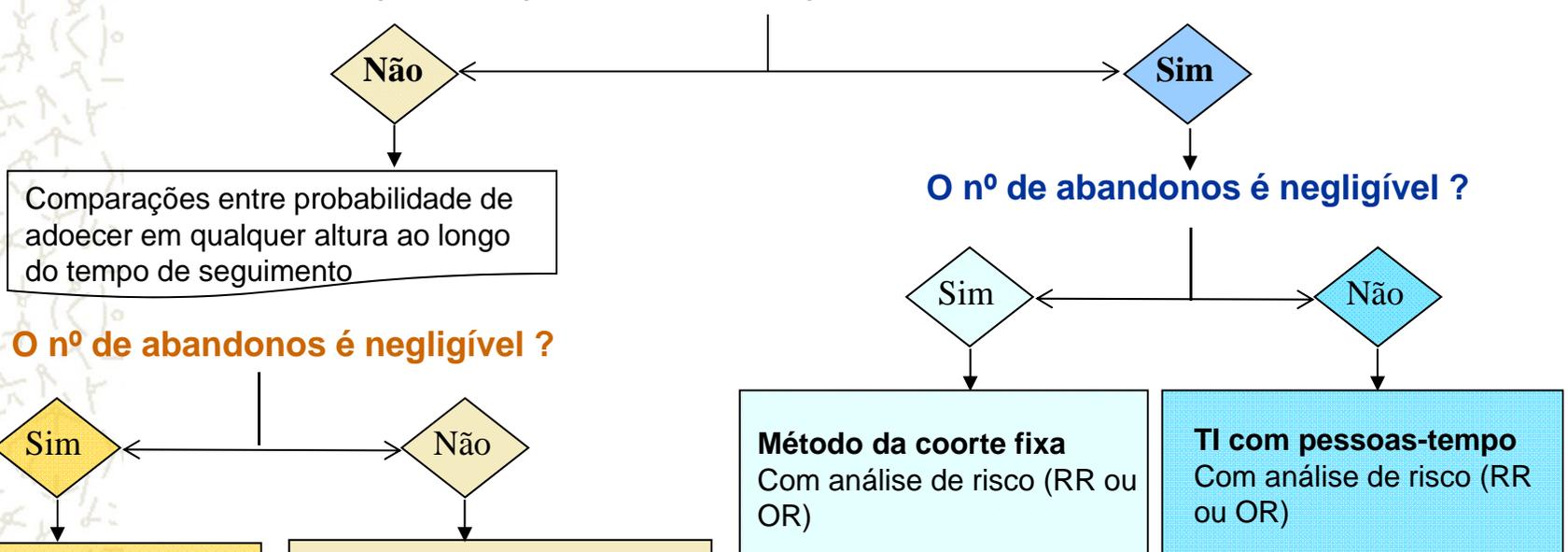
Estes intervalos podem ser todos iguais – análise da "life table" da coorte

ou

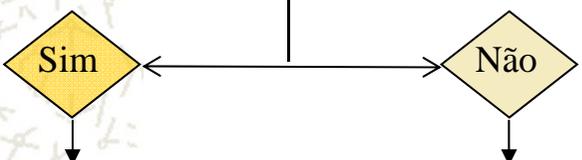
Os intervalos variam e são definidos pelos próprios acontecimentos de doença – métodos Kaplan-Meier

Métodos em estudos de coortes

Comparar expostos e não-expostos só no fim do estudo ?



O nº de abandonos é negligível ?



Método da Life Table da Coorte
O seguimento é dividido em intervalos de tempo arbitrários e em geral iguais

Método actuarial
Intervalos de tempo arbitrários e iguais
Métodos Kaplan-Meier
O seguimento é dividido em intervalos de tempo determinados pelos próprios episódios em estudo

Método da coorte fixa
Com análise de risco (RR ou OR)

TI com pessoas-tempo
Com análise de risco (RR ou OR)

Análise de Sobrevivência