

## Medidas de Associação.

O cálculo de medidas apropriadas da frequência de uma doença é a base para a comparação de populações, e, conseqüentemente, para a identificação de determinantes da doença. Para fazer isto de maneira mais eficaz e informativa, as duas frequências que vão ser comparadas podem ser combinadas num único parâmetro que estime a associação entre uma exposição e o risco de desenvolver a doença. Isto pode ser feito através de:

- **Medidas de associação** – quantificam a relação entre uma dada exposição e uma consequência. (*Dizem-nos quão mais susceptível está um grupo de desenvolver a doença do que outro*)
- **Medidas de impacto** – quantificam o impacto da mudança de exposição num dado grupo. (*Indicam-nos numa escala absoluta quão maior é a frequência de uma doença num grupo comparado com outro*).

Nesta aula vamos apenas abordar as medidas de associação.

Para ajudar no cálculo de medidas de associação, os dados epidemiológicos são muitas vezes apresentados numa tabela 2 por 2, também chamada de tabela de contingência:

	Doentes	Não doentes
Expostos	a	B
Não expostos	c	D

### A - Risco Relativo:

(para os estudos de coorte)

- estima a magnitude da associação entre exposição e doença;
- indica a probabilidade de desenvolver a doença no grupo exposto relativamente àqueles que não estão expostos;
- é definido como a razão entre a incidência da doença no grupo dos expostos ( $I_e$ ) e a incidência da doença no grupo dos não expostos ( $I_o$ ).

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

- um RR de 1,4 significa que os expostos têm 1,4 vezes mais risco de desenvolver a doença que os não expostos (que é o mesmo que dizer 40% mais probabilidade de desenvolver a doença).
- O cálculo do risco relativo, visto que é uma razão entre 2 riscos ou incidências cumulativas, depende do período de tempo no qual os riscos foram calculados. Este valor pode mudar dependendo do tempo de observação; o risco relativo após dez anos pode ser bastante

diferente daquele calculado após 1 ano. Não nos podemos esquecer de ter um tempo de seguimento relevante: por exemplo, se estivermos a comparar taxas de mortalidade de 2 populações e o tempo de seguimento for muito grande, vamos ter 100% de mortalidade tanto no grupo dos expostos como no grupo dos não expostos, e o RR seria de 1,0.

(para os estudos de casos-controlos)

Não podemos calcular incidências em estudos de casos-controlos porque não temos população em risco no início do estudo, excepto em estudos de casos-controlos tipo populacionais (a população base é equivalente à população em risco). Assim, a fórmula aplicada para o cálculo do RR num estudo de coorte não pode ser aplicada aos dados de um estudo de casos e controlos. O risco relativo pode, no entanto, ser estimado pelo cálculo de uma razão entre a odds da exposição entre os casos e o odds da exposição entre os controlos:

### **B - Odds ratio (OR) – estimativa do Risco Relativo**

O OR tanto pode ser calculado num estudo de coorte como num estudo de casos e controlos.

Um Odds Ratio vai ser a razão entre 2 Odds. Um odds (não confundir com odds ratio!) calcula-se da seguinte maneira:

$$\text{Odds} = \frac{\text{probabilidade de um acontecimento ocorrer}}{\text{probabilidade de não ocorrer}}$$

Como num estudo de coorte nós partimos da exposição vendo depois quantos é que desenvolvem a doença tanto nos expostos como nos não expostos, o OR vai ser a razão entre o odds de desenvolver a doença nos expostos e o odds de desenvolver a doença nos não expostos.

$$\text{Odds ratio}_{\text{coorte}} = \frac{\text{odds de desenvolver a doença nos expostos}}{\text{odds de desenvolver a doença nos não expostos}}$$

Utilizando a nomenclatura das tabelas de contingência, podemos ver como calcular esses Odds:

$$\text{Odds}_{\text{doença expostos}} = \frac{\frac{a}{a+b}}{\frac{b}{a+b}} = \frac{a}{b} \qquad \text{Odds}_{\text{doença não expostos}} = \frac{\frac{c}{c+d}}{\frac{d}{c+d}} = \frac{c}{d}$$

$$\text{Assim, OR}_{\text{coorte}} = \frac{\frac{a}{b}}{\frac{c}{d}} = \frac{\mathbf{a \times d}}{\mathbf{b \times c}}$$

Já num **estudo de casos e controlos** o nosso ponto de partida é a doença; assim o OR vai ser a razão entre os Odds de casos nos expostos e os Odds de controlos nos expostos.

$$\text{Odds ratio}_{\text{casos e controlos}} = \frac{\text{Odds da exposição dos casos}}{\text{Odds da exposição dos controlos}}$$

$$\text{Odds}_{\text{exposição casos}} = \frac{\frac{a}{a+c}}{\frac{c}{a+c}} = \frac{a}{c}$$

$$\text{Odds}_{\text{exposição controlos}} = \frac{\frac{b}{b+d}}{\frac{d}{b+d}} = \frac{b}{d}$$

$$\text{OR}_{\text{casos e controlos}} = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a \times d}{b \times c}$$

Ou seja, no final, tanto num estudo de coorte, como num estudo de casos e controlos obtemos a mesma equação. Concluimos assim que, para calcular um Odds Ratio basta utilizar a fórmula:

$$\text{OR} = \frac{a \times d}{b \times c}$$

o que, numa tabela de contingência, equivale a fazer a razão entre o produto dos cruzamentos.

Note-se porém que apenas costumamos calcular o OR num estudo de casos e controlos, visto que num estudo de coorte temos a possibilidade de calcular o Risco relativo.

A interpretação do Odds Ratio é igual à interpretação do RR. Se a exposição não estiver relacionada com a doença,  $\text{OR} = 1$ ; se a exposição estiver relacionada positivamente com a doença,  $\text{OR} > 1$ ; se a exposição estiver negativamente relacionada com a doença,  $\text{OR} < 1$ .

Já referimos que o OR obtido num estudo de casos e controlos é uma estimativa do RR.

**Mas quando é que esta estimativa representa uma boa aproximação?**

Quando se verificam as seguintes condições:

- Os casos estudados são representativos no que diz respeito à exposição comparativamente a todos os casos de doença da população da qual os casos foram retirados;
- Os controlos estudados são representativos no que diz respeito à exposição comparativamente a todas as pessoas sem doença da população da qual os casos foram retirados;

- A doença a ser estudada não ocorre frequentemente. Isto porque há “ $a+b$ ” pessoas expostas. Como a maior parte das doenças com as quais estamos a lidar ocorrem pouco frequentemente, muito poucas pessoas na população exposta vão desenvolver de facto a doença; conseqüentemente, “ $a$ ” é muito pequeno comparado com “ $b$ ” e podemos aproximar “ $a+b$ ” a “ $b$ ” ( $a+b \approx b$ ). Similarmente, muito poucas pessoas não expostas ( $c+d$ ) desenvolvem a doença, e podemos aproximar “ $c+d$ ” a “ $d$ ” ( $c+d \approx d$ ). Assim podemos calcular o RR da seguinte maneira:

$$\frac{\left(\frac{a}{a+b}\right)}{\left(\frac{c}{c+d}\right)} = \frac{\left(\frac{a}{b}\right)}{\left(\frac{c}{d}\right)}$$

Ao resolver esta equação obtemos:

$$\frac{a \times d}{b \times c}$$

que é o OR. Assim, quando a ocorrência da doença é pouco frequente. O OR é uma aproximação muito boa do RR.

### Resolução dos exercícios:

1)

a)

Estudo de coorte prospectivo.

b)

	<b>Doentes</b>	<b>Não doentes</b>	<b>Total</b>
<b>Expostos</b>	299	4503	<b>4802</b>
<b>Não expostos</b>	107	2312	<b>2419</b>
<b>Total</b>	<b>406</b>	<b>6815</b>	<b>7221</b>

c)

A melhor medida de associação, visto que se trata de um estudo de coorte, é o risco relativo.

**Incidência nos expostos (E):**

$$I_E = 299 / 4802 = 6,2 \%$$

**Incidência nos Não Expostos (NE):**

$$I_{NE} = 107 / 2419 = 4,4\%$$

**Risco Relativo:**

$$RR = 0,062 / 0,044 = 1,41$$

$RR^{-1} = 41\% \rightarrow$  ou seja, o risco aumenta 41% nos expostos.

d)

*Odds ratio*, mas num estudo de coorte não faz sentido visto que temos uma medida melhor, o risco relativo. Só se costuma calcular o *Odds ratio* num estudo de casos e controlos.

2)

a)

Estudo transversal. (Nota: o desenho não é casos e controlos mas vamos analisá-lo como tal)

b)

	Hipertenso	Não Hipertenso	Total
Café	250	550	800
Não Café	350	450	800
Total	600	1000	1600

c)

$$\begin{aligned} \text{Odds ratio} &= (250 \times 450) / (350 \times 550) = \\ &= 112500 / 192500 = 0,58 = 58\% \end{aligned}$$

Segundo este resultado, os consumidores de café têm aproximadamente um risco 40% inferior (100%-58%) de ter Hipertensão do que os não consumidores. Isto poder-nos-ia indicar que o consumo de café é protector relativamente à hipertensão.

$$R_{\text{CAFÉ}} = 250 / 600 = 0,31 = 31\%$$

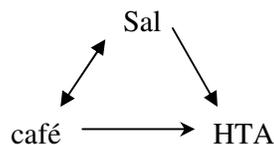
$$R_{\text{NÃO CAFÉ}} = 350 / 600 = 0,44 = 44\%$$

d)

É uma situação em que a exposição se pode alterar devido à consequência (quem é hipertenso deixa de beber café) – **erro de causalidade inversa**.

Podem também existir factores confundidores que influenciem esta relação.

ex;



O sal é um factor confundidor.

## 3) Atenção que esta tabela não está como estamos habituada a vê-la!

	<b>Casos (+)</b>	<b>Controlos (-)</b>
<b>Não bebedores (-)</b>	<b>c</b>	<b>d</b>
<b>Bebedores (+)</b>	<b>a</b>	<b>b</b>

Normalmente temos os expostos em cima e os não expostos em baixo. Assim, para calcular o OR NÃO podemos cruzar os números de maneira igual ao que costumamos fazer. Utilizar a seguinte fórmula:

$$\frac{a \times d}{c \times b} \rightarrow \begin{array}{l} \text{Números que favorecem a associação (doença - exposição)} \\ \text{Números que vão contra a associação} \end{array}$$

a)

$$\begin{aligned} \text{Odds ratio} &= (245 / 53) / (244 / 66) = \\ &= (245 \times 66) / (53 \times 244) = 1,25 \end{aligned}$$

b)

(0 – 15 gramas por dia)

	<b>Doente</b>	<b>Não Doente</b>
<b>Exposto</b>	45	86
<b>Não Exposto</b>	53	66

$$\text{OR}_{(0-15)} = (45 \times 66) / (53 \times 86) = 0,65$$

(16 – 30 gramas por dia)

	<b>Doente</b>	<b>Não Doente</b>
<b>Exposto</b>	70	76
<b>Não Exposto</b>	53	66

$$\text{OR}_{(16-30)} = (70 \times 66) / (53 \times 76) = 1,15$$

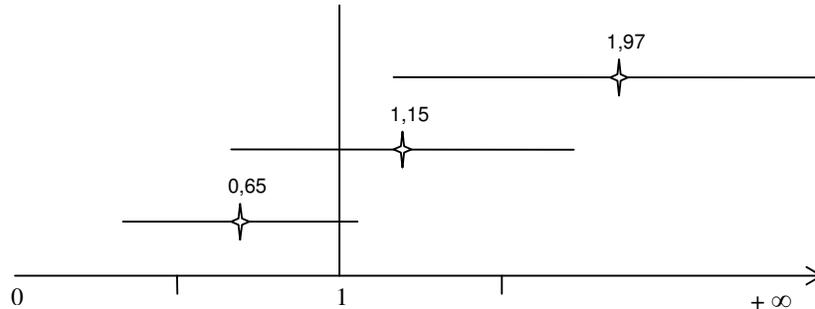
	<b>Doente</b>	<b>Não Doente</b>
<b>Exposto</b>	130	82
<b>Não Exposto</b>	53	66

$$\text{OR}_{(>30)} = (130 \times 66) / (53 \times 82) = 1,97$$

c)

Para as classes de consumo  $< 30\text{g}$ , o **IC 95%** inclui a **unidade**. Logo, estes dados não se revestem de significância estatística.

Apenas para consumos  $> 30\text{g/dia}$  é que se pode afirmar uma associação positiva entre o consumo e a doença (risco aumentado).



4)

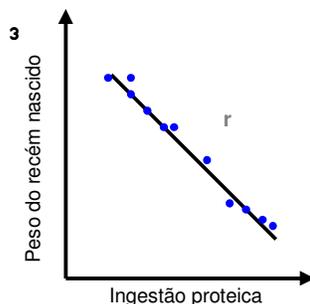
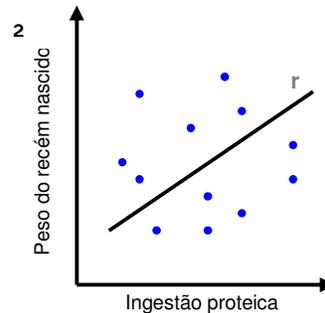
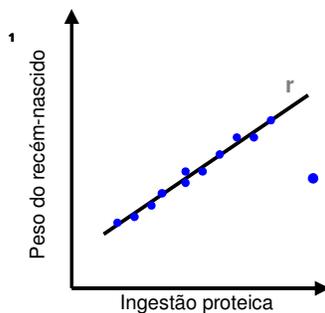
Ambas as variáveis são contínuas. Poderíamos categorizá-las mas há maneiras de as relacionar sem ser necessária a categorização. Podemos calcular o coeficiente de correlação ou o coeficiente de regressão, mas o que nos permite quantificar a relação é o coeficiente de regressão.

#### Coeficiente de correlação

A correlação é usada para medir a “força” da relação linear entre duas variáveis. Este coeficiente é normalmente representado pela letra  $r$ .

O coeficiente de correlação varia entre  $-1$  e  $1$ . O valor  $0$  (zero) significa que não há relação **linear**, o valor  $1$  indica uma relação linear perfeita e o valor  $-1$  também indica uma relação linear perfeita mas inversa, ou seja, quando uma das variáveis aumenta a outra diminui. Quanto mais próximo estiver de  $1$  ou  $-1$ , mais forte é a associação linear entre as duas variáveis.

Nota: uma correlação  $0$  ou próxima de  $0$  não implica obrigatoriamente que as duas variáveis não estão relacionadas mas apenas que as duas variáveis não estão relacionadas de uma forma linear.



- 1 – correlação próxima de 1
- 2 – correlação próxima de 0
- 3 – correlação próxima de -1

**Coeficiente de regressão:**

Esta medida é também uma medida de associação. Enquanto que a correlação é usada para medir a força da relação linear entre duas variáveis, a regressão linear é usada para estudar a **natureza** dessa relação. Ao contrário da correlação, é necessário distinguir qual a variável que se tenta prever (variável dependente) e a variável que prevê (variável independente).

Graficamente o modelo de regressão linear é apresentado como a recta que melhor aproxima a relação entre a variável dependente e a variável independente. Esta é a mesma recta representada no coeficiente de correlação; no entanto, o coeficiente de regressão permite-nos construir a recta; a representação matemática do modelo é a equação dessa recta:

$$y = a + bx$$

Ou seja, esta medida diz-nos quanto varia y por variação de x.

*Resumindo:*

Coeficiente de correlação: diz-nos se existe relação linear ou não.

Coeficiente de regressão: dá-nos a magnitude da relação entre as duas variáveis; diz-nos qual a quantidade de proteínas que temos de ingerir para fazer modificar o peso do recém-nascido.