

Epidemiologia – Soluções dos problemas

Epidemiologia Descritiva – Mod 2

1. a) Falso, é a subpopulação da área delimitada pelo estuário
b) Falso, é a população de aves na RLVT
c) Verdadeiro. Inferência: tirar conclusões sobre a população a partir da amostra. Pressupõe que a região delimitada é representativa de toda a RLVT, por exemplo porque tem aves de todas as espécies e idades etc de toda a RLVT.
2. Incidência: 8 casos por 100 mil; Prevalência: 2 casos por 100 mil
3. a) 973.7 / 100 mil
b) 18.9 / 100 mil (subestimação)
c) 0.103 mortes por doente
d) 2228.6 novos casos
e) 10.8 meses
f) A incidência e o número de indivíduos em tratamento (prevalência). As mortes são fiáveis.
4. a) LVT: 10.89 por mil; Alentejo: 15.14 por mil
b) Taxas padronizadas pelo método directo, LVT: 11.24, Alentejo: 12.48
c) A diferença de mortalidade na LVT e no Alentejo não é tão grande como a alínea a) sugere. Pode-se verificar que a proporção de idosos no Alentejo é maior que na LVT.
d) LVT: 2.5 por mil, Alentejo: 2.9 por mil
e) LVT: 3.03 por mil, Alentejo: 2.6 por mil
No Alentejo a mortalidade infantil é maior (pior assistência hospitalar e no parto ?) mas morre-se menos passada a fase crítica dos 0-4 anos. Os >69 anos também morrem mais no Alentejo que na LVT (respectivamente, 60.36 e 51.43 por mil).
5. e) O conceito de risco de ser infectado está associado a incidência, mas os dados fornecidos são de prevalência.
6. b) Confunde-se mortalidade com prevalência. Nada é dito no enunciado sobre o tempo de duração da pneumonia até à morte, em cada cidade.

RISCO – Mod 3

1. a) $r = 10/40 = 0.25$, IC: [0.1158, 0.3842]

b) $r = 8/18 = 0.44$,

c) RR = 4.89, OR = 8

d) IC para RR: [1.18, 20.20]

IC para OR: [1.425, 44.92]

e) Existe associação estatisticamente significativa pois, com 95% de confiança, os IC's estão inteiramente colocados à direita de 1. Esta associação não implica necessariamente relação causa-efeito entre água e doença.

2. a) $r = 58/7927 = 0.007 = 0.7\%$

b) $RR = (52/2872)/(6/5055) = 15.3$

c) RR = 0.07

erro-padrão do $\ln(RR) = 0.4305$, IC a 95%: [0.03, 0.163]. Ter tido varicela confere protecção contra vir a contrair doença benigna, uma vez que o IC não inclui 1.

OBSERVAÇÃO:

Note-se que se pode calcular o RR de quem teve varicela vir a desenvolver doença severa (RR=0), mas não se pode calcular o RR de quem não teve vir a desenvolver doença severa.

3. Estima-se que a prevalência é zero. Para representar a incerteza inerente a esta estimativa os ICs por aproximação da Normal não são apropriados. Usando os ICs de Agresti-Coull, conclui-se que o seguinte IC inclui a verdadeira prevalência com 95% de probabilidade: [0, 0.00384], ou seja, 0 a 0.384 por 1000 indivíduos.

4. a) $r = 220/425 = 0.52$

b) ovos e atum: $r = 60/75 = 0.8$

c) atum, $r = 70/200 = 0.35$.

5. a) Tabela da direita, onde a incidência é $(50+25)/200 = 37.5\%$,

b) RR = 2 nas duas tabelas

c) Na esquerda OR= 2.02, onde a doença é mais rara.

Na esquerda, $a/(a+b)$ é aproximadamente $=a/b = 0.02$ e $c/(c+d)$ é aproximadamente $= c/d = 0.01$.

Na direita os valores do risco são muito diferentes dos valores do odds.

6. a) Risco, RR e IC para cada tipo de consumo,

Nulo: $r = 0.088$, RR = 1. O IC é irrelevante, pois inclui obrigatoriamente 1. Esta é a exposição de referência, só os outros ICs têm interesse.

Ocasional: $r = 142/1845 = 0.077$; RR=0.077/0.088 = 0.875; IC: [0.627, 1.22]

Leve: $r = 0.056$, RR=0.639, IC: [0.458, 0.891]

Moderado: $r = 0.057$, RR=0.646, IC: [0.459, 0.909]

Alto: $r = 0.075$, RR=0.847, IC: [0.58, 1.236]

b) o RR de morrer é menor em níveis intermédios de consumo

c) Nada é dito sobre o passado de consumo destes indivíduos. Além disso, os indivíduos com consumo nulo-ocasional devem incluir pessoas que, por razões de saúde, não podem consumir álcool. A existência de outras doenças que se distribuem não uniformemente pelas classes de consumo, com maior predominância nos abstémios, pode funcionar como confundimento e explicar os resultados.

d) Tentar estender o estudo a "toda uma vida" dos inquiridos, excluindo indivíduos que tiveram um passado muito diferente da situação actual. Tentar que outras variáveis que se sabe aumentarem o risco de morrer (tabagismo, consumo de gorduras, idade, outras doenças...) se distribuam de forma uniforme pelas categorias de consumo.

7. a) RR=1,97

b) RRA = 0,093

Por cada 100 doentes que fizessem rivaroxaban em vez de enoxaparina, evitar-se-iam 9,3 eventos trombóticos.

c) NNT = 1/RRA = 10,75

d) RRA = 0,043; NNT = 23,26

Por cada 100 mudanças para rivaroxaban podemos esperar: menos 9,3 trombozes, mais 4,3 transfusões.

8. a) $r_{total} = 0.1134$, $r_{fumadores} = 0.1237$, $r_{n\grave{a}o-fumadores} = 0.0888$, $RR_{fumadores} = 1.393$

b) RA = 0.217. Significa que cerca de 22% da incidência de doença respiratória em toda a população é atribuível ao tabaco.

c) IC: [0.087, 0.445]

d) RA = 0.217

e) A incidência da doença em 10 anos deve ser 567000 novos casos (=0.1134 x 5 milhões). Se se anulasse totalmente o tabagismo, 22% destes casos (= 124740 casos) desapareceriam. Se apenas 10% do tabagismo é anulado, eliminam-se apenas 12474 casos em 10 anos e, portanto, esperar-se-iam cerca de menos 1247 casos de doença por ano.

9. a) A visualização deste problema em termos de tabelas de contingência é esta:

cancro				coronárias		
	morte	não morte	total	morte	não morte	total
fumador	140		100000	fumador	669	100000
não fumador	10		100000	não fumador	413	100000

Note-se que não existe uma amostra $n=200\ 000$ a partir da qual se construiu cada tabela. Os 100000 de cada linha não são independentes e não podem ser somados. Pode-se, contudo, calcular risco de morte ao longo da horizontal. Pode-se também somar as mortes por 100mil. Por exemplo, o risco total de morte por cancro é $(140+10)/100000$.

Os calculos do RR de morte dos fumadores conduzem a: Cancro: $RR=14$, Coronária: $RR=1.62$

b) Usar a fórmula mais geral: $RA = (\text{risco total} - \text{risco em não-expostos})/\text{risco total}$. Aplicado a este contexto, tem-se:

Cancro: risco total = $150/100^5$, risco em não expostos = $10/100^5$; $RA = (150-10)/150 = 0.933$

Coronária: risco total = $1082/100^5$, risco em não expostos = $413/100^5$;

$RA = (1082-413)/1082 = 0.618$

c) Embora o RR e o RA de cancro sejam mais elevados, o número de mortes evitadas por doença coronária seria muito mais alto, pois esta doença é mto mais prevalente.

Mortes evitáveis por cancro: 140/100 mil

Mortes evitáveis por doença coronária: 669/100 mil

10. a) A tabela é rearranjada ignorando as idades:

	Sero +	Sero -	total
1969	163	777	940
1989	330	670	1000
	493	1447	1940

risco em 1969 = 0.1734, risco em 1989 = 0.33, risco global = 0.2541

b) $RR = 1.9$; $OR = (330/670)/(163/777) = 2.35$

c) Sim. Porque em DSTs, a probabilidade de infecção aumenta com o número de encontros sexuais e, portanto, com a idade (idade está associada à infecção). Além disso, a tabela indica que a idade média de 1ª gravidez aumentou de 1969 para 1989, logo a idade não se distribui da mesma forma nas duas categorias do factor de risco sob análise. Assim, a aparente maior prevalência em 1989 pode ser provocada por se tratarem de pessoas mais velhas.

d) <26: $RR=1.91$, 26-30: $RR=1.72$, 31-35: $RR=1.9$, >35: $RR=1.36$

e) RR médio = 1.72 não é adequado, pois o número de mulheres difere muito entre idades

f) A resposta pode ser dada usando o RR global de 1.9 (da alínea b)). A idade das mães é irrelevante para esta pergunta, não sendo de confundimento neste contexto.

Caso-Controllo Mód 4

1. OR = 0.67 ratio de usar, relativamente a não usar. IC a 95% : [0.52, 0.88]
As luvas conferem efeito protector.

2. Resultados da análise de risco para cada item

Item	Casos (n=19)	Controlos (n=17)	OR	erro-pad	LI	LS
Galinha panada	14	11	1,53	0,73	0,37	6,35
Outra galinha	16	16	0,33	1,21	0,03	3,55
Omoleta	14	3	13,07	0,82	2,61	65,48
Arroz frito	14	9	2,49	0,71	0,62	10,06

A omeleta é a principal suspeita.

3. OR para ser de etnia X relativamente a não ser de etnia X, erro padrão e limites de confiança por etnia:

Etnia	Casos	Controlos	OR	erro	LI	LS
Caucasianas	514	541	0.93	0.162	0.68	1.28
Afro-americanas	25	25	1.04	0.289	0.59	1.84
Hispânicas	13	5	2.75	0.529	0.97	7.75
Asiáticas	32	21	1.62	0.287	0.92	2.84
Outras	20	37	0.55	0.284	0.31	0.96
	604	629				

a) Existe associação porque saber qual a etnia pode dar informação sobre a probabilidade de infecção. Neste caso, a associação é causada pelas "outras", que aparentam estar protegidas contra *E coli*.

b) A etnia que sirva de base de comparação é arbitrária. Do ponto de vista estatístico, é recomendável usar como base de comparação o grupo com a amostra maior, neste caso as caucasianas:

Etnia	Casos	Controlos	OR	erro	LI	LS
Caucasianas	514	541	1.00			
Afro-americanas	25	25	1.05	0.289	0.60	1.86
Hispânicas	13	5	2.74	0.530	0.97	7.73
Asiáticas	32	21	1.60	0.288	0.91	2.82
Outras	20	37	0.57	0.284	0.33	0.99
	604	629				

O resultado qualitativo é o mesmo: as 'outras' parecem estar protegidas.

4. a) OR = 1.21, IC : [1.012, 1.447]

Deve ser uma boa aproximação pois a doença é rara: $1021/75000 = 1.4\%$

b) Assumindo que: (1) OR é uma boa aproximação ao RR e (2) A proporção de expostos ao factor de risco nos controlos, $b/(b+d)$, é uma boa estimativa dessa proporção na população. (Parece verdade, pois quase todas as grávidas vão à maternidade).

Então: $RA = (ad-bc)/[d(a+c)] = 0.1098$. Cerca de 11% dos casos são atribuíveis à rubéola.

c) IC: [0.044, 0.25]

5. Será difícil conseguir que os controlos informem telefonicamente acerca de factores de risco socialmente sensíveis, como o uso de drogas intravenosas ou comportamento sexual de risco, com o mesmo grau de honestidade que os infectados. O resultado vai ser que a exposição a factores de risco terá prevalência muito mais elevada nos infectados, comparativamente aos controlos, apenas devido ao enviesamento nas respostas.

Coortes Mód 5

1. a) 157,9 pessoa-mês
 b) TI = 0.0697 casos por criança-mês; 0.836 casos por criança-ano
 c) em 1 mês: $r = 0.067$, em 6 meses: 0.342, pressuposto: a TI permanece constante
 d) semi-parábola que tende para uma assintota em 1
2. TI = 0.5195 casos por dia em 80 pessoas
 $r = 1 - \exp(-0.0065 \cdot 2) = 0.013$ em dois dias
3. a) IC por 1000 pessoas-tempo pela aproximação à Normal: [3.04, 7.57]
 O mesmo IC pela fórmula de Byar: [3.38, 7.95]
 b) RR=3.038. O IC é: [1.043, 8.85], não inclui 1 sugerindo que a vacina confere protecção.
 c) A eficácia reporta-se a 1 ano lectivo e é a proporção de vacinados que não contraiu doença: 0.9475, ou seja ~95%, uma eficácia habitual em vacinas da infância.
 d) Não porque quando regressam já não estão em risco de contrair doença com a mesma probabilidade q tinham antes de contrair.
4. a) Estudo de coortes prospectivo. Não é caso-controlo porque o critério de formação dos dois grupos é o factor de risco (por oposição a ser/não-ser doente). À partida não há ainda “doentes” no sentido usado no estudo (seroconversão HIV).
 b) Todos deviam ter sido testados para excluir a possibilidade de já serem HIV+ antes do estudo iniciar.
 c) O problema pode ser tratado como uma coorte fixa:

	sero +	sero -	total	risco	erro	LI	LS
Úlceras genitais	21	128	149	0.14	0.03	0.09	0.20
Outros problemas	3	141	144	0.02	0.01	-0.002	0.04
	24	269	293	0.08	0.02	0.05	0.11
				RR	erro	LI	LS
				6.77	0.61	2.06	22.19

As úlceras estão estatisticamente associadas à seroconversão para HIV

5. a) Estudo de coortes em que se segue um grupo exposto e se toma a população geral como controlo.
 b) RR = 1.026, IC: [0.9044, 1.163]. O RR parece muito baixo: os dados disponíveis não permitem rejeitar a não-associação.
 c) - São homens relativamente saudáveis q tendem a estar empregados nestas fabricas, enquanto na população geral há de tudo. Não se usaram variáveis de confundimento (idade, tabaco, alimentação) para comparar c/ a população.
 - Não se considerou a dose de exposição dentro da fábrica
 - Só se usaram mortes, a exposição pode causar doença sem morte
 - Pode haver abandonos do estudo pela população seguida, p.e. devido a problemas de saúde ocorridos antes da morte, coisa que não existe na população “controlo”.

6.

Nível de toma	risco	RR	erro Ln(RR)	LI	LS
1	0,0041	1			
2	0,0038	0,930	0,081	0,79	1,09
3	0,0039	0,963	0,123	0,76	1,23
4	0,0032	0,793	0,142	0,60	1,05
5	0,0031	0,764	0,209	0,51	1,15

Tomados individualmente, o efeito de cada dose não é significativo, mas a toma de vitamina C diminui o risco de forma sistemática. Parece haver um efeito dose-resposta consistente.

7. Apenas a a) não é uma vantagem

8. Os problemas são: a, e, f

Ensaio Clínico Mód 17

1. a) É ensaio clínico, paralelo, aleatorizado e com controlo (não-cego)

b) Idealmente, a comparação entre proporções é feita criando uma TC 2x2 e usando as medias de associação e teste conhecidas. Começando pelo RR...

RR = 2.605, IC: [0.954, 7.118] Indica que o risco de não-reversão é 2.6 vezes maior nos que não mudaram de estilo de vida. O IC não permite rejeitar a H_0 , mas está no borderline.

Uma vez que a amostra não é grande, podia-se tentar também usar o teste exacto de Fisher (com 1 cauda):

H_0 : Não há diferença entre expostos e não expostos qto a contrair doença

H_1 : Os expostos têm maior probabilidade (daí ser 1 cauda) de contrair doença

Usando o OpenEpi, o teste de Fisher com 1 cauda dá $p=0.047$, rejeitando a H_0 também sobre o borderline.

Seria recomendável aumentar o tamanho da amostra para conseguir resultados mais convincentes.

c) As pessoas do grupo de controlo podem espontaneamente mudar de estilo de vida, assustadas com o seu estado.

A amostra é pequena, atendendo a que os efeitos do tratamento não são muito fortes. Além disso, tratados e não-tratados podem diferir muito qto a variáveis de confundimento apenas porque a amostra é pequena.

2. a) Paralelo = Formam-se grupos e cada indivíduo só recebe o tratamento do grupo onde foi colocado.

Aleatorizado = Distribuição aleatória de indivíduos pelos grupos

Controlado = Há um grupo de controlo

Cego = Os pacientes não sabiam que tratamento receberam (mas quem aplicou os tratamentos sabia)

b) Vai-se comparar diferenças médias entre os dois grupos. Com dados em escala contínua, assumindo que as amostras são independentes uma da outra, parece apropriado usar o t-student (T):

$$T = (X_1 - X_2) / s_d$$

X_1 = média no grupo 1; X_2 = média no grupo 2

S_d = erro-padrão da diferença entre as médias

T tem distrib t-student com 14 gdl

T = 1.776, o valor tabelado para $T_{14 \text{ gdl}, 0.05} = 2.14$

3. a) Estudo cruzado, se possível duplamente cego. Evita a grande variabilidade intra-tratamentos que é expectável. Possível por ser doença com características crónicas.

b) Estudo paralelo. Atribuição sequencial aleatória de tratamentos. Não necessita ser cego.

c) Estudo cruzado devido à presumível grande variabilidade intra-tratamentos. Grande risco de "efeito de transporte", pelo que o estudo se pode estender por muito tempo.

- d) Estudo paralelo, com atribuição aleatória sequencial de tratamentos. Em princípio, não necessita ser cego.
 e) Estudo paralelo com dois grandes grupos. Ensaio cego ou duplamente cego.

4. a) Estudo cruzado, com reposta binária, analisável pelo método de Gart. Qui-quadrado=6,349 rejeita H_0 , logo os dados fornecem evidência de que as drogas produzem efeito diferente.

b) Qui-quadrado=0,4103. Não existe evidência de que a ordem por que são tomadas faça diferença.

5. Formar 4 estratos (dois sexos x duas Idades) e aleatorizar *dentro* de cada estrato

Testes de diagnóstico Mód 6

1. a) $s=60\%$, $p=86,15\%$

b) Há muitos doentes infectados que não são detectados pelo teste (baixo s). A ausência de cheiro pútrido, porém, é um indicador razoável de ausência de infecção (alto p). Parece aceitável não receitar antibiótico aos que são negativos ao cheiro. O valor preditivo dos negativos é 77,8%.

2. a) 999; b) 8955; c) 45; d) 96%;

e) 17%. Apesar dos elevados valores de s e p , e apesar do número de falsos positivos não ter variado muito entre as duas populações, o teste é praticamente inútil num ambiente de tão baixa prevalência.

3. a) $RR= 3,281$, $IC:[1,266; 8,505]$, existe associação estatisticamente significativa

b) Se a sensibilidade fosse total, deveriam ter sido detectados 17 casos positivos. Distribuindo aleatoriamente o adicional de indivíduos por expostos e não-expostos, obtem-se $RR= 3,33$, $IC:[1,374; 8,086]$. A associação aumentou e a precisão também. Um dos efeitos habituais de erros aleatórios na classificação feita pelo teste é reduzir a força de associação do estudo quando esta existe.

4. a) 27%

b) 69%; O valor preditivo de uma doença pouco prevalente pode ser muito melhorado se a especificidade do teste aumentar. Manipular a sensibilidade, se fosse possível, pouco alteraria.

c) Valor preditivo (com $p=70\%$) = 89%; Valor preditivo (com $p=95\%$) = 98%. Aumentar a especificidade numa doença muito prevalente não altera tanto o valor preditivo.

5. a) $s=100\%$; $p=0\%$

b) cut-off em 5: $s=97,14\%$; $p=24,95\%$; cut-off em 10: $s=94,84\%$; $p=42,85\%$

e) 40

f) w deve ser menor

g) O número de resultados correctos do teste é maximizado colocando w =prevalência.

Regressão Logística Mód 18

1. a)

Idade	CHD=1	CHD=0	
>55	21	6	27
≤ 55	22	51	73
	43	57	

- b) $\ln(p/(1-p)) = -0.841 + 2.094 \text{ Idade}$
- c) Na tabela: odds = 21/6=3.5; risco = 21/27 = 0.778
No modelo: Logit = -0.841 + 2.094 (1)= 1.253; odds = exp(1.253) = 3.5
Risco = 1/(1 + exp(-1.253)) = 0.778
- d) No modelo: OR= exp(2.09(1-0)) = 8.12
Na tabela: (21/6)/(22/51) = 8.11
- e) No modelo: RR=0.7777/0.3013 = 2.58
Na tabela: RR = (21/27)/(22/73) = 2.58
- f) $\text{Exp}(2.094(1-0) \pm 1.96(0.529)(1-0))$
donde IC:[2.88, 22.9], é estatisticamente significativo
- g) $W = 2.094/0.529 = 3.96$, é significativo (superior ao valor tabelado 1.96, numa Normal para $\alpha=0.05$)

2. a) $\text{Logit} = b_0 + b_1^{(1)}\text{estirpe1} + b_1^{(2)}\text{estirpe2} + b_1^{(3)}\text{estirpe3} + b_1^{(4)}\text{estirpe4}$, ou ainda,
= -1.386 + 2.079 estirpe1 + 1.792 estirpe2 + 1.386 estirpe3 + 0 estirpe4

b) $\text{Odds}_{\text{est1}} = \exp(-1.386 + 2.079(1)) = 2$; $\text{Risco}_{\text{est1}} = 1/(1+(1/2)) = 0.666$
 $\text{Odds}_{\text{est4}} = \exp(-1.386 + 0) = 0.25$; $\text{Risco}_{\text{est4}} = 1/(1+(1/0.25)) = 0.2$

- c) $\text{OR}_{\text{est1}} = \exp(2.079-0) = 8$; $\text{OR}_{\text{est4}} = \exp(0) = 1$
- d) $\text{OR}(\text{est1}/\text{est2}) = \exp(2.079-1.792) = 1.332$
- e) $\text{Exp}(2.079 \pm 1.96(0.633))$, donde IC:[2.3, 27.6]

3. a) $\text{Logit} = -4.5995 + 0 \text{ SBP1} + 0.6092 \text{ SBP2} + 0.8697 \text{ SBP3} + 1.0297 \text{ SBP4} +$
 $+ 1.3425 \text{ SBP5} + 0 \text{ COL1} + 0.2089 \text{ COL2} + 0.8229 \text{ COL3} + 1.0066 \text{ COL4}$
 $+ 1.2957 \text{ COL5}$.

- b) Odds=0.1407; risco=0.1233
- c) OR=13.99; RR=12.38
- d) graus de liberdade (de cima para baixo): 24, 20, 20, 16
- e) Deviance=18.86 (para 16 gdl), $p=0.276$, Não se rejeita H_0 , o modelo está bem ajustado.
- f) $\Delta \text{Deviance}=28.1$, para 4 gdl, $p<0.0001$. Altamente significativo.
- g) $\Delta \text{Deviance}=35.35$, para 4 gdl, $p<0.0001$. Altamente significativo.
- h) $\Delta \text{Deviance}=65.97$, para 8 gdl, $p<0.0001$. Altamente significativo.
- i) $\Delta \text{Deviance}=37.87$, para 4 gdl, $p<0.0001$. Altamente significativo.

- 4. a) Quantidade de variação (aumento) do log(odds) quando a idade varia 1 ano, ajustado para o confundimento das outras variáveis
- b) odds= 0.03; risco=0.029
- c) OR=2,5
- d) Não, depende também da variabilidade da variável
- e) coeficientes padronizados: idade=0.26; colesterol=0.37. O colesterol parece mais importante.
- f) $\exp(0.0171)=1.02$

- 5.
- a) $\text{Logit} = -1.18 - 0.52\text{SOC} + 0.04\text{SBP} - 0.56\text{SMK} - 0.033(\text{SOC} \times \text{SBP}) + 0.175(\text{SOC} \times \text{SMK})$
- b) risco=0.262
- c) Pessoa 1: 0.352; Pessoa 2: 0.473
- d) Modelo 1 $r=0.262$; Modelo 2 $r=0.352$. A presença da interação explica esta diferença apesar de ser a mesma pessoa.
- e) RR=0.744, o nível social confere protecção

- f) Procuravam-se mortes por CVD e investigava-se o seu passado em vivos. Procuravam-se controlos vivos com a mesma idade dos mortos e questionavam-se sobre as mesmas variáveis. Não se pode calcular riscos, apenas odds e OR.
- g) $OR=0.6065 = 1/1.65$. O risco de morte em pessoas $SOC=1$ é inferior ao de pessoas com $SOC=0$.
- h) $\exp(-0,520 - 0,033 SBP + 0,175 SMK)$

6. a) V
 b) F; b_1 controla para efeitos de confundimento das outras variáveis sobre X_1
 c) V
 d) F; o OR é $\exp(b)$, sendo b o coeficiente da variável binária
 e) F; O OR é $\exp(b_{tab} + b_{tab \times idade} Idade)$
 f) V

7.
 a) Logit =
 $= b_0 + b_1 PER + b_2 PAR + b_3 PS + b_4 NMCS + b_5 (PER \times PAR) + b_6 (PER \times PS) + b_7 (PER \times NMCS)$

- b) $OR = \exp(b_1 + b_5 PAR + b_6 PS + b_7 NMCS)$

8.
 a) $\Delta Deviance = 43.4$, com 1 gdl, $p < 0.0001$; rejeita-se H_0 : a interacção $CAT \times CHL$ é significativa
 b) $Wald = 4.77$, $p < 0.0001$
 c) $OR = \exp(-14.0809 + 0.0683 CHL)$
 d) $CHL = 220$, dá $OR = 2.57$; $CHL = 240$, dá $OR = 10.09$

9.
 a) $Logit = b_0 + b_1 HT + b_2 HS + b_3 CT + b_4 AGE + b_5 SEX + b_6 (HT \times HS) + b_7 (HT \times CT)$

- b) $Logit =$
 $= b_0 + b_1 HT + b_2 HS + b_3 CT + b_4 AGE + b_5 SEX + b_6 (HT \times HS) + b_7 (HT \times CT) + b_8 (HT \times AGE) + b_9 (HT \times SEX) + b_{10} (AGE \times SEX) + b_{11} (HT \times AGE \times SEX)$

- c) Não é apropriado. Primeiro porque existe um termo de ordem mais elevada onde HS intervém em interacção. Se este termo for significativo, HS tem de se manter no modelo. Segundo, o efeito directo de HS é de confundimento e, para isso, não há testes estatísticos apropriados.

- d)
 1º- Testar as interacções de ordem mais elevada globalmente, com teste estatísticos
 2º- Se não eliminar globalmente essas interacções, testá-las separadamente, com testes estatísticos
 3º- Testar os efeitos directos de confundimento, por comparação dos OR estimados
 4º Se algum termo de interacção tiver permanecido no modelo, as variáveis correspondentes não são testadas para confundimento
 5º Se a avaliação do confundimento sugerir que alguma variável deve sair, avaliar a precisão do modelo com e sem a variável. Se a precisão baixar a variável não sai.

- e) São elegíveis todas excepto HT, HS e CT.