© drtamil@gmail.com 2020

Cochran-Mantel-Haenszel

FK6193

Dr Azmi Mohd Tamil

© drtamil@gmail.com 2020

Historically

- Pearson’s Chi-Square Test (1904)
- Likelihood Ratio Test
- Cochran Test (1954)
- Mantel-Haenszel Test (1959)
- Breslow-Day Test (1980)
- Tarone (1985)

© drtamil@gmail.com 2020

Mantel-Haenszel

- An excellent method for adjusting for confounding factors when analysing the relationship between a dichotomous risk factor and a dichotomous outcome.

© drtamil@gmail.com 2020

Example

- Those with high catecholamine are believed to be of high risk for coronary heart disease. However age & ECG changes are probable confounders.
  - RF - Catecholamine (Low / High)
  - Outcome - CHD (Present / Absent)
  - Confounders
    - Age (<55, 55+)
    - ECG (+, -)

© drtamil@gmail.com 2020

Combine All

	CHD +	CHD -	
High Cat	27(22.1%)	95	122
Low Cat	44(9.0%)	443	487
	71	538	609

Crude OR = 2.86, X^2=16.25

© drtamil@gmail.com 2020

Stratification

- To control confounding factors, we divide the sample into a series of strata, which are now internally homogenous with regards to the confounding factors.
- The odds ratio calculated within each stratum are free of bias arising from confounding.

© drtamil@gmail.com 2020

Age < 55, ECG -

	CHD +	CHD -	
High Cat	1(12.5%)	7	8
Low Cat	17(6.2%)	257	274
	18	264	282

OR = (1 x 257)/(7 x 17) = 2.16

© drtamil@gmail.com 2020

Age < 55, ECG +

	CHD +	CHD -	
High Cat	3(17.6%)	14	17
Low Cat	7(11.9%)	52	59
	10	66	76

OR = (3 x 52)/(14 x 7) = 1.59

© drtamil@gmail.com 2020

Age 55+, ECG -

	CHD +	CHD -	
High Cat	9(23.1%)	30	39
Low Cat	15(12.3%)	107	122
	24	137	161

OR = (9 x 107)/(30 x 15) = 2.14

© drtamil@gmail.com 2020

Age 55+, ECG +

	CHD +	CHD -	
High Cat	14(24.1%)	44	58
Low Cat	5(15.6%)	27	32
	19	71	90

OR = (14 x 27)/(44 x 5) = 1.72

© drtamil@gmail.com 2020

Odds Ratio

Stratum	Risk +	Risk -	OR
<55, ECG+	3(17.6%)	7(11.9%)	1.59
55+, ECG+	14(24.1%)	5(15.6%)	1.72
55+, ECG-	9(23.1%)	15(12.3%)	2.14
<55, ECG-	1(12.5%)	17(6.2%)	2.16
Combined	27(22.1%)	44(9.0%)	2.86

© drtamil@gmail.com 2020

$CI = OR.\exp \pm 1.96\sqrt{1/a+1/b+1/c+1/d}$

Stratum	OR	Lower	Higher
<55, ECG+	1.59	0.36	6.96
55+, ECG+	1.72	0.56	5.31
55+, ECG-	2.14	0.85	5.37
<55, ECG-	2.16	0.25	18.58
Combined	2.86	1.69	4.85

© drtamil@gmail.com 2020

Odds Ratio

Stratum	OR
<55, ECG+	1.59
55+, ECG+	1.72
55+, ECG-	2.14
<55, ECG-	2.16
Combined	2.86

- Despite stratification, stress constantly leads to higher odds (but not significant) of getting CHD.
- There seems to be little effect modification due to age and ECG. The odds are similar. But combined table stronger & highly significant OR=2.86;1.69<OR<4.85.
- Need an adjusted summary measure & adjust for effect of age & ECG.

© drtamil@gmail.com 2020

Chi-Square Cochran-Mantel-Haenszel

$$X^2_{MH} = \frac{\sum\left(\frac{ad-cb}{n}\right)^2}{\sum\left(\frac{(a+b)\times(c+d)\times(a+c)\times(b+d)}{(n-1)\times n^2}\right)}$$

© drtamil@gmail.com 2020

Testing for Overall Association

	D+	D-	
E+	a	b	a+b
E-	c	d	c+d
	a+c	b+d	n

$$X^2_{MH} = \frac{\sum\left(\frac{ad - cb}{n}\right)^2}{\sum\left(\frac{(a+b)\times(c+d)\times(a+c)\times(b+d)}{(n-1)\times n^2}\right)}$$

© drtamil@gmail.com 2020

Age < 55, ECG -

	CHD +	CHD -	
High Cat	1	7	8
Low Cat	17	257	274
	18	264	282

$$X^2_{MH} = \frac{\sum \left(\frac{ad - cb}{n}\right)^2}{\sum \left(\frac{(a+b)\times(c+d)\times(a+c)\times(b+d)}{(n-1)\times n^2}\right)}$$

$$\frac{\left[\frac{1\cdot 257 - 17\cdot 7}{282} + \frac{3\cdot 52 - 7\cdot 14}{76} + \frac{9\cdot 107 - 15\cdot 30}{161} + \frac{14\cdot 27 - 5\cdot 44}{90}\right]^2}{\left[\frac{8\cdot 274\cdot 18\cdot 264}{281\cdot 282^2} + \frac{17\cdot 59\cdot 10\cdot 66}{75\cdot 76^2} + \frac{39\cdot 122\cdot 24\cdot 137}{160\cdot 161^2} + \frac{58\cdot 32\cdot 19\cdot 71}{89\cdot 90^2}\right]}$$

© drtamil@gmail.com 2020

Age < 55, ECG +

	CHD +	CHD -	
High Cat	3	14	17
Low Cat	7	52	59
	10	66	76

$$X^2_{MH} = \frac{\sum \left(\frac{ad - cb}{n}\right)^2}{\sum \left(\frac{(a+b) \times (c+d) \times (a+c) \times (b+d)}{(n-1) \times n^2}\right)}$$

$$\frac{\left[\frac{1 \cdot 257 - 17 \cdot 7}{282} + \frac{3 \cdot 52 - 7 \cdot 14}{76} + \frac{9 \cdot 107 - 15 \cdot 30}{161} + \frac{14 \cdot 27 - 5 \cdot 44}{90}\right]^2}{\left[\frac{8 \cdot 274 \cdot 18 \cdot 264}{281 \cdot 282^2} + \frac{17 \cdot 59 \cdot 10 \cdot 66}{75 \cdot 76^2} + \frac{39 \cdot 122 \cdot 24 \cdot 137}{160 \cdot 161^2} + \frac{58 \cdot 32 \cdot 19 \cdot 71}{89 \cdot 90^2}\right]}$$

© drtamil@gmail.com 2020

Age 55+, ECG -

	CHD +	CHD -	
High Cat	9	30	39
Low Cat	15	107	122
	24	137	161

$$X^2_{MH} = \frac{\sum\left(\frac{ad - cb}{n}\right)^2}{\sum\left(\frac{(a+b)\times(c+d)\times(a+c)\times(b+d)}{(n-1)\times n^2}\right)}$$

$$\frac{\left[\frac{1\cdot257-17\cdot7}{282} + \frac{3\cdot52-7\cdot14}{76} + \frac{9\cdot107-15\cdot30}{161} + \frac{14\cdot27-5\cdot44}{90}\right]^2}{\left[\frac{8\cdot274\cdot18\cdot264}{281\cdot282^2} + \frac{17\cdot59\cdot10\cdot66}{75\cdot76^2} + \frac{39\cdot122\cdot24\cdot137}{160\cdot161^2} + \frac{58\cdot32\cdot19\cdot71}{89\cdot90^2}\right]}$$

© drtamil@gmail.com 2020

Age 55+, ECG +

	CHD +	CHD -	
High Cat	14	44	58
Low Cat	5	27	32
	19	71	90

$$X_{MH}^2 = \frac{\sum \left(\frac{ad - cb}{n} \right)^2}{\sum \left(\frac{(a+b) \times (c+d) \times (a+c) \times (b+d)}{(n-1) \times n^2} \right)}$$

$$\frac{\left[\frac{1 \cdot 257 - 17 \cdot 7}{282} + \frac{3 \cdot 52 - 7 \cdot 14}{76} + \frac{9 \cdot 107 - 15 \cdot 30}{161} + \frac{14 \cdot 27 - 5 \cdot 44}{90}\right]^2}{\left[\frac{8 \cdot 274 \cdot 18 \cdot 264}{281 \cdot 282^2} + \frac{17 \cdot 59 \cdot 10 \cdot 66}{75 \cdot 76^2} + \frac{39 \cdot 122 \cdot 24 \cdot 137}{160 \cdot 161^2} + \frac{58 \cdot 32 \cdot 19 \cdot 71}{89 \cdot 90^2}\right]}$$

© drtamil@gmail.com 2020

Example

$$\frac{\left[\frac{1\cdot 257-17\cdot 7}{282}+\frac{3\cdot 52-7\cdot 14}{76}+\frac{9\cdot 107-15\cdot 30}{161}+\frac{14\cdot 27-5\cdot 44}{90}\right]^2}{\left[\frac{8\cdot 274\cdot 18\cdot 264}{281\cdot 282^2}+\frac{17\cdot 59\cdot 10\cdot 66}{75\cdot 76^2}+\frac{39\cdot 122\cdot 24\cdot 137}{160\cdot 161^2}+\frac{58\cdot 32\cdot 19\cdot 71}{89\cdot 90^2}\right]}$$

$$=\frac{(.4894+.7632+3.1863+1.7556)^2}{.4661+1.5281+3.7721+3.4731}$$

$$=4.15(p=.04,\text{two-sided})$$

Table 3 : Percentage point of X^2

d.k.	P Value							
	0.5	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	1.39	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	2.37	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	3.36	5.39	7.78	9.49	11.14	13.28	14.86	18.47
5	4.35	[illegible]	9.24	11.07	12.83	15.09	[illegible]	20.52
6	5.35	7.84	10.64	[illegible]	[illegible]	16.81	18.55	22.46
7	6.35	9.04	12.02	14.07	16.01	18.48	20.28	24.32
8	7.34	10.22	13.36	15.51	[illegible]	20.09	21.96	26.13
9	8.34	11.39	14.68	16.92	[illegible]	21.67	23.59	27.88
10	9.34	12.55	15.99	18.31	[illegible]	[illegible]	25.19	29.59
11	10.34	13.70	17.28	19.68	21.92	[illegible]	26.76	31.26
12	11.34	14.85	18.55	21.03	23.34	[illegible]	28.30	32.91
13	12.34	15.98	19.81	22.36	24.74	[illegible]	29.69	34.53
14	13.34	17.12	21.06	23.68	26.12	[illegible]	31.32	36.12
15	14.34	18.25	22.31	25.00	27.49	[illegible]	32.80	37.70
16	15.34	19.37	23.54	26.30	28.85	[illegible]	[illegible]	39.25
17	16.34	20.49	24.77	27.59	30.19	[illegible]	[illegible]	40.79
18	17.34	21.60	25.99	28.87	31.53	[illegible]	[illegible]	42.31
19	18.34	22.72	27.20	30.14	32.85	36.19	[illegible]	43.82
20	19.34	23.83	28.41	31.41	34.17	37.57	[illegible]	45.32
21	20.34	24.93	29.62	32.67	35.48	38.93	[illegible]	46.80
22	21.34	26.04	30.81	33.92	36.78	40.29	[illegible]	[illegible]
23	22.34	27.14	32.01	35.17	38.08	41.64	[illegible]	[illegible]
24	23.34	28.24	33.20	36.42	39.36	42.98	[illegible]	[illegible]
25	24.34	29.34	34.38	37.65	40.65	44.31	46.93	[illegible]
26	25.34	30.43	35.56	38.89	41.92	45.64	48.29	[illegible]
27	26.34	31.53	36.74	40.11	43.19	46.96	49.64	[illegible]
28	27.34	32.62	37.92	41.34	44.46	48.28	50.99	[illegible]
29	28.34	33.71	39.09	42.56	45.72	49.59	52.34	[illegible]
30	29.34	34.80	40.26	43.77	46.98	50.89	53.67	[illegible]
40	39.34	45.62	51.81	55.76	59.34	63.69	66.77	[illegible]
50	49.33	56.33	63.17	67.50	71.42	76.15	79.49	[illegible]
60	59.33	66.98	74.40	79.08	83.30	88.38	91.95	[illegible]
70	69.33	77.58	85.53	90.53	95.02	100.43	104.22	112.32
80	79.33	88.13	96.58	101.88	106.63	112.33	116.32	124.84
90	89.33	98.65	107.57	113.15	118.14	124.12	128.30	137.21
100	99.33	109.14	118.50	124.34	129.56	135.81	140.17	149.45

Refer to Table 3.

Look at df = 1.

X^2_{MHtest} = 4.15, larger than 3.84 (p=0.05) but smaller than 5.02 (p=0.025).

5.02>4.15>3.84

Therefore if X^2_{MHtest}=4.15, 0.025<p<0.05.

d.k.	0.5	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	1.39	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	2.37	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	3.36	5.39	7.78	9.49	11.14	13.28	14.86	18.47

© drtamil@gmail.com 2020

Interpretation

- There is a significant relationship between CAT and CHD, adjusted simultaneously for age and ECG ($p < 0.05$; X^2_{MHtest}).

Important: In the numerator, sum before squaring.

Under the null hypothesis X^2_{MHtest} ~ Chi square (1 df)

© drtamil@gmail.com 2020

Mantel-Haenszel Adjusted Odds Ratio

$$\widehat{OR}_{MH} = \sum_{g=1}^{G} \frac{a_g \cdot d_g}{n_g} \Big/ \sum_{g=1}^{G} \frac{b_g \cdot c_g}{n_g}$$

© drtamil@gmail.com 2020

Estimating The Adjusted OR

Stratum	OR	Lower	Higher
<55, ECG+	1.59	0.36	6.96
55+, ECG+	1.72	0.56	5.31
55+, ECG-	2.14	0.85	5.37
<55, ECG-	2.16	0.25	18.58
Crude OR	2.86	1.69	4.85

© drtamil@gmail.com 2020

Mantel-Haenszel Estimator of Common Odds Ratio

$$\hat{\theta}_{MH} = \frac{\sum \left(\frac{ad}{n} \right)}{\sum \left(\frac{bc}{n} \right)}$$

$$\widehat{OR}_{MH} = \sum_{g=1}^{G} \frac{a_g \cdot d_g}{n_g} \Big/ \sum_{g=1}^{G} \frac{b_g \cdot c_g}{n_g}$$

© drtamil@gmail.com 2020

Common/Average Odds Ratio

	D+	D-	
E+	a	b	a+b
E-	c	d	c+d
	a+c	b+d	n

$$\widehat{OR}_{MH} = \sum_{g=1}^{G} \frac{a_g \cdot d_g}{n_g} \Big/ \sum_{g=1}^{G} \frac{b_g \cdot c_g}{n_g}$$

© drtamil@gmail.com 2020

Common/Average Odds Ratio

$$\widehat{OR}_{MH} = \sum_{g=1}^{G} \frac{a_g \cdot d_g}{n_g} \Big/ \sum_{g=1}^{G} \frac{b_g \cdot c_g}{n_g}$$

$$\widehat{OR}_{MH} = \frac{1 \cdot 257/282 + 3 \cdot 52/76 + 9 \cdot 107/161 + 14 \cdot 27/90}{17 \cdot 7/282 + 7 \cdot 14/76 + 15 \cdot 30/161 + 5 \cdot 44/90}$$

$$= 1.89$$

© drtamil@gmail.com 2020

Age < 55, ECG -

	CHD +	CHD -	
High Cat	1	7	8
Low Cat	17	257	274
	18	264	282

$$\widehat{OR}_{MH} = \frac{1 \cdot 257/282 + 3 \cdot 52/76 + 9 \cdot 107/161 + 14 \cdot 27/90}{17 \cdot 7/282 + 7 \cdot 14/76 + 15 \cdot 30/161 + 5 \cdot 44/90}$$

$$= 1.89$$

© drtamil@gmail.com 2020

Age < 55, ECG +

	CHD +	CHD -	
High Cat	3	14	17
Low Cat	7	52	59
	10	66	76

$$\widehat{OR}_{MH} = \frac{1 \cdot 257/282 + 3 \cdot 52/76 + 9 \cdot 107/161 + 14 \cdot 27/90}{17 \cdot 7/282 + 7 \cdot 14/76 + 15 \cdot 30/161 + 5 \cdot 44/90}$$

$$= 1.89$$

© drtamil@gmail.com 2020

Age 55+, ECG -

	CHD +	CHD -	
High Cat	9	30	39
Low Cat	15	107	122
	24	137	161

$$\widehat{OR}_{MH} = \frac{1 \cdot 257/282 + 3 \cdot 52/76 + 9 \cdot 107/161 + 14 \cdot 27/90}{17 \cdot 7/282 + 7 \cdot 14/76 + 15 \cdot 30/161 + 5 \cdot 44/90}$$

$$= 1.89$$

© drtamil@gmail.com 2020

Age 55+, ECG +

	CHD +	CHD -	
High Cat	14	44	58
Low Cat	5	27	32
	19	71	90

$$\widehat{OR}_{MH} = \frac{1 \cdot 257/282 + 3 \cdot 52/76 + 9 \cdot 107/161 + 14 \cdot 27/90}{17 \cdot 7/282 + 7 \cdot 14/76 + 15 \cdot 30/161 + 5 \cdot 44/90}$$

$$= 1.89$$

© drtamil@gmail.com 2020

Conf. Interval, OR=1.89, X^2=4.15

$$\widehat{OR}_{MH}^{1 \pm Z_{1-\alpha/2}/\chi_{MH}}$$

$$1.89^{1 \pm 1.96/\sqrt{4.15}}$$

$$1.89^{1 \pm 1.96/2.04}$$

$$1.89^{1 \pm 0.962}$$

$$1.89^{0.038} \quad to \quad 1.89^{1.962}$$

$$(1.02, 3.49)$$

© drtamil@gmail.com 2020

Conclusion

- There is a significant relationship between CAT and CHD, adjusted simultaneously for age and ECG ($p < 0.05; X^2_{MHtest}$).
- The adjusted OR is 1.89 (1.02, 3.49). Since the CI did not include the value of 1, therefore it is significant.
- Those who are stressed have significantly higher 2 times risk of developing CHD compared to those not stressed, after adjusting for age and ECG changes.

© drtamil@gmail.com 2020

Breslow-Day Test

Azmi Mohd Tamil

© drtamil@gmail.com 2020

Introduction

- Breslow & Day provided a test for assessing the homogeneity of the odds ratios across many tables/stratum.
- Its derivation involves solving a quadratic equation, therefore not advisable to calculate manually.
- I used an Excel trick to bypass the need for quadratic equation.

© drtamil@gmail.com 2020

Breslow & Day proposed a statistic (Equation 4.32) for testing the null hypothesis of homogeneity of the K true odds ratios. It sums up the squared deviations of observed and fitted values, each standardized by its variance

$$\sum_{k=1}^{K} \frac{(a_k - A_k(\hat{\psi}))^2}{\text{var}(a_k;\hat{\psi})},$$

where Ak(ψ) and var(ak ; ψ), denote the expected number and the asymptotic variance of exposed cases based on the MH adjusted odds ratio ψ , respectively.

Yep, the words doesn’t make any sense at all. You will hopefully understand it once you see the calculation in action.

© drtamil@gmail.com 2020

Equation 4.32

$$\chi^2_{\text{Breslow-Day}} = \sum_{i=1}^{\text{Kstrata}} \frac{\left[a_i - A_i(\text{using } OR_{MH})\right]^2}{\text{Var}(a_i;\text{null})}$$

$$\text{Var}(a_i;\text{null}) = \left(\frac{1}{A_i} + \frac{1}{B_i} + \frac{1}{C_i} + \frac{1}{D_i}\right)^{-1}$$

Breslow-Day uses the Mantel-Haenszel Odds Ratio to generate the expected tables. The most optimum would be to use conditional maximum likelihood estimator but that would need computing power.

© drtamil@gmail.com 2020

Step 1

- Calculate the Mantel-Haenszel adjusted Odds Ratio.

$$\widehat{OR}_{MH} = \frac{1 \cdot 257/282 + 3 \cdot 52/76 + 9 \cdot 107/161 + 14 \cdot 27/90}{17 \cdot 7/282 + 7 \cdot 14/76 + 15 \cdot 30/161 + 5 \cdot 44/90}$$

$$= 1.89$$

© drtamil@gmail.com 2020

Step 2

Observed Data			
Stress	CHD+	CHD-	Total
High	1	7	8
Low	17	257	274
Total	18	264	282

- MH OR=1.89. If for every stratum, the expected Odds Ratio is 1.89, what is the expected value of cell a for all tables?
- This is where you need computers to calculate for you. Shown only for 1st table.
- $1.89 = ad/bc = \frac{A \times (274-18+A)}{(8-A) \times (18-A)}$
- A = 0.8941.

Expected Data				OR
Stress	CHD+	CHD-	Total	
High	0.8941	7.1059	8	1.890
Low	17.1059	256.8941	274	
Total	18	264	282	

© drtamil@gmail.com 2020

Quadratic Equation (1st Stratum)

- ad/bc = 1.89
- 1.89 = $\frac{A \times (274-18+A)}{(8-A) \times (18-A)}$ where
  - a = A
  - b = (8 – A)
  - c = (18 – A)
  - d = 274 – c = (274 – 18 + A)

Observed Data			
Stress	CHD+	CHD-	Total
High	1 A	7	8
Low	17	257	274
Total	18	264	282

Credit to Dr Ihsan Zamzuri
p102428@siswa.ukm.edu.my

© drtamil@gmail.com 2020

Quadratic Equation

- $1.89 = \dfrac{A \times (274-18+A)}{(8-A) \times (18-A)}$
- $1.89 = \dfrac{A^2 + 256A}{A^2 – 26A+144}$
- $1.89A^2 – 49.14A + 272.16 = A^2 + 256A$
- $1.89A^2 – A^2 – 49.14A – 256A + 272.16 = 0$
- $0.89A^2 – 305.14A + 272.16 = 0$

Observed Data			
Stress	CHD+	CHD-	Total
High	1 A	7	8
Low	17	257	274
Total	18	264	282

Credit to Dr Ihsan Zamzuri
p102428@siswa.ukm.edu.my

© drtamil@gmail.com 2020

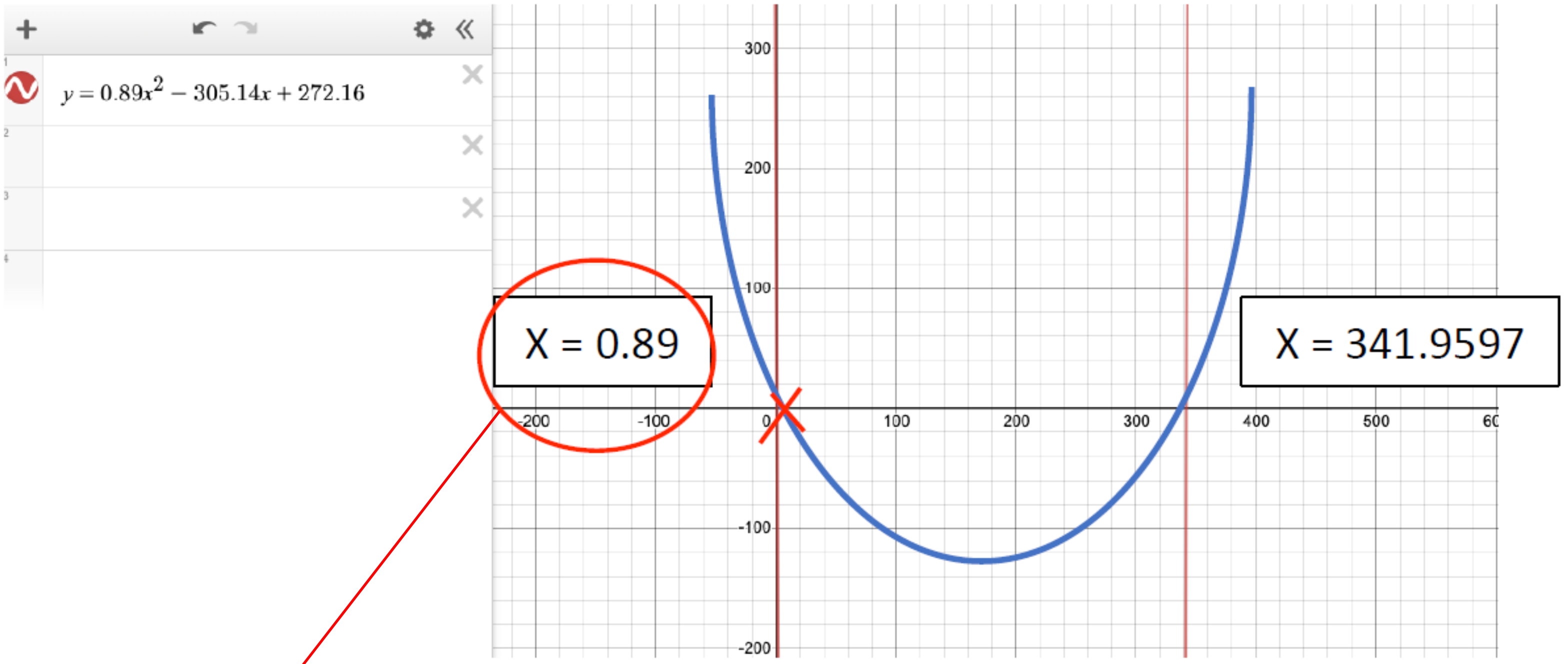
Quadratic Equation Using fx-570

- $0.89A^2 – 305.14A + 272.16 = 0$
- $y = 0.89x^2 – 305.14x + 272.16$
- Press Mode 3x & select EQN for equation.
- For “Unknowns”, press right to display “degree” then select 2 for quadratic equation.
- Enter 0.89 for a, -305.14 for b and 272.16 for c.
- Answer x1=341.959682, x2=0.89425089.

© drtamil@gmail.com 2020

Credit to Dr Ihsan Zamzuri p102428@siswa.ukm.edu.my

We take X=0.89425 since 341.95968 is too big for Table 1



Observed Data			
Stress	CHD+	CHD-	Total
High	1 A	7	8
Low	17	257	274
Total	18	264	282

Stress	CHD+	CHD-	Total	
High	0.8943	7.1057	8	1.890
Low	17.1057	256.8943	274	
Total	18	264	282	

© drtamil@gmail.com 2020

Step 3

- For each stratum, obtain the null hypothesis variance of the cell count.

Expected Data				OR
Stress	CHD+	CHD-	Total	
High	0.8943	7.1057	8	1.890
Low	17.1057	256.8943	274	
Total	18	264	282	

$$\mathrm{Var}(a_i;\mathrm{null}) = \left(\frac{1}{A_i} + \frac{1}{B_i} + \frac{1}{C_i} + \frac{1}{D_i}\right)^{-1}$$

- Var=(1/0.8943+1/7.1057+1/17.1057+1/256.8943)$^{-1}$

  = 0.75684.

© drtamil@gmail.com 2020

Step 4

$$\chi^2_{\text{Breslow-Day-}} = \sum_{i=1}^{\text{Kstrata}} \frac{[a_i - A_i(\text{using } OR_{MH})]^2}{\text{Var}(a_i;\text{null})}$$

- a = observed value
- A = expected value
- $\frac{(1-0.8943)^2}{0.75684}$

= 0.014762.

- Repeat for all tables.

Observed Data			
Stress	CHD+	CHD-	Total
High	1	7	8
Low	17	257	274
Total	18	264	282

Expected Data				OR
Stress	CHD+	CHD-	Total	
High	0.8943	7.1057	8	1.890
Low	17.1057	256.8943	274	
Total	18	264	282	

© drtamil@gmail.com 2020

Excel Spreadsheet

Stratum	Observed Data				OR	Expected Data				OR	A€	V€	Breslow
	Stress	CHD+	CHD-	Total		Stress	CHD+	CHD-	Total				
Young ECG-	High	1	7	8	2.16	High	0.8943	7.1057	8	1.890	0.8943	0.7568	0.014762
	Low	17	257	274		Low	17.1057	256.8943	274				
	Total	18	264	282		Total	18	264	282				
	Stress	CHD+	CHD-	Total		Stress	CHD+	CHD-	Total				
Young ECG+	High	3	14	17	1.59	High	3.309	13.691	17	1.890	3.3090	1.8388	0.051924
	Low	7	52	59		Low	6.691	52.309	59				
	Total	10	66	76		Total	10	66	76				
	Stress	CHD+	CHD-	Total		Stress	CHD+	CHD-	Total				
Old ECG-	High	9	30	39	2.14	High	8.442	30.558	39	1.890	8.4420	4.4474	0.07001
	Low	15	107	122		Low	15.558	106.442	122				
	Total	24	137	161		Total	24	137	161				
	Stress	CHD+	CHD-	Total		Stress	CHD+	CHD-	Total				
Old ECG+	High	14	44	58	1.72	High	14.284	43.716	58	1.890	14.2840	2.9276	0.02755
	Low	5	27	32		Low	4.716	27.284	32				
	Total	19	71	90		Total	19	71	90				
	Stress	CHD+	CHD-	Total									
TOTALS	High	27	95	122	2.86						26.929	9.971	**0.164**
	Low	44	443	487									
	Total	71	538	609									

Breslow-Day Chi Square

© drtamil@gmail.com 2020

Step 5

- Sum up all the differences and check the p value from the chi square table (df = 3, since 4 stratum).
- χ^2_{BDTest} = 0.164; (d.f.=3) therefore p > 0.5.
- Since the test of homogeneity is not significant, all the OR of the stratums are homogenous.

Table 3 : Percentage point of X^2

	P Value							
d.k.	0.5	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	1.39	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	2.37	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	3.36	5.39	7.78	9.49	11.14	13.28	14.86	18.47
5	4.35	6.63	9.24	11.07	12.83	15.09	16.75	20.52
6	5.35	7.84	10.64	12.59	14.45	16.81	18.55	22.46
7	6.35	9.04	12.02	14.07	16.01	18.48	20.28	24.32
8	7.34	10.22	13.36	15.51	[illegible]	20.09	21.96	26.13
9	8.34	11.39	14.68	16.92	[illegible]	21.67	23.59	27.88
10	9.34	12.55	15.99	18.31	[illegible]	[illegible]	25.19	29.59
11	10.34	13.70	17.28	19.68	21.92	[illegible]	26.76	31.26
12	11.34	14.85	18.55	21.03	23.34	[illegible]	28.30	32.91
13	12.34	15.98	19.81	22.36	24.74	[illegible]	29.69	34.53
14	13.34	17.12	21.06	23.68	26.12	[illegible]	31.32	36.12
15	14.34	18.25	22.31	25.00	27.49	[illegible]	32.80	37.70
16	15.34	19.37	23.54	26.30	28.85	[illegible]	[illegible]	39.25
17	16.34	20.49	24.77	27.59	30.19	[illegible]	[illegible]	40.79
18	17.34	21.60	25.99	28.87	31.53	34.81	[illegible]	42.31
19	18.34	22.72	27.20	30.14	32.85	36.19	[illegible]	43.82
20	19.34	23.83	28.41	31.41	34.17	37.57	[illegible]	45.32
21	20.34	24.93	29.62	32.67	35.48	38.93	[illegible]	46.80
22	21.34	26.04	30.81	33.92	36.78	40.29	[illegible]	[illegible]
23	22.34	27.14	32.01	35.17	38.08	41.64	[illegible]	[illegible]
24	23.34	28.24	33.20	36.42	39.36	42.98	[illegible]	[illegible]
25	24.34	29.34	34.38	37.65	40.65	44.31	46.93	[illegible]
26	25.34	30.43	35.56	38.89	41.92	45.64	48.29	[illegible]
27	26.34	31.53	36.74	40.11	43.19	46.96	49.64	[illegible]
28	27.34	32.62	37.92	41.34	44.46	48.28	50.99	[illegible]
29	28.34	33.71	39.09	42.56	45.72	49.59	52.34	[illegible]
30	29.34	34.80	40.26	43.77	46.98	50.89	53.67	[illegible]
40	39.34	45.62	51.81	55.76	59.34	63.69	66.77	[illegible]
50	49.33	56.33	63.17	67.50	71.42	76.15	79.49	[illegible]
60	59.33	66.98	74.40	79.08	83.30	88.38	91.95	[illegible]
70	69.33	77.58	85.53	90.53	95.02	100.43	104.22	112.32
80	79.33	88.13	96.58	101.88	106.63	112.33	116.32	124.84
90	89.33	98.65	107.57	113.15	118.14	124.12	128.30	137.21
100	99.33	109.14	118.50	124.34	129.56	135.81	140.17	149.45

Refer to Table 3.

Look at df = 3.

X^2_{BDtest} = 0.164, smaller than 2.37 (p=0.5)

0.164>2.37

Therefore if X^2_{BDtest}=0.164, p>0.5.

d.k.	0.5	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83
2	1.39	2.77	4.61	5.99	7.38	9.21	10.60	13.82
3	2.37	4.11	6.25	7.81	9.35	11.34	12.84	16.27
4	3.36	5.39	7.78	9.49	11.14	13.28	14.86	18.47

© drtamil@gmail.com 2020

Same result as SPSS

A€	V€	Breslow
0.8941	0.7567	0.01482
3.3090	1.8388	0.05192
8.4420	4.4474	0.07001
14.2840	2.9276	0.02755
26.929	9.971	**0.164**

Tests of Homogeneity of the Odds Ratio

	Chi-Squared	df	Asymp. Sig. (2-sided)
Breslow-Day	.164	3	.983
Tarone's	.164	3	.983

© drtamil@gmail.com 2020

Tarone Adjustment

Should subtract Tarone correction from Breslow-Day statistic to get better chi-square approximation.

Tarone correction =

Thus, for a consistent estimator $\tilde{\psi}$, the modified heterogeneity score test statistic is

$$X^2(\tilde{\psi}) = \sum_{k=1}^{K} \{x_k - e_k(\tilde{\psi})\}^2 / v_k(\tilde{\psi}) - \{x_. - e_.(\tilde{\psi})\}^2 / v_.(\tilde{\psi}).$$

Thus, let $\hat{e}_k(\psi)$ be the appropriate solution to the quadratic equation

$$\frac{\hat{e}_k(n_{2k} - m_k + \hat{e}_k)}{(m_k - \hat{e}_k)(n_{1k} - \hat{e}_k)} = \psi,$$

and let

$$\hat{v}_k(\psi) = \{\hat{e}_k^{-1} + (m_k - \hat{e}_k)^{-1} + (n_{1k} - \hat{e}_k)^{-1} + (n_{2k} - m_k + \hat{e}_k)^{-1}\}^{-1}$$

© drtamil@gmail.com 2020

Why Tarone?

- *Tarone noted that by using the MH Odds Ratio estimator instead of the better conditional maximum likelihood estimator, the Breslow–Day test statistic becomes like the conditional likelihood score test. Since the MH estimator is inefficient, Tarone noted that the test statistic is stochastically larger than a χ2 random variable under the homogeneity hypothesis.*
- *Tarone wrote in 1985; "this paper derives the appropriate modification of the heterogeneity score test when the parameter of interest is estimated by an inefficient, but consistent, estimator."*

© drtamil@gmail.com 2020

Summary

- CMH test assumes common odds ratio θ and tests if it is 1.
- Mantel-Haenszel estimate of the odds ratio averages numerators and denominators before taking the ratio.
- Breslow-Day test checks if odds ratios are indeed common using discrepancies in (observed – expected) cell counts.
- Tarone's adjustment claims that using MH Odds Ratio estimator for the test of homogeneity, is inefficient, therefore needs to be corrected. But the formula is all Greek to me, so I give up.

© drtamil@gmail.com 2020

In SPSS

- In this data, we are trying to see the relationship between CAT & CHD and see whether AGE & ECG changes are Confounders.

© drtamil@gmail.com 2020

Data For Exercise

CATvsCHD-CMH.sav [DataSet1] - SPSS Statistics Data Editor

File Edit View Data Transform Analyze Graphs Utilities Add-o

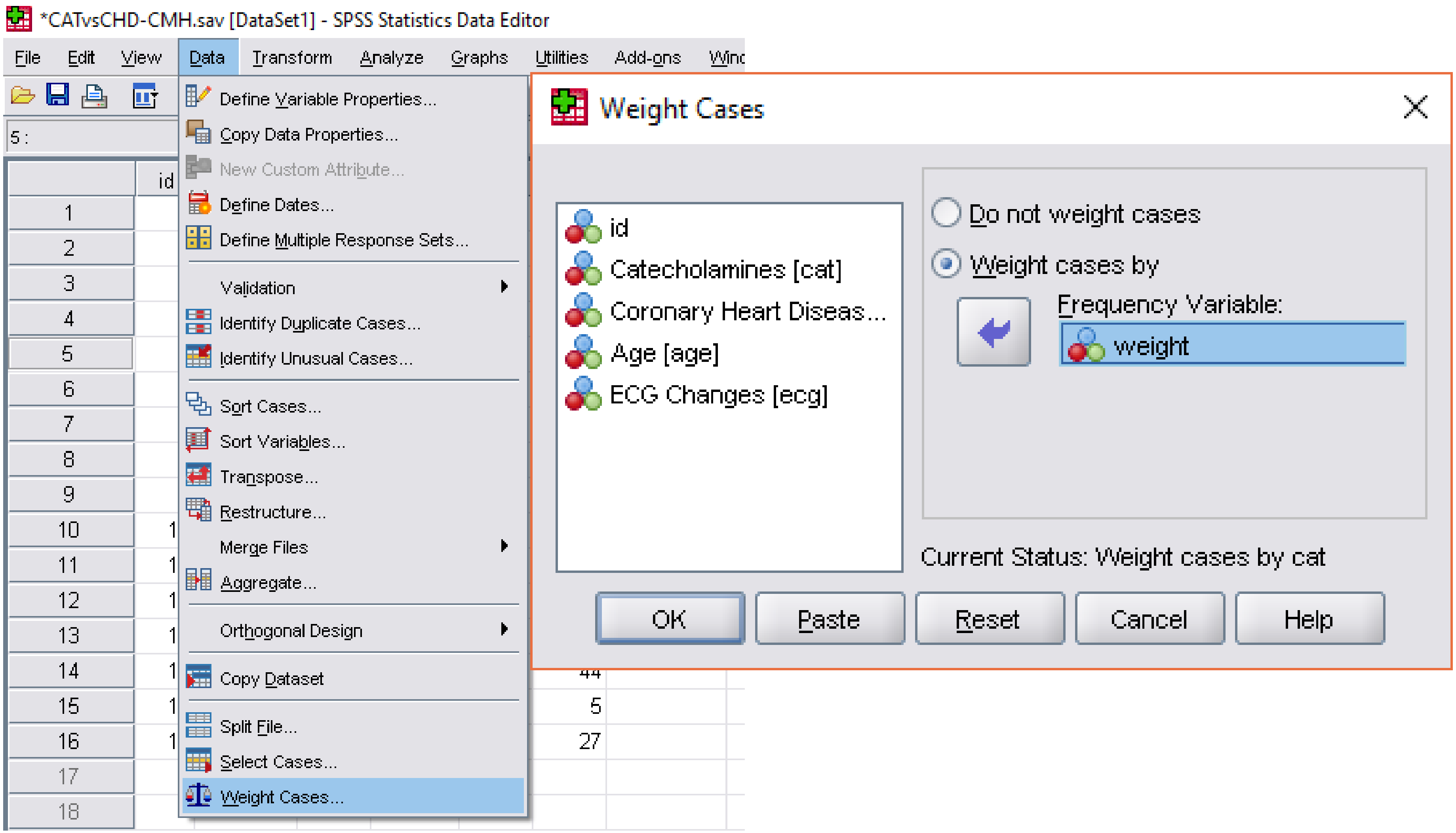
5 :

	id	cat	chd	age	ecg	weight
1	1	High Cat	CHD+	Age <55	ECG-	1
2	2	High Cat	CHD-	Age <55	ECG-	7
3	3	Low Cat	CHD+	Age <55	ECG-	17
4	4	Low Cat	CHD-	Age <55	ECG-	257
5	5	High Cat	CHD+	Age <55	ECG+	3
6	6	High Cat	CHD-	Age <55	ECG+	14
7	7	Low Cat	CHD+	Age <55	ECG+	7
8	8	Low Cat	CHD-	Age <55	ECG+	52
9	9	High Cat	CHD+	Age 55+	ECG-	9
10	10	High Cat	CHD-	Age 55+	ECG-	30
11	11	Low Cat	CHD+	Age 55+	ECG-	15
12	12	Low Cat	CHD-	Age 55+	ECG-	107
13	13	High Cat	CHD+	Age 55+	ECG+	14
14	14	High Cat	CHD-	Age 55+	ECG+	44
15	15	Low Cat	CHD+	Age 55+	ECG+	5
16	16	Low Cat	CHD-	Age 55+	ECG+	27

https://wp.me/p4mYLF-81

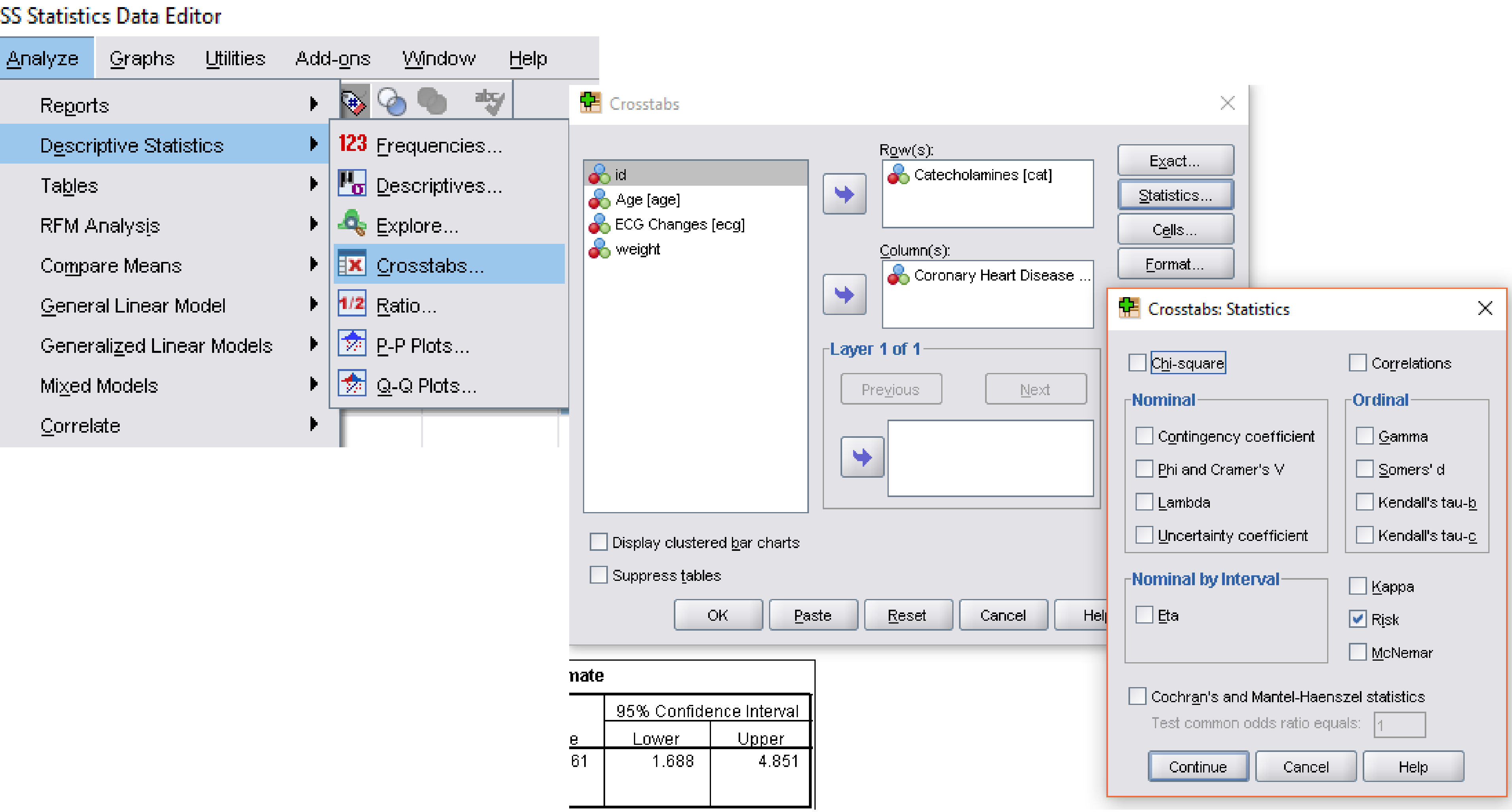
© drtamil@gmail.com 2020

Weighted Analysis



© drtamil@gmail.com 2020

Combined Analysis



© drtamil@gmail.com 2020

Combine

Catecholamines * Coronary Heart Disease Crosstabulation

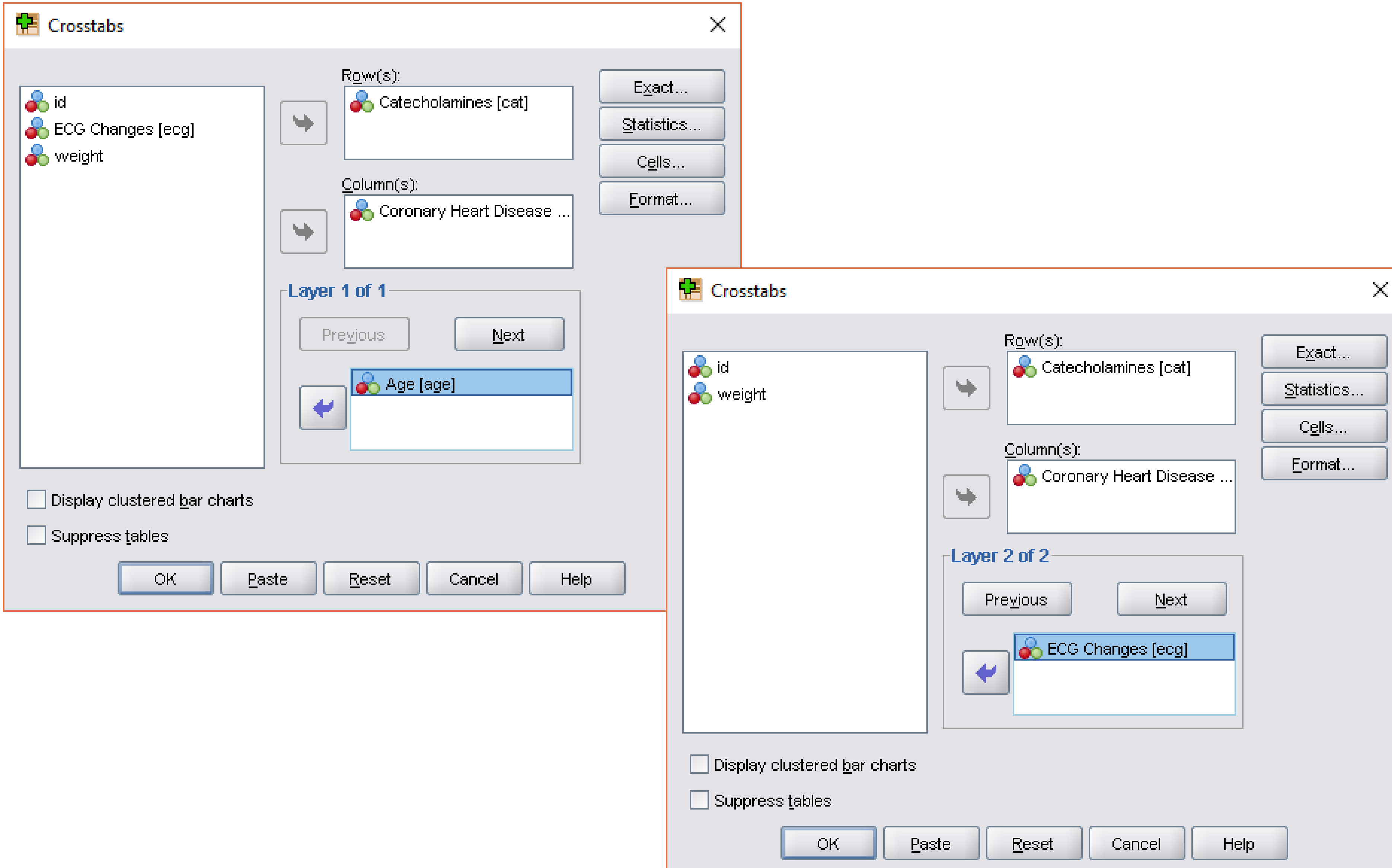
			Coronary Heart Disease		Total
			CHD-	CHD+	
Catecholamines	Low Cat	Count	443	44	487
		% within Catecholamines	91.0%	9.0%	100.0%
	High Cat	Count	95	27	122
		% within Catecholamines	77.9%	22.1%	100.0%
Total		Count	538	71	609
		% within Catecholamines	88.3%	11.7%	100.0%

Risk Estimate

	Value	95% Confidence Interval	
		Lower	Upper
Odds Ratio for Catecholamines (Low Cat / High Cat)	2.861	1.688	4.851

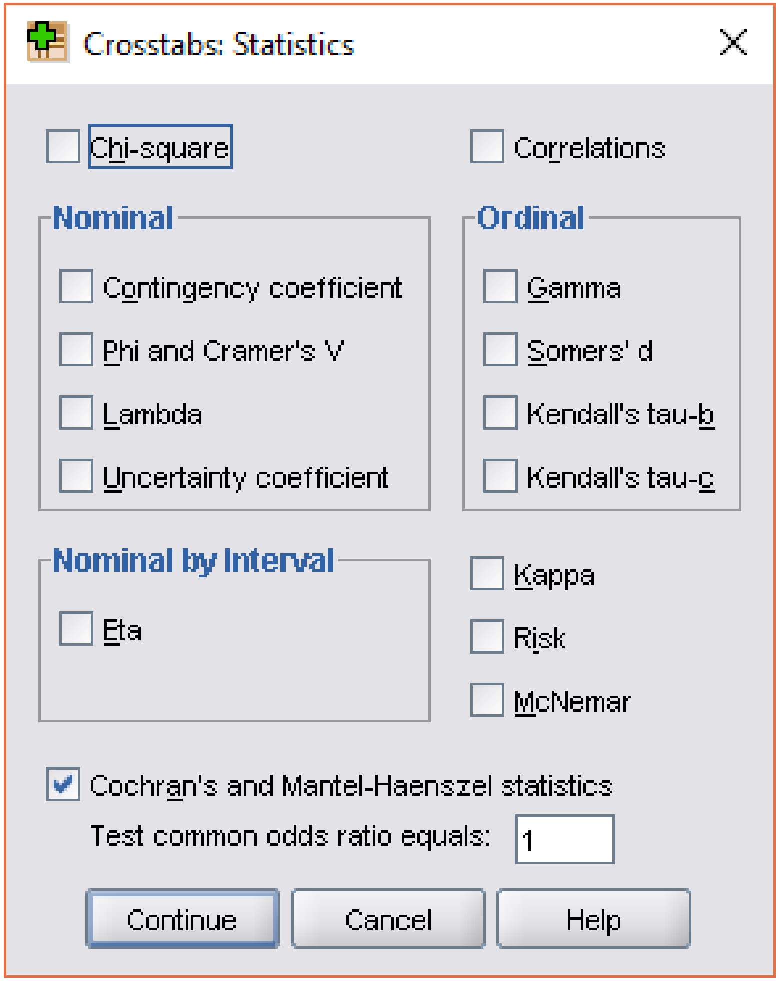
© drtamil@gmail.com 2020

Analyse->Descriptives->Crosstab



© drtamil@gmail.com 2020

Select CMH statistics



© drtamil@gmail.com 2020

Catecholamines * Coronary Heart Disease * ECG Changes * Age Crosstabulation

Age	ECG Changes				Coronary Heart Disease		Total
					CHD-	CHD+	
Age <55	ECG-	Catecholamines	Low Cat	Count	257	17	274
				% within Catecholamines	93.8%	6.2%	100.0%
			High Cat	Count	7	1	8
				% within Catecholamines	87.5%	12.5%	100.0%
		Total		Count	264	18	282
				% within Catecholamines	93.6%	6.4%	100.0%
	ECG+	Catecholamines	Low Cat	Count	52	7	59
				% within Catecholamines	88.1%	11.9%	100.0%
			High Cat	Count	14	3	17
				% within Catecholamines	82.4%	17.6%	100.0%
		Total		Count	66	10	76
				% within Catecholamines	86.8%	13.2%	100.0%
Age 55+	ECG-	Catecholamines	Low Cat	Count	107	15	122
				% within Catecholamines	87.7%	12.3%	100.0%
			High Cat	Count	30	9	39
				% within Catecholamines	76.9%	23.1%	100.0%
		Total		Count	137	24	161
				% within Catecholamines	85.1%	14.9%	100.0%
	ECG+	Catecholamines	Low Cat	Count	27	5	32
				% within Catecholamines	84.4%	15.6%	100.0%
			High Cat	Count	44	14	58
				% within Catecholamines	75.9%	24.1%	100.0%
		Total		Count	71	19	90
				% within Catecholamines	78.9%	21.1%	100.0%

© drtamil@gmail.com 2020

Odds Ratio by Stratum

Risk Estimate

Age	ECG Changes		Value	95% Confidence Interval Lower	95% Confidence Interval Upper
Age <55	ECG-	Odds Ratio for Catecholamines (Low Cat / High Cat)	2.160	.251	18.578
	ECG+	Odds Ratio for Catecholamines (Low Cat / High Cat)	1.592	.364	6.962
Age 55+	ECG-	Odds Ratio for Catecholamines (Low Cat / High Cat)	2.140	.853	5.371
	ECG+	Odds Ratio for Catecholamines (Low Cat / High Cat)	1.718	.556	5.308

Risk Estimate

	Value	95% Confidence Interval Lower	95% Confidence Interval Upper
Odds Ratio for Catecholamines (Low Cat / High Cat)	2.861	1.688	4.851

© drtamil@gmail.com 2020

Odds Ratio

Stratum	OR
<55, ECG+	1.59
55+, ECG+	1.72
55+, ECG-	2.14
<55, ECG-	2.16
Crude OR	2.86

- There seems to be little effect modification due to age and ECG. But combined table stronger & highly significant.
- Need to adjust for effect of age & ECG.

© drtamil@gmail.com 2020

No Interaction between Age & ECG Changes with Catecholamine Level

Tests of Homogeneity of the Odds Ratio

	Chi-Squared	df	Asymp. Sig. (2-sided)
Breslow-Day	.164	3	.983
Tarone's	.164	3	.983

Since the test of homogeneity is not significant, all the OR of the stratums are homogenous. The changing level of Age & ECG did not change CHD OR much.

© drtamil@gmail.com 2020

Adjusted OR = 1.891, different than unadjusted OR=2.86. p value is significant, indicating OR sig. since Confidence Interval did not include 1.

Mantel-Haenszel Common Odds Ratio Estimate

Estimate			1.891
ln(Estimate)			.637
Std. Error of ln(Estimate)			.316
Asymp. Sig. (2-sided)			.044
Asymp. 95% Confidence Interval	Common Odds Ratio	Lower Bound	1.017
		Upper Bound	3.516
	ln(Common Odds Ratio)	Lower Bound	.017
		Upper Bound	1.257

The Mantel-Haenszel common odds ratio estimate is asymptotically normally distributed under the common odds ratio of 1.000 assumption. So is the natural log of the estimate.

© drtamil@gmail.com 2020

Magnitude of Confounding > 10%

- May cause an overestimate (positive confounding) or an underestimate (negative confounding).
- Can be quantified by computing the percentage difference between the crude and adjusted measures.
- If Adjusted OR = 1.891, Crude OR=2.86.
  - Epid; (2.86 – 1.891)/1.891 = 51.24%
  - Stats; (2.86 - 1.891)/2.86 = 33.88%
- % larger than 10%, therefore Age/ECG changes are positive confounding factors for CAT.

© drtamil@gmail.com 2020

X^2_{MH}

Tests of Conditional Independence

	Chi-Squared	df	Asymp. Sig. (2-sided)
Cochran's	4.191	1	.041
Mantel-Haenszel	3.510	1	.061

- Even after adjusting for Age & ECG changes, X^2_{CMH} is 4.19, p=0.041, therefore sig association between CAT level & CHD.

© drtamil@gmail.com 2020

Using SPSS X^2_{CMH}

$$\frac{\left[\frac{1\cdot 257-17\cdot 7}{282}+\frac{3\cdot 52-7\cdot 14}{76}+\frac{9\cdot 107-15\cdot 30}{161}+\frac{14\cdot 27-5\cdot 44}{90}\right]^2}{\left[\frac{8\cdot 274\cdot 18\cdot 264}{281\cdot 282^2}+\frac{17\cdot 59\cdot 10\cdot 66}{75\cdot 76^2}+\frac{39\cdot 122\cdot 24\cdot 137}{160\cdot 161^2}+\frac{58\cdot 32\cdot 19\cdot 71}{89\cdot 90^2}\right]}$$

$$=\frac{(.4894+.7632+3.1863+1.7556)^2}{.4661+1.5281+3.7721+3.4731}$$

$$=4.15(p=.04,\text{ two-sided})$$

Tests of Conditional Independence

	Chi-Squared	df	Asymp. Sig. (2-sided)
Cochran's	4.191	1	.041
Mantel-Haenszel	3.510	1	.061

© drtamil@gmail.com 2020

Using SPSS X^2_{CMH}

continuity correction? enter 'y' or 'n':		n		
	chi-square:	4.153		
	d.f.:	1		
	P-value:	0.042		

stratum↓	outcome risk ↓→	CHD+	CHD-	proportion CHD+
Young, ECG-	High	1	7	0.125
	Low	17	257	0.062
		CHD+	CHD-	
Young, ECG+	High	3	14	0.176
	Low	7	52	0.119
		CHD+	CHD-	
Old, ECG-	High	9	30	0.231
	Low	15	107	0.123
		CHD+	CHD-	
Old, ECG+	High	14	44	0.241
	Low	5	27	0.156

$$X^2_{MH} = \frac{\sum\left(\frac{ad - cb}{n}\right)^2}{\sum\left(\frac{(a+b)\times(c+d)\times(a+c)\times(b+d)}{(n-1)\times n^2}\right)}$$

Tests of Conditional Independence

	Chi-Squared	df	Asymp. Sig. (2-sided)
Cochran's	4.191	1	.041
Mantel-Haenszel	3.510	1	.061

© drtamil@gmail.com 2020

Using Continuity Correction X^2_{MH}

continuity correction? enter 'y' or 'n': y

chi-square:	3.51
d.f.:	1
P-value:	0.061

$$\chi^2_{MH} = \frac{\{|\sum[a-(a+b)(a+c)/n]|-0.5\}^2}{\sum(a+b)(a+c)(b+d)(c+d)/(n^3-n^2)}$$

stratum↓	outcome risk ↓→	CHD+	CHD-	proportion CHD+
Young, ECG-	High	1	7	0.125
	Low	17	257	0.062
		CHD+	CHD-	
Young, ECG+	High	3	14	0.176
	Low	7	52	0.119
		CHD+	CHD-	
Old, ECG-	High	9	30	0.231
	Low	15	107	0.123
		CHD+	CHD-	
Old, ECG+	High	14	44	0.241
	Low	5	27	0.156

Tests of Conditional Independence

	Chi-Squared	df	Asymp. Sig. (2-sided)
Cochran's	4.191	1	.041
Mantel-Haenszel	3.510	1	.061

© drtamil@gmail.com 2020

When to use Continuity Correction?

- https://www.statsdirect.com/help/meta_analysis/mh.htm
- If any cell count in any of the stratum tables is zero, then the continuity correction should be applied.
- $$\chi^2_{MH} = \frac{(|\sum(a-(a+b)(a+c)/n)|-0.5)^2}{\sum(a+b)(a+c)(b+d)(c+d)/(n^3-n^2)}$$

© drtamil@gmail.com 2020

Using StatCalc X^2_{MH}

	+ Disease	−	
+	27	95	122
−	44	443	487
Exposure	71	538	609

Mantel-Haenszel test (MH)

$$\chi^2(MH) = \frac{(n-1)(ad-bc)^2}{(n_1 \times n_0 \times m_1 \times m_0)} =$$

$$= \frac{608*((27*443)-(95*44))^2}{(71*538*122*487)}$$

=16.21978128

```
Analysis of Single Table
Odds ratio = 2.86 (1.63 <OR< 5.01)
Cornfield 95% confidence limits for OR
Relative risk = 2.45 (1.58 <RR< 3.79)
Taylor Series 95% confidence limits for RR
Ignore relative risk if case control study.
```

	Chi-Squares	P-values
Uncorrected :	16.25	0.0000556
Mantel-Haenszel:	16.22	0.0000564
Yates corrected:	15.00	0.0001075

d.k.	0.5	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83

- http://web1.sph.emory.edu/activepi/Instructors/Kevin_MSword/lesson_12boh.htm
- The Mantel-Haenszel Test in StatCalc is a large-sample version of Fisher's Exact Test, not the same as CMH Chi-square.

© drtamil@gmail.com 2020

Using StatCalc X^2_{MH}

	+ Disease	−	
+	27	95	122
−	44	443	487
	71	538	609

Exposure

Mantel-Haenszel test (MH)

$$\chi^2(MH) = \frac{(n-1)(ad - bc)^2}{(n_1 \times n_0 \times m_1 \times m_0)} =$$

$$= \frac{608*((27*443)-(95*44))^2}{(71*538*122*487)}$$

$= 16.21978128$

```
        Analysis of Single Table
   Odds ratio = 2.86 (1.63 <OR< 5.01)
 Cornfield 95% confidence limits for OR
  Relative risk = 2.45 (1.58 <RR< 3.79)
Taylor Series 95% confidence limits for RR
Ignore relative risk if case control study.
```

	Chi-Squares	P-values
Uncorrected :	16.25	0.0000556
Mantel-Haenszel:	16.22	0.0000564
Yates corrected:	15.00	0.0001075

d.f.	0.5	0.25	0.1	0.05	0.025	0.01	0.005	0.001
1	0.45	1.32	2.71	3.84	5.02	6.63	7.88	10.83

Table 17.3 Crude Analysis of CAT–CHD Data

	HI CAT	*LO CAT*	
CHD	27	44	71
$\overline{CHD}$	95	443	538
	122	487	609

$$c\hat{R}R = \frac{27/122}{44/487} = 2.45$$

$$\chi^2_{MH} = \frac{608[(27)(443) - (44)(95)]^2}{(122)(487)(71)(538)}$$

$= 16.22 \quad (P < .001)$

© drtamil@gmail.com 2020

Conclusion

- There is a significant relationship between CAT and CHD, adjusted simultaneously for age and ECG ($p < 0.05$; X^2_{CMH}).
- The adjusted OR is 1.89 (1.02, 3.49). Since the CI did not include the value of 1, therefore it is significant.
- Those who are stressed have significantly higher 2 times risk of developing CHD compared to those not stressed, after adjusting for age and ECG changes.

© drtamil@gmail.com 2020

References

- David G. Kleinbaum, Lawrence L. Kupper, Hal Morgenstern. 1982. *Epidemiologic Research: Principles and Quantitative Methods.* John Wiley & Sons. (pages 325, 447-460)
- N. E. Breslow & N. E. Day. 1980. *Statistical Methods In Cancer Research Volume 1 - The Analysis Of Case-control Studies.* International Agency For Research On Cancer, World Health Organization. (pages 136-146)
- Robert E. Tarone. 1985. *On Heterogeneity Tests Based On Efficient Scores*. Biometrika, Volume 72, Issue 1, April 1985. (pages 91–95).