ARISTOTLE UNIVERSITY
OF THESSALONIKI

FACULTY OF HEALTH SCIENCES - SCHOOL OF MEDICINE
MSc Health Statistics and Data Analytics

Association of categorical variables

Anna-Bettina Haidich
Associate Professor of Medical Statistics –Epidemiology
haidich@auth.gr

THESSALONIKI 2021-22

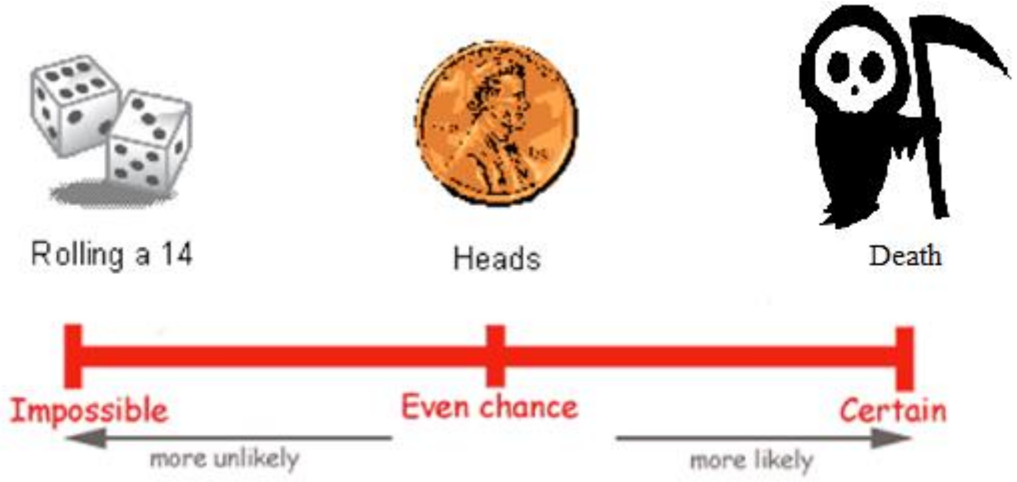
MSc Health Statistics & Data Analytics

Learning objectives

- Upon completion of this lecture you will be able to:
  - Define and explain the idea of risk and its relationship with probability
  - Define and explain the idea of odds
  - Test whether two categorical variables are associated using the chi-square test or the Mc Nemar's test
  - Calculate a Risk Ratio and its confidence interval
  - Calculate an Odds Ratio and its confidence interval

Probability = Proportion = Risk

- $0 \leq P \leq 1$



- If the probability of an event happening is p, then the probability of the event not happening is 1 – p.
- In medicine, the term "risk" is often used in place of "proportion" or "probability"

Risk

Number of patients that died after therapy

	Death	No death	Total
Patients	30	70	100

Risk (probability) of death after therapy:

30 / 100 = 0.30 30%

The numerator **<u>is included</u>** in the denominator

Odds

Number of patients that died after therapy

	Death	No death	Total
Patients	30	70	100

Odds of death after therapy:

(30 /100)/ (70/100) = 30/70 = 3/7 = 0.43

The numerator **is not included** in the denominator

Risk and Odds

The probability of something happening can be expressed as a risk or as an odds:

$$\text{RISK} = \frac{\text{the probability of something happening}}{\text{the probability of } \textit{all} \text{ things happening}}$$

$$\text{ODDS} = \frac{\text{the probability of something happening}}{\text{the probability of it } \textit{not} \text{ happening}}$$

Risk

- What is the probability of getting a six when rolling a dice?

1 in 6 (1/6 = 16.7%)

The numerator **is included** in the denominator

Odds

- What is the probability of getting a six when rolling a dice compared to not getting a six?

1 to 5 (1/5 = 20%)

Numerator: Event / Denominator: No event

Odds

$$\frac{\text{Probability of getting a 6}}{\text{Probability of not getting a 6}} = \frac{1/6}{5/6} = 1{:}5$$

Numerator + Denominator: 1/6 + 5/6 = 1

Risk and Odds

Thus a risk is a **proportion**,

But an odds is a **ratio**.

An odds is a special type of ratio one in which the numerator and denominator sum to one.

Risk and Odds

- odds = p/(1 – p)
- risk or probability (p)= odds/(1 + odds)
- Note:
  - $0 \le p$ and $1 - p \le 1$.
  - $0 \le \text{odds} \le \infty$
  - p=0.5 → odds = 0.5 / (1 - 0.5) = 1
  - The odds always take larger values than the probability (since 1 - probability < 1)
  - Probability <0.1 → probability ≈ odds

  (rare diseases)

The link between probability and odds

- The difference between risk and odds is small when the events are rare but can be large for common events

Event	Total	Risk	Odds
5	100	0.05	0.0526
50	100	0.5	1
95	100	0.95	19

Measures of association: Risk Ratio

- Measures of association quantify the relationship between two variables; usually the outcome variable and the exposure variable.

	Disease		Total
Exposure	Yes	No	
Yes	a	b	a+b
No	c	d	c+d
Total	a+c	b+d	n

- Among those exposed to the risk factor, the risk of disease = a/(a + b).
- Among those not exposed, the risk of disease = c/(c + d).

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

H_0: RR=1
H_1: RR≠1

Measures of association: Odds Ratio

- The Odds Ratio (OR) is the ratio between the odds of the exposure among those with the outcome and the odds of the exposure among those without the outcome.

Exposure	Disease: Yes	Disease: No	Total
Yes	a	b	a+b
No	c	d	c+d
Total	a+c	b+d	n

The odds of exposure to the risk factor among those with the disease = a/c

The odds of exposure to the risk factor among those with no disease = b/d.

$$Odds\ Ratio(OR) = \frac{\frac{a}{c}}{\frac{b}{d}} = \frac{a \times d}{b \times c}$$

H_0: OR=1

H_1: OR≠1

The odds ratio (OR) is usually called the cross-product term.

Direction of results

- $0 \leq RR/OR < \infty$

- RR = 1 or OR = 1 there is no difference in the risk of event between two groups

- RR > 1 or OR > 1 the event rate **is larger** in the group in the numerator (OR > RR)

- RR < 1 or OR < 1 the event rate **is lower** in the group in the numerator (OR < RR)

Chi-square test (Independence test)

- Test for an association between 2 categorical variables.

  - **H_0**: There is no association between two categorical variables (they are independent)
  - **H_1**: There is an association between two categorical variables (they are dependent)

- If two categorical variables are related, it means the chance that an individual falls into a particular category for one variable depends upon the particular category they fall into for the other variable.

Assumptions of chi-square test

- Independent observations
  - Each individual is represented only once in the study. The rows (and columns) of the table are mutually exclusive, implying that each individual can belong in only one row and only one column.
  - Mc Nemar’s test for dependent observations
    - e.g. cross-over trials or matched case-control studies

- Large expected frequencies.
  - A) All **expected counts** should be > 1.
  - B) At least 80% of **expected counts** should > 5.

Vaccination and Influenza Example

- Association between vaccination (yes, no) and influenza (yes, no). When we test if there is an association between these two variables, we are trying to determine if subjects being vaccinated are less likely to contract influenza. If that is the case, then we can say that vaccination and influenza are ***related*** or ***associated***.

Chi-Square Analysis Details

The 5 Steps in a Chi-Square Test:

- **Step 1: Write the null and alternative hypothesis.**

  H_0: There is no association between the variables.
  H_1: There is an association between the variables.

- **Step 2: Check conditions.**

  A) All **expected counts** should be > 1.

  B) At least 80% of **expected counts** should be > 5.

- **Step 3: Calculate Test Statistic and p-value.**

  The test statistic measure the difference between the observed counts and the expected counts assuming independence.

$$\chi^2 = \sum_{\text{all cells}} \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} = \sum_{i,j} \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

This is called **chi-square statistic** because if the null hypothesis is true, then it has a *chi-square distribution* with (r-1)x(c-1) degrees of freedom.

Chi-Square Analysis Details

<u>The 5 Steps in a Chi-Square Test:</u>

- **Step 4: Decide whether or not the result is statistically significant.**
  The results are statistically significant if the p-value < α, where α is the significance level (usually α = 0.05).

- **Step 5: Report the conclusion in the context of the situation.**

  - The ***p-value*** <0.05, this result ***is statistically significant***. **Reject the H0.** Conclude that (the two variables) are related.

  - The ***p-value*** > 0.05, this result ***is NOT statistically significant***. **We cannot reject the H0** . Cannot conclude that (the two variables) are related.

Vaccination & Influenza example

	Influenza		
	Yes	No	Total
Vaccine	20 (8.3%)	220 (91.7%)	240
Placebo	80 (36.4%)	140 (63.6%)	220
Total	100 (21.7%)	360 (78.3%)	460

- **Step 1. Define the null and alternative hypotheses under study**
- **H_0**: There is **no association** between the influenza vaccine and the contraction of influenza.
- **H_1**: There is **an association** between the influenza vaccine and the contraction of influenza.

Influenza example (R do the work for us)

	Influenza		Total
	Yes	No	
Vaccine	O_{11} = 20	O_{12} = 220	R_1 = 240
Placebo	O_{21} = 80	O_{22} = 140	R_2 = 220
Total	C_1 = 100	C_2 = 360	N=460

E_{11}=(240x100)/460=52.2 E_{12}=(240x360)/460=187.8

E_{21}=(220x100)/460=47.8 E_{22}=(220x360)/460=172.2

Step 2: All expected counts > 1

Step 3: Calculate chi-square

$$\chi^2 = \frac{(20-52.2)^2}{52.2} + \frac{(80-47.8)^2}{47.8} + \frac{(220-187.8)^2}{187.8} + \frac{(140-172.2)^2}{172.2} = 19.86 + 21.69 + 5.52 + 6.02 = 53.09$$

shortcut formula

$$\chi^2 = \frac{460 \times (20 \times 140 - 80 \times 220)^2}{100 \times 360 \times 240 \times 220} = 53.01$$

$$\chi^2 = \frac{N(O_{11} \cdot O_{22} - O_{12} \cdot O_{21})^2}{R_1 C_1 R_2 C_2}$$

Vaccination & Influenza example (cont.)

- **Step 4. : Decide whether or not the result is statistically significant**
  χ^2=53.01 P<0.001 H_0 is rejected

- **Step 5. Interpretation**
  There is an association between the influenza vaccine and the contraction of influenza. The percentage contracting influenza was significantly lower in the vaccine group than in the placebo group (8.3% vs. 36.4%; P<0.001).

Chi-square test and z-test

- $Z^2 = \chi_1^{\ 2}$

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

	Influenza		
	Yes	No	Total
Vaccine	20 (8.3%)	220 (91.7%)	240
Placebo	80 (36.4%)	140 (63.6%)	220
Total	100 (21.7%)	360 (78.3%)	460

2-proportions

H0: p1 – p2 = 0

H1: p1 – p2 ≠ 0

Chi-Square

H0: There is no association between vaccination and influenza.

H1: There is an association between vaccination and influenza.

Chi-square test and z-test

$$z = \frac{0.0833 - 0.3636}{\sqrt{0.2174 \times (1 - 0.2174) \times (\frac{1}{240} + \frac{1}{220})}} = \frac{-0.2803}{\sqrt{0.2174 \times 0.7826 \times (0.004167 + 0.004546)}}$$

$$z = \frac{-0.2803}{\sqrt{0.2174 \times 0.7826 \times 0.008713}} = \frac{-0.2803}{\sqrt{0.001482}} = \frac{-0.2803}{0.0385} \approx -7.281$$

$z^2=(-7.281)^2 \approx$ 53 which is approximately the χ^2 value.

The absolute z-value 7.281 is larger than the critical $z_{0.05}$=1.96, so the H0 is rejected.

The percentage contracting influenza was significantly lower in the vaccine group (8.3%), than in the placebo group (36.4%).

Article table

	IgG positive (*n* = 68)	IgG negative (*n* = 1543)	p value
Age, median value (IQR)	74.0 (70.3–80.0) 4	58 (42.0–71.0) 76	<0.001
Missing data			
Sex			0.036
Male	39 (5.4%)	681 (94.6%)	
Female	29 (3.3%)	857 (96.7%)	
Missing data	0	5	
Occupation			<0.001
Farmer	50 (6.6%)	706 (93.4%)	
Other	11 (1.9%)	562 (98.1%)	
Missing data	7	275	
Contact with animals			0.020
Yes	46 (5.6%)	777 (94.4%)	
No	22 (3.1%)	687 (96.9%)	
Missing data	0	79	
Contact with goats			0.001
Yes	21 (8.8%)	219 (91.3%)	
No	47 (3.6%)	1245 (96.4%)	
Missing data	0	79	
Contact with sheep			<0.001
Yes	26 (11.6%)	199 (88.4%)	
No	42 (3.2%)	1265 (96.8%)	
Missing data	0	79	
Contact with cows			<0.001
Yes	12 (12.1%)	87 (87.9%)	
No	56 (3.9%)	1377 (96.1%)	
Missing data	0	79	
Contact with dogs			0.096
Yes	11 (2.9%)	369 (97.1%)	
No	57 (4.9%)	1095 (95.1%)	
Missing data	0	79	

Sidira T, et al. Seroepidemiological study of Crimean- Congo haemorrhagic fever in Greece, 2009–2010. Clinical Microbiology and Infection, Volume 18 Number 2, February 2012

Assumptions of chi-square test

- Independent observations
  - Each individual is represented only once in the study. The rows (and columns) of the table are mutually exclusive, implying that each individual can belong in only one row and only one column.
- Large expected frequencies.
  - A) All **expected counts** should be > 1.
  - B) At least 80% of **expected counts** should > 5.

Fisher's exact test when there are small samples

Assumptions of chi-square test

- Independent observations
  - Each individual is represented only once in the study. The rows (and columns) of the table are mutually exclusive, implying that each individual can belong in only one row and only one column.
- Large expected frequencies.
  - A) All **expected counts** should be > 1.
  - B) At least 80% of **expected counts** should > 5.

Fisher's exact test when there are small samples

Example with Fisher's exact test

- When expected counts <5

	Bleeding complications		
Treatment regimen	Yes	No	Total
A (group 1)	1	12	13
B (group 0)	3	9	12
Total	4	21	25

- **Step 1. Define the null and alternative hypotheses under study**
  - **H0:** There is no relationship between the treatment regimens and the bleeding complications.
  - **H1:** There is a relationship between the treatment regimens and the bleeding complications.
- Fisher's exacte test P= **0.3217 > 0.05** do not reject H0

Vaccination & Influenza example

	Influenza		
	Yes	No	Total
Vaccine	20 (8.3%)	220 (91.7%)	240
Placebo	80 (36.4%)	140 (63.6%)	220
Total	100 (21.7%)	360 (78.3%)	460

2-proportions

H0: $p1 - p2 = 0$

H1: $p1 - p2 \neq 0$

Chi-Square

H0: There is no association between vaccination and influenza.

H1: There is an association between vaccination and influenza.

- The table shows that the percentage contracting influenza was much lower in the vaccine group (8.3%), than in the placebo group (36.4%).

- p1-p2 = 0.083-0.364 = **-0.281**

-The risk of influenza was 28% lower in the vaccine group than in the placebo group (P<0.001).

Vaccination & Influenza example

	Influenza		
	Yes	No	Total
Vaccine	20 (8.3%)	220 (91.7%)	240
Placebo	80 (36.4%)	140 (63.6%)	220
Total	100 (21.7%)	360 (78.3%)	460

H_0: RR=1
H_1: RR≠1

RR = 0.08 / 0.36= **0.22**

-The vaccine group has 0.22 times the risk (of influenza) of the placebo group
-The vaccine group has (0.22-1=-0.78) 78% lower risk of influenza than the placebo group

95%CI: 0.15, 0.36 **reject H_0**

H_0: OR=1
H_1: OR≠1

OR = (20*140) / (220*80) = **0.16**

-The vaccine group has 0.16 times the odds (of influenza) of the placebo group
-The vaccine group has (0.16-1=-0.84) 84% lower odds of influenza than the placebo group

95%CI: 0.09, 0.27 **reject H_0**

Attendance in day care & respiratory infection

- *Is attendance at a day care associated with an increase in respiratory infections?*

	Respiratory infection					
	Yes	%	No	%	Total	%
Daycare	27	30.3	62	69.7	89	100
No daycare	53	9.6	499	90.4	552	100
Total	80	12.5	561	78.3	641	100

H_0: RR=1
H_1: RR≠1

RR = 0.303 / 0.096 = **3.16**

-The risk of respiratory infection is 3.16 times higher in children who attended daycare than in those who didn't.

H_0: OR=1
H_1: OR≠1

OR = (27*499) / (62*53) = **4.10**

-The odds of respiratory infection is 4.10 times higher in children who attended daycare than in those who didn’t.

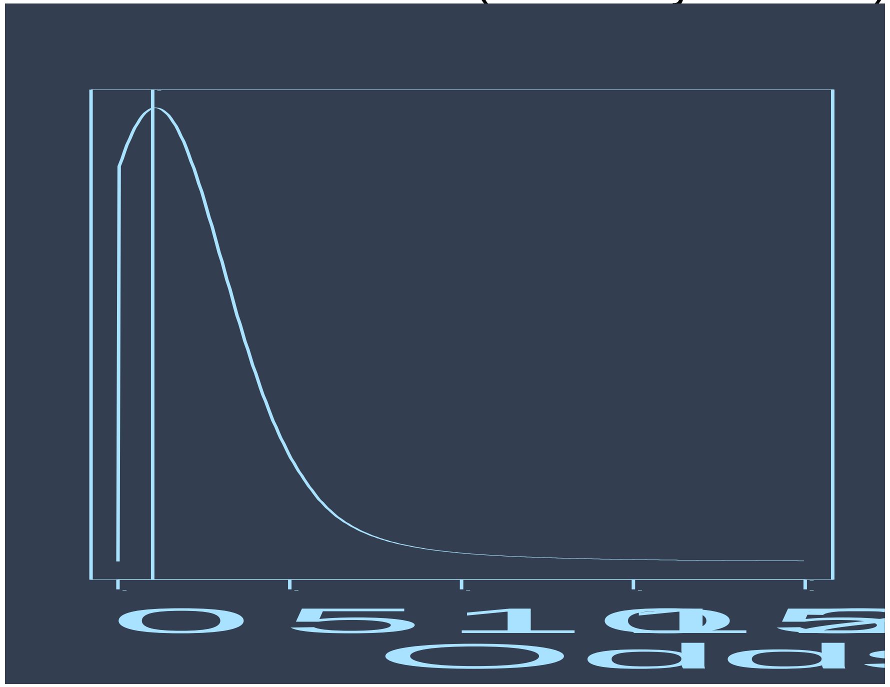
H_0: p1-p2=0
H_1: p1-p2≠0

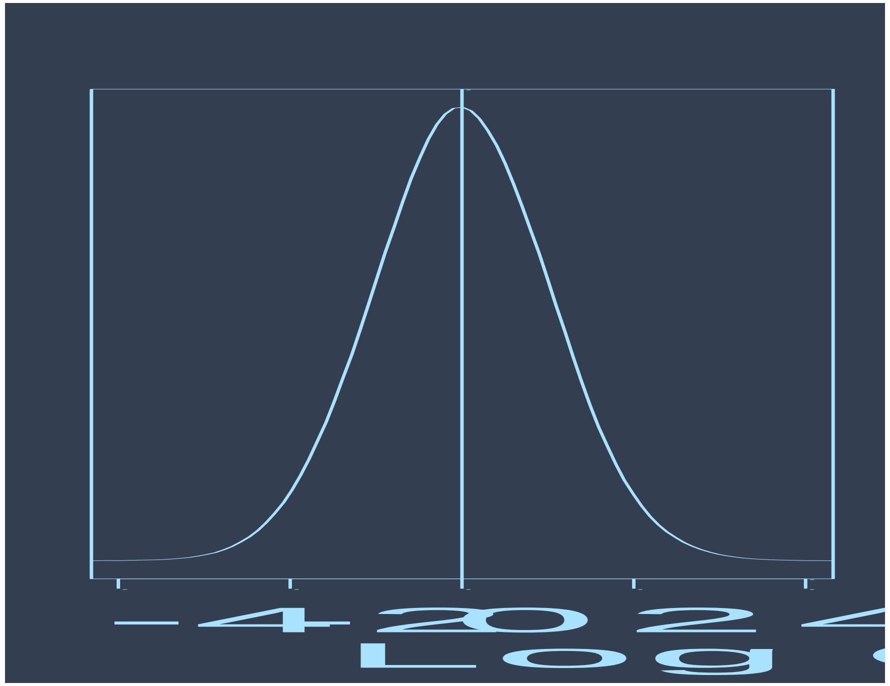
Risk difference 0.303 – 0.096 = **0.207**

The risk of respiratory infection is 21% higher in children who attended daycare than in those who didn't.

More on the Odds Ratio

- Sometimes, we see the log odds ratio instead of the odds ratio.
- Tends to be skewed (i.e. not symmetric)





log OR > 0: increased risk

log OR = 0: no difference in risk

log OR < 0: decreased risk

Use for the calculation of the 95% confidence intervals

95% confidence intervals OR

- $(1-a)\%$ $CI\ lnOR: lnOR \pm Z\_a\ \times SE(lnOR)$

- where $SE(logOR)= \sqrt{\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d}}$

Exposure	Respiratory infection Yes	No	Total
Yes	27	62	89
No	53	499	552
Total	80	561	641

- $SE(logOR)$= √(1/27+ 1/53+ 1/62+ 1/499)=0.272
- lnOR = ln(4.10) = 1.41
- 95% confidence interval of lnOR = lnOR ± 1.96*SE(lnOR) = 1.41 ± 1.96*0.272
- Lower limit = 1.41 – 1.96*0.272 = 0.878
- Upper limit=1.41 + 1.96*0.272 = 1.944

95% confidence intervals OR

- Exponentiating: $e^{0.878} = 2.41$, $e^{1.944} = 6.99$

- H_0: OR=1
- H_1: OR≠1

- Therefore the 95% confidence interval of the odds ratio 4.10 was 2.41 – 6.99, and this association is significant since the confidence interval does not include the value of 1 ($P<0.05$).

95% confidence intervals RR

- Similarly,

$$SE(logRR) = \sqrt{\frac{1}{a} - \frac{1}{a+b} + \frac{1}{c} - \frac{1}{c+d}}$$

$$SE(logRR) = \sqrt{\frac{1}{27} - \frac{1}{89} + \frac{1}{53} - \frac{1}{552}} = 0.207$$

	Respiratory infection		Total
Exposure	Yes	No	
Yes	27	62	89
No	53	499	552
Total	80	561	641

H_0: RR=1
H_1: RR≠1

- lnRR = ln(3.16) = 1.15
- 95% confidence interval of lnRR = lnRR ± 1.96*SE(lnRR) = 1.15 ± 1.96*0.207
- Lower limit = 1.15 – 1.96*0.207 =0.745
- Upper limit=1.15 + 1.96*0.207 = 1.556

- The 95% confidence interval of the risk ratio is given by exponentiating the upper and the lower limits: $e^{0.745}$ = 2.11, $e^{1.556}$= 4.74

- Therefore the 95% confidence interval of the risk ratio 3.16 was 2.11 - 4.74, and this association is significant since the confidence interval does not include the value of 1 (P<0.05).

Which measure is best?

- RR are usually calculated for cross-sectional studies and cohort studies.
- OR are always calculated for case-control studies and often for cross-sectional studies.
- Note that in case-control studies it is not generally possible to calculate risks, and therefore RR
- OR is a consistent measure of effect, unlike the RR. (1/0.16 = 6.25 higher odds of influenza in placebo vs. vaccine group)
- OR is difficult to understand

Why do we so often see OR and not RR?

(1) Logistic regression:

- Allows us to look at association between two variables, adjusted for other variables. (Linear models course)

(2) Can be more globally applied. Design of study does not restrict usage.

Article table

Table 3. Associations of meta-analysis characteristics with focus on specific agent(s)

	Focus on specific agent(s), *n* (%)		Univariable analysis		
Characteristics	**Yes**	**No**	**OR**	**95% CI**	***P*-value**
Publication venue					
Journal	157 (48.3)	168 (51.7)	1.92	1.31–2.82	0.001
Cochrane	57 (32.8)	117 (67.2)	1.00		
Funding					
Industry involved	37 (75.5)	12 (24.5)	5.28	2.62–10.66	<0.001
Not reported	60 (39.0)	94 (61.0)	1.09	0.72–1.66	0.676
No funding	30 (50.0)	30 (50.0)	1.71	0.97–3.03	0.065
No industry involved	87 (36.9)	149 (63.1)	1.00		

As shown in Table 3, in univariable analyses, meta-analysis publications focusing on specific named agent(s) were more likely to be published in journal venues rather than the *Cochrane Database of Systematic Reviews*, they were significantly more common when there was industry involvement, when harms only were addressed, and when individual patient data analyses were presented. Type of

A.-B. Haidich et al. / Journal of Clinical Epidemiology 66 (2013) 371–378

Mc Nemar's test

- Paired categorical variables
  - Before-after studies
  - Cross-over studies

		After treatment		Row Total
		present	absent	
Before treatment	present	a	b	a+b
	absent	c	d	c+d
Column Total		a+c	b+d	N=a+b+c+d

- Matched categorical data
  - Matched case-control study

		Control		Row Total
		exposed	unexposed	
Case	exposed	a	b	a+b
	unexposed	c	d	c+d
Column Total		a+c	b+d	N=a+b+c+d

Mc Nemar's test

- **Assumptions**
  - The sample data have been randomly selected and consist of matched pairs.
  - Large samples and b+c>25

  (in the case that b+c<25, the **exact binomial probability** should be computed).

Steps for conducting the test

- **Step 1. Define the null and alternative hypotheses under study**
  - **H_0**: The two marginal probabilities for each outcome are the same, $p_a + p_b = p_a + p_c$ and $p_c + p_d = p_b + p_d$ which means that $p_b = p_c$.
  - **H_1**: The two marginal probabilities for each outcome are not the same, $p_a + p_b \neq p_a + p_c$ and $p_c + p_d \neq p_b + p_d$ which means that $p_b \neq p_c$.

		After treatment		Row Total
		present	absent	
Before treatment	present	a	b	a+b
	absent	c	d	c+d
Column Total		a+c	b+d	N=a+b+c+d

Mc Nemar's test

		Control		Row Total
		exposed	unexposed	
Case	exposed	a	b	a+b
	unexposed	c	d	c+d
Column Total		a+c	b+d	N=a+b+c+d

Step 2. Calculate the value of the test statistic specific to H_0

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

Step 3. Compare the value of the test statistic to values from the χ^2- distribution

- If $\chi^2 > 3.84$ (which means that p-value<0.05) then H_0 is rejected.

Step 4. Interpretation

Children's knowledge of asthma management

- In a survey 86 children with asthma attended a camp to learn how to self-manage their asthmatic episodes

		Knowledge at end of camp		Row Total
		Yes	No	
Knowledge at beginning of camp	Yes	24	6	30
	No	29	27	56
Column Total		53	33	N=86

Step 1. Define the null and alternative hypotheses under study

H_0: there was no change in children's knowledge of asthma management between the beginning and completion of the health camp.

H_1: there was change in children's knowledge of asthma management between the beginning and completion of the health camp.

Children's knowledge of asthma management

Step 2. Calculate the value of the test statistic specific to H_0

$$\chi^2 = \frac{(6-29)^2}{6+29} = \frac{529}{35} = 15.11$$

Step 3. Compare the value of the test statistic to values from the χ^2-distribution

X^2 =15.11 > 3.84 ➡ H_0 is rejected ➡ $P<0.05$

Children's knowledge of asthma management

		Knowledge at end of camp		Row Total
		Yes	No	
Knowledge at beginning of camp	Yes	24	6	30
	No	29	27	56
Column Total		53	33	N=86

Step 4. Interpretation

- (30/86)x100=34.9% of children had appropriate knowledge at the beginning of the camp and (53/86)x100=61.6% at the end of the camp.
- The analysis show that the number of children who knew how to manage their asthma episodes appropriately increased significantly on completion of the camp.

Children's knowledge of asthma management

		Knowledge at end of camp		Row Total
		Yes	No	
Knowledge at beginning of camp	Yes	24	6	30
	No	29	27	56
Column Total		53	33	N=86

Step 4. Interpretation

- (30/86)x100=34.9% of children had appropriate knowledge at the beginning of the camp and (53/86)x100=61.6% at the end of the camp.
- The analysis show that the number of children who knew how to manage their asthma episodes appropriately increased significantly on completion of the camp.

Article table

24-h Intraocular pressure control with evening-dosed travoprost/timolol, compared with latanoprost/timolol, fixed combinations in exfoliative glaucoma

AGP Konstas[1], DG Mikropoulos[1], TA Embeslidis[1], AT Dimopoulos[1], A Papanastasiou[1], A-B Haidich[2] and WC Stewart[3]

Eye (2010) 24, 1606–1613

Patients

Patients were recruited for this prospective, single-masked, crossover study from the Glaucoma Unit of the First University Department of Ophthalmology, AHEPA

Table 4 Adverse events recorded in the study

Adverse eventr[a]	*LTFC*	*TTFC*			*P-value*
		Yes	*No*	*Total*	
Conjunctival hyperemia	Yes	2	2	4 (10)	0.29
	No	6	30	36 (90)	
	Total	8 (20)	32 (80)	40 (100)	
Stinging	Yes	3	3	6 (15)	0.62
	No	1	33	34 (85)	
	Total	4 (10)	36 (90)	40 (100)	
Hypertrichosis	Yes	3	0	3 (7)	0.25
	No	3	34	37 (93)	
	Total	6 (15)	34 (85)	40 (100)	
Skin pigmentation	Yes	0	0	0 (0)	0.25
	No	3	37	40 (100)	
	Total	3 (7)	37 (93)	40 (100)	

Flowchart of statistical tests