

Measures of Association in Epidemiology

Dr. Sireen Alkhaldi

Community Medicine 2023/ 2024

Faculty of Medicine, The University of Jordan



21% of boys and 30% of girls support me; therefore I'll get 51% of the vote.



Measures of Association in Epidemiology

- Chi square
- Odds Ratio (OR)
- Relative Risk or Risk Ratio (RR)
- Attributable Risk (AR)



2X2 Table (contingency table)

		Disease		
		Yes (+)	No (-)	Total
Exposure	Yes (+)	a	b	a+ b
	No (-)	c	d	c+ d
Total		a+ c	b+ d	a+b+c+d



Cells

Event forecast	Event observed		
	Yes	No	Marginal total
Yes	a	b	a + b
No	c	d	c + d
Marginal total	a + c	b + d	a + b + c + d = n

A= Exposed, and diseased

B= Exposed, Not diseased

C= Not exposed, diseased

D= Not exposed, Not diseased

A+B+C+D=Total



Totals

Event forecast	Event observed		
	Yes	No	Marginal total
Yes	a	b	a + b
No	c	d	c + d
Marginal total	a + c	b + d	a + b + c + d = n

Marginal totals

$a+b =$ Exposed

$c+d =$ Non-exposed

$a+c =$ Diseased

$b+d =$ Non-diseased

Grand total

$$n = a+b+c+d$$

1. Chi-square in Cross-sectional studies

Chi-square tests whether there is an association between two categorical variables.

يختبر مربع كاي ما إذا كان هناك ارتباط بين متغيرين قاطعين.

For a 2X2, table:

$$X^2 = \frac{n(ad - bc) - n/2)^2}{(a+b)(a+c)(c+d)(b+d)}$$

إذا كانت قيمة مربع كاي المحسوبة أكبر من القيمة الحرجة أو $P < 0.05$ ، فإننا نقول أن هناك ارتباطاً كبيراً بين عامل الخطر والمرض (يتم حسابه عادةً باستخدام برامج مثل SPSS أو Excel).

تخبرنا إحصائية مربع كاي فقط ما إذا كان هناك ارتباط أم لا. لا يخبرنا بمدى قوة الارتباط.

If the calculated chi-square value is greater than the critical value or $P < 0.05$, we say that there is a significant association between the risk factor and the disease (usually calculated using software like SPSS or excel).

Chi-square statistic tells only whether there is association. It doesn't tell us how strong an association is.



2. Relative risk (RR) or Risk Ratio (RR) In a cohort study

RR: The estimation of disease risk associated with exposure (strength of association)

تقدير مخاطر الأمراض المرتبطة
بالتعرض (قوة الارتباط)

يعبر RR عن خطر الإصابة بمرض في
المجموعة المعرضة (أ + ب) مقارنة
بالمجموعة غير المعرضة (ج + د)

RR Expresses risk of developing a disease in exposed group (a + b) as compared to non-exposed group (c + d)

RR= Incidence (risk) among exposed
Incidence (risk) among non-exposed

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

Event forecast	Event observed		
	Yes	No	Marginal total
Yes	a	b	a + b
No	c	d	c + d
Marginal total	a + c	b + d	a + b + c + d = n

Analysis in Cohort studies

In a Cohort Study, we can calculate Incidence.

So, Relative Risk can be obtained from a cohort study.

في دراسة الكوهورت ، يمكننا حساب معدل الإصابة. لذلك، يمكن الحصول على المخاطر النسبية من دراسة الكوهورت.

Cigarette smoking (Exposure)	Disease (with Ca lung)	No Disease (without Ca lung)	Total
Yes	70 (a)	6930 (b)	7000 (a+b)
No	3(c)	2997(d)	3000 (c+d)



RR in a Cohort Study

Incidence rates :

Risk among exposed (smokers) = $70/7000 = 10 / 1000$.

Risk among non-exposed(non smokers) = $3/3000 = 1 / 1000$.

$RR = \frac{\text{Risk (Incidence) among exposed}}{\text{Risk (Incidence) among non exposed}}$.

RR = 10/1=10

The exposed have 10 times the risk of developing the disease when compared to non-exposed

يتعرض الأشخاص المعرضون لخطر الإصابة بالمرض
10 أضعاف مقارنة بغير المعرضين



Interpretation of relative risk

ماذا يعني RR من 2؟

خطر التعرض = 2X خطر عدم التعرض

وبالتالي فإن الخطر النسبي 2 يعني أن
المجموعة المعرضة معرضة لخطر الإصابة
بالمرض بمقدار الضعف مقارنة بغير المعرضين

What does a RR of 2 mean?

Risk in exposed=2X Risk in non-exposed

Thus a relative risk of 2 means the exposed group is two times at a higher risk of developing the disease when compared to non-exposed



Strength of association

In general **strength of association** can be considered as:

High association if $RR \geq 3$

Moderate if RR is between 1.5 & 2.9

Weak association if RR is between 1.2 & 1.4

No association exists if RR is 1

Negative association (protective effect) if $RR < 1$



3. ODDS RATIO (OR)

Odds Ratio (OR) is a measure of the strength of the association between risk factor & outcome.

- ✓ The odds ratio is the cross product of the entries in table.
- ✓ OR can be calculated in case-control studies instead of RR

2 by 2 table

	Diseased - Cases	Non-diseased - Controls	Total
Exposed	A	B	A+B
Non-exposed	C	D	C+D
Total	A+C	B+D	A+B+C+D

نسبة الأرجحية (OR) هي مقياس لقوة الارتباط بين عامل الخطر والنتيجة.

* نسبة الأرجحية هي المنتج الاتجاهي للإدخالات في الجدول.

* يمكن حساب OR في دراسات الحالات والشواهد بدلاً من RR



Odds ratio (OR)

يمكن أن تكون نسبة الأرجحية تقديراً جيداً لـ RR.
نسبة الأرجحية هي نسبة احتمالات التعرض بين المرضى إلى
احتمالات التعرض بين غير المرضى

Odds Ratio can be a good estimate of RR.

Odds ratio is the ratio of odds of exposure among diseased to odds of exposure among non-diseased

$$\text{OR} = \frac{\text{Odds of exposure among diseased}}{\text{Odds of exposure among non-diseased}}$$
$$= (a/c)/(b/d) = ad/bc$$

تفسير OR هو نفس تفسير RR

Interpretation of OR is the same as that of RR



Odds ratio...

يمكن تقدير RR بشكل أفضل بواسطة OR إذا تم استيفاء الشروط التالية:

1. الضوابط تمثل عامة السكان

2. الحالات المختارة تمثل جميع الحالات

3. المرض نادر

RR can be best estimated by OR if the following conditions are fulfilled:

- 1. Controls are representative of general population**
- 2. Selected cases are representative of all cases**
- 3. The disease is rare**



ANALYSIS in case-control studies

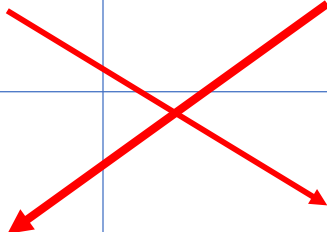
تقدير احتمالات التعرض بين المعرضين (نسبة الأرجحية)

Estimation of the odds of exposure among the exposed (odds ratio)



OR in a case-control study

	Cases (with Ca lung)	Controls (without Ca lung)
Smokers	33 (a)	55 (b)
Non smokers	2 (c)	27 (d)
Total	35 (a+c)	82 (b+d)



OR in Case-control studies

Odds ratio is a key parameter in the analysis of Case-control studies

نسبة الأرجحية هي معلّمة أساسية في تحليل دراسات الحالات والشواهد

	Disease	
	Cases (Lung Ca)	Controls (No Lung Ca)
Smokers	<i>a</i> (33)	<i>b</i> (55)
Not smokers	<i>c</i> (2)	<i>d</i> (27)

$$\text{Odds Ratio} = ad/bc = \frac{33 \times 27}{55 \times 2} = 8.1$$



Attributable Risk (AR)

ويشير AR إلى مقدار المخاطر الناجمة عن (المنسوبة إلى) التعرض.
يحدد حجم المخاطر الزائدة في المكشوفة التي يمكن أن تكون التي تعزى إلى التعرض، عن طريق إزالة خطر المرض الذي حدث لأسباب أخرى.
الخطر (الحدوث) في المكشوفة- الخطر (الحدوث) في غير المكشوفين = AR

AR indicates how much of the risk is due to (attributable to) the exposure.

Quantifies the excess risk in the exposed that can be attributable to the exposure, by removing the risk of the disease that occurred due to other causes.

AR= Risk (incidence) in exposed- Risk (incidence) in non-exposed

$$AR = [a/(a+b)] - [c/(c+d)]$$

Attributable risk is also called risk difference. وتسمى المخاطر المنسوبة أيضاً بفارق المخاطر.



Attributable risk percent (AR%)

تقدير نسبة المرض بين المعرضين والتي تعزى إلى التعرض (على سبيل المثال التدخين والرئة C).

Estimates the proportion of disease among the exposed that is attributable to the exposure (example of smoking and lung C).

$$AR\% = \frac{\text{Risk in exposed} - \text{Risk in non-exposed}}{\text{Risk in exposed}} \times 100\%$$

$$AR\% = \frac{10 - 1}{10} \times 100\% = 90\% \text{ (as in the previous Cohort study example)}$$

90% من حالات سرطان الرئة بين المدخنين كانت بسبب تدخينهم.

يشير هذا إلى حجم المرض الذي يمكن القضاء عليه إذا أمكن السيطرة على التدخين أو القضاء عليه.

- ❑ 90% of the lung cancer among smokers was due to their smoking.
- ❑ This suggests the amount of disease that might be eliminated if smoking could be controlled or eliminated.



Possible outcomes in studying the relationship between exposure & disease

النتائج المحتملة في دراسة العلاقة بين التعرض والمرض

1. No association

$$RR=1$$

$$AR=0$$

2. Positive association

$$RR>1$$

$$AR>0$$

3. Negative association

$$RR<1 \text{ (fraction)}$$

$$AR<0 \text{ (Negative)}$$



Risk Vs Preventive factors

عامل الخطر هو أي عامل يرتبط بشكل إيجابي بالمرض ($RR > 1$). ويرتبط بزيادة حدوث المرض.
العامل الوقائي هو أي عامل يرتبط سلباً بالمرض ($RR < 1$). ويرتبط مع انخفاض حدوث المرض.
قد (لا) تكون عوامل الخطر والعوامل الوقائية قابلة للتغيير (مثل التدخين والعمر)

A **risk factor** is any factor positively associated with a disease ($RR > 1$). It is associated with an increased occurrence of a disease

A **preventive factor** is any factor negatively associated with a disease ($RR < 1$). It is associated with a decreased occurrence of a disease.

Risk and preventive factors **may (not)** be amenable to change (e.g. Smoking, age)

