


## Types of studies

* Last time we talked about introduction to epidemiology, starting from the definition of epidemiology, studying the distribution, and the determinants of disease in population, and we talked a little bit about the distribution of disease in relation to person, place \& time, talked about 3 types of variations of morbidity or/and mortality according to time.
- Determinant factor: increase or decrease of occurrence of disease in human population.
- Risk factor: the factor which is significantly associated with occurrence of a disease, which means that $P$ value is less than or equal 0.05 .
$P$ value is calculated according to statistical tests.
$\rightarrow$ So if $\mathrm{P}<=0.05 \rightarrow$ the factor is a risk factor.
Ex.: If we have a study about smoking \& bronchitis, \& the $P$ value of that study was found to be very small, say 0.0005 , Now, the question is, does smoking cause bronchitis? Ans: We can't judge if smoking causes bronchitis or not, except if we go back to criteria of causation.


## Criteria of causation:

- Some call them Bradford Hill's criteria of causation (established by Bradford Hill).
- They are 9 , we are interested here in 5 criteria of causation :
- First 5 criteria of causation (also termed as causality, cause-effect relationship) are :

1. The type of epidemiological study: they are 3 types of those studies that investigate causation or causality: 1. Case control studies.
2. Cohort (follow up studies) دراسات تتبعية
3. Clinical trials: doing experiments on individuals to find out either:
4. Causality or causation.
5. Efficacy or efficiency of intervention.

Ex: if you have a new antibiotic and want to find how deficient is this antibiotic, you can make a clinical trial دراسة تجريية طىى الانسان, to be able to say that the efficacy rate of this new antibiotic is $90 \%$ (for example), means that this antibiotic is able to cure $90 \%$ of cases of that disease.
2. Dose response or dose effect relationship:

Dose means: quantity of the risk factor.*Effect or response is the outcome of a disease.

Ex: in case of smoking and bronchitis, if number of cigarettes smoked daily is increasing, there will be a higher probability of having bronchitis.
3. Temporality or time sequence of both the risk factor and the disease. -the risk factor should come before the onset of the disease.
Ex: to be able to say that your colleague is having bronchitis due to smoking, he should be a smoker before the onset of bronchitis.
4. Amount or degree of association between disease and risk factor: in our example: degree of association between bronchitis (disease) \& smoking (risk factor).
-can be measured by relative risk and odds ratio.
-relative risk is used in Cohort studies دراسات تتبعية while odds ratio is used in case of Case Control studies.
5. Consistency : تطابق نتائج در استكا مع نتائج الار اسات الأخرى

Ex: if you conduct a study on smoking and bronchitis and you find out that the relative risk is 4 , then go back to other studies of other parts of the world to find out what is their relative risk. If most of studies are having relative risk between 35 you will be able to say that my finding is consistent with findings of other studies.
$\rightarrow$ If you find that your study is consistent with almost all other studies except few studies, this will increase the probability that the risk factor causes that disease.
-This part of the research is the most part that takes time from us (till you get your results, you need several days, and then, you'll have to look for other studies \& compare your result with).

Those 5 criteria are the ones that are actually needed in any research.

## * Risk measurement :

- Risk means: probability of having a bad outcome which could be either morbidity or mortality.
- It can be measured, Measures of risk can be classified into 4 types :

1. Absolute measures of risk:

Include: 1. Prevalence rate 2. Incidence rate 3. Attack rate of the disease.
$\rightarrow$ The presence of 3 different terms means that there are differences between them.
2. Relative risk and Odds ratio :
-both of them measure the degree of association between risk factor and disease.
3. Attributable risk percentage: ( population attributable risk percentage )

Ex: -assume that you are "lecturing" some smokers about the risks of smoking: 1.lung cancer 2.bronchitis 3.cardiovascular diseases.
-One of these smokers asked you what he should do?? Your answer is to stop smoking .... Then he asked if he stops smoking how much the risk of lung cancer will decrease?? $\rightarrow$ Then he is talking about "Attributable risk percentage". You'll answer that it will decrease based on certain studies (AR \%) by 20\% 4 example.

- Regression or correlation coefficient : "R" value معامل الانحدار و الارتباط
- It reflects the degree of association/correlation between the disease \& a certain factor.
- $\quad R$ value range in any study is between -1 to +1 .
- If any value is above zero $\rightarrow$ then it's positive association (positive regression or correlation).
But if below zero $\rightarrow$ then it's negative association.
- Positive association means that if one variable increases the other variable must increase

Ex: medically, if the number of cigarettes smoked per day increases, the risk of bronchitis will increase (same direction).

- While negative association means that if one variable increases the other will decrease.

Ex: 1. if the income increases the morbidity will decrease.
2. Women who use oral contraceptive methods the risk for them to have osteoarthritis will decrease, although, oral contraceptive methods are one of the risk factors to so many health problems.
Ex: if you read an article about smoking \& bronchitis and there were these 2 values ( $\mathbf{R}=\mathbf{0 . 6}, \underline{P=0.01}$ ), what would you understand from this? This means that there is a significant positive association between cigarettes \& bronchitis.
Ex: Another article is talking about oral contraceptive methods \& osteoarthritis, and you found these 2 values ( $\mathbf{R}=\mathbf{- 0 . 4}, \underline{P=0.01}$ ), what does this mean? That there is a significant negative correlation between oral contraceptive methods \& osteoarthritis.

- As the value of $R$ becomes closer to zero $\rightarrow$ the correlation will becomes more likely to be not significant, and $P$ value will increase, at $R=0, P$ is almost 1 (0.9999)
- So if $R$ increases to more positive $\rightarrow$ the $P$ value will decrease (very high significance).
- And if $R$ decreases to more negative $\rightarrow$ the $P$ value will also decrease (very high significance).


## * Prevalence and incidence rate of a disease :

- How to calculate them of a disease??
- Any rate can be calculated by using the formula of ((X/Y)*K).

While: X : number of cases of certain disease.
Y : the sample that you are studying.
K: a constant, could be 100, 1000, 10000,100000 ...etc.
$\rightarrow$ For this constant, if you remember the infant mortality rate is always per 1000, maternal mortality rate per 100000, crude death rate per 1000, for rates regarding diseases k is always 100 (ex.: if there was an outbreak of a disease in Jordan \& you wanted to know the attack rate, you'll call the MOH \& they'll tell you that it is 10 for example, this means 10 per 100)

- Way of calculating prevalence and incidence rate :

Ex: someone says that he wants to make a study on 10 persons to calculate the prevalence and incidence rate of bronchitis in December (from 1 December to 31 Dec ) , he took each individual and asked him if he has bronchitis ,
$\rightarrow$ the first said that : yes, from few days and I still have bronchitis ( had bronchitis before 1 December/ before the onset of the study) ..... And we followed him up for the whole month \& he was cured during that month.
$\rightarrow$ The second says that: l'm not sick, but through the study, after 1 week of Dec $1^{\text {st }}$, he got sick, and he is cured within few days.
$\rightarrow$ The third says that: no I'm not sick.... And he was followed up for the whole month \& he didn't have bronchitis at all.
$\rightarrow$ The fourth say that: I had bronchitis before 1 December... and he still has bronchitis for $3 / 4$ the month.
$\rightarrow$ The fifth says that: bronchitis has developed during the month.
And so on....
*Here is a presentation of the 4 cases that had bronchitis just like the Dr. drew them in the lecture: where the line ---- means "having bronchitis":


Now :- all cases of bronchitis that started before December ( before the onset of the study ) \& lasted till the $1^{\text {st }}$ of Dec or after, are called old or ongoing cases $\rightarrow$ individual number 1,4 ( as in the example above ), so 2 old cases.
-all cases that developed during the month (study) are called new case $\rightarrow$ individual number 2, 5, so 2 new cases.
The study is finished with 2 old cases \& 2 new cases of bronchitis and the rest (6) didn't develop bronchitis before or during the study.
-prevalence rate $=($ old cases + new cases $/$ study population (sample)) * 100\% $\rightarrow(2+2 / 10) * 100 \%=40 \%$ (in the example above).
-incidence rate $=($ new cases $/($ sample size (N) - old cases)) * 100\% $\rightarrow(2 /(10-2)) * 100 \%=25 \%$.

- Notes: - always in incidence rate we look for new cases of disease.
-prevalence rate appears/comes out from Cross Sectional studies, while incidence rate appears/comes out from Cohort Studies and clinical trials.

Q: Is there a relationship between Prevalence rate \& Incidence rate? -they found in studies that prevalence rate of a disease almost equals the incidence rate of the same disease multiplied by D .
$\rightarrow$ Prevalence rate = incidence rate * $D$
D: average duration of the disease from the start of signs and symptoms to either to be cured or died.
(From example) $\rightarrow 40 \%=25 \%{ }^{*} \mathrm{D} \rightarrow \mathrm{D}=1.6$.
-D value differs, in case of bronchitis the $D$ value could be in weeks, in salmonella the $D$ value could be in hours, some diseases like tuberculosis $D$ value could be in months, other chronic diseases like diabetes $D$ value could be in years .
*In this formula $P . R=I . R * D$, if you know 2 values, you are able to know the third (trust me... it's a Q in the final ;) peace of cake).

The greatest pleasure in life is doing what people say you cannot do. :D

