جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

THUMBAY
Growth Through Innovation

Data analysis

Research made easy

Continuing Quality Education Deries - CAHQ

Dr. Jayadevan Sreedharan
Professor of Epidemiology and Biostatistics
Program Director - MPH
Gulf Medical University, Ajman, United Arab Emirates
Adjunct Professor- University of Arizona, USA

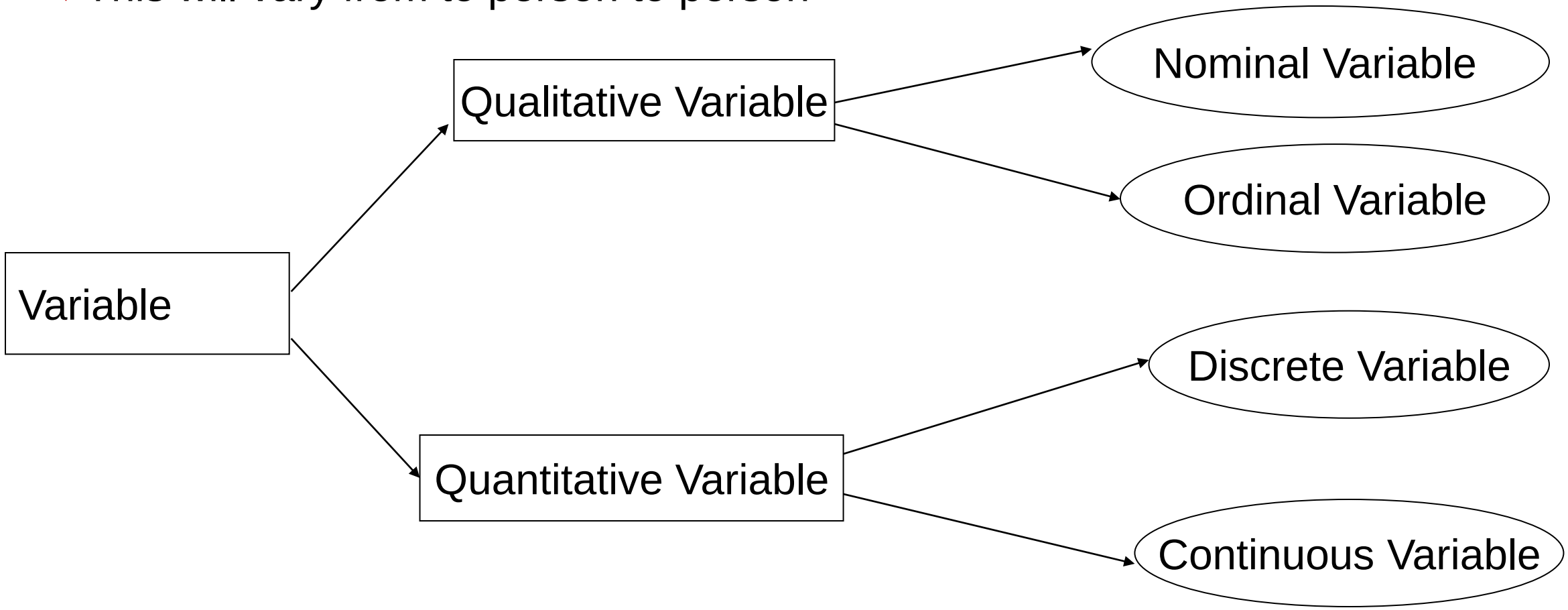
June 25, 2022

- Variables
- Descrptive statistics
- Inferential statistics

Variables

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

- The item or characteristic on which observations are made.
- This will vary from to person



Independent and Dependent Variables

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Independent variable

- One which influences the dependent variable.
- This is otherwise called explanatory variable.
- Variable that is varied or manipulated by the researcher.

Dependent variable

- One which depends on independent variable.
- This is otherwise called response variable.
- Variable which the response is measured.

Descriptive statistics

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Descriptive Statistics

Describe the Sample: Describe important variables like sociodemographic, confounders etc. in your analyses

- Frequency
- Percentage
- Central tendency
- Dispersion

Heart disease on rise among young expats, says study:

Indian expatriates at higher risk than others:

Published: Gulf News - October 22, 2018

- https://gulfnews.com/news/uae/health/heart-disease-on-rise-among-young-expats-says-study-1.2292529

Heart disease on rise among young expats, says study:

Indian expatriates at higher risk than others:

Published: Gulf News - October 22, 2018

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Dubai: Coronary Artery Disease (CAD) in the UAE is striking younger people in the age group of 35-45 with the incidence being highest in Asian expatriates and Indian expatriates being the most vulnerable, a recent clinical study has indicated.

The ongoing clinical study conducted by Aster Hospitals, Dubai, covers the period between mid-June to mid-September 2018 and covered 142 patients who were referred to their Cardiac Cath Laboratory with manifestation of CAD.

Of these nearly 106 patients were under the age of 55.

Nearly 66.2% of the patients were Indians with 14.2 being from Pakistan. Other expatriates were from Bangladesh, the Philippines, UK , Egypt, Sri Lanka, Nepal, Nigeria and Serbia.

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

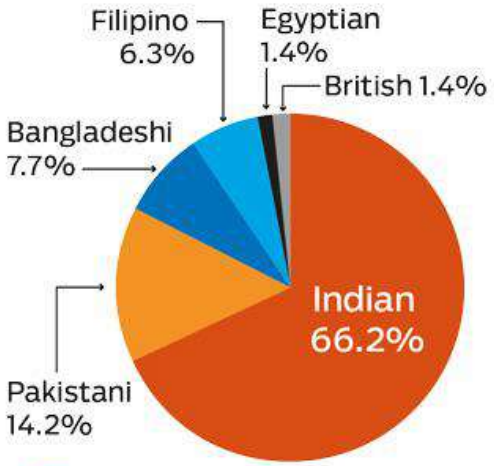
Elaborating on the nature of the data, Dr. Sachin Upadhyaya, specialist cardiologist, said: “Heart disease is striking nearly a decade earlier in people of the Indian subcontinent. This study is based on pure empirical data and it is a cause for concern that now 30-40- year-olds are stricken by CAD.”

Giving further details on the nature and scope of the study, Dr. Naveed Ahmad, specialist interventional cardiologist at Aster Hospitals, said: “The main cause for worry is the fact that the mean age for CAD in this region is much lower than the rest of the world where the high risk group falls in the age group of 55 and above. In our study of the 142 patients who were referred to us in various stages of CAD requiring surgical intervention, 57% of them were in the age group of 40-60 years, 74.6% were in less than 55 years of age group followed by 31.7% who were less than 45 years of age with only 11.3% being above the age of 61 years. Nearly 88% of patients were from India, Pakistan, Bangladesh and neighbouring countries.”

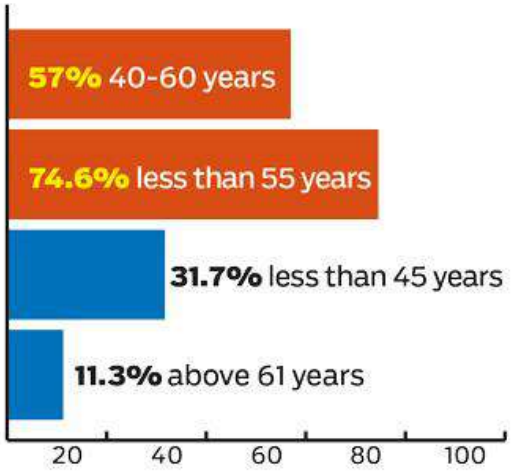
Coronary Artery Disease

- Aster Hospitals conducted a study on 142 patients referred to its cardiac cath laboratory from mid-June to mid-September 2018. Of these 106 patients were under the age of 55.

Nationality of patients



Age of patients



WHAT YOU CAN DO

- Stop smoking. Even one cigarette can put you in the high risk category.
- Reduce harmful lipids by eating more fruits and vegetables and avoiding oily, fatty food.
- Diabetics must monitor sugar intake get checked regularly.
- Make physical fitness a part of your lifestyle.
- High blood pressure is another high risk factor for CAD.
- Avoid stress and loneliness by participating in community activity
- Go in for preventive screening and take the TreadMill Test.

Source: Aster Hospitals, Dubai

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Conclusion: Coronary Artery Disease (CAD) in the UAE is striking younger people in the age group of 35-45 with the incidence being highest in Asian expatriates and Indian expatriates being the most vulnerable.

Do you agree with the conclusion of this study?

If Yes, why; if No, why

Descriptive statistics

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Data of 1000 patients with Cardiovascular Disease (CVD) showed that 800 (80%) consume non-vegetarian diet.

Conclusion: Consumption of non-vegetarian diet is a cause for CVD.

Do you agree?

Measures of Central tendency

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Common measures of central tendency:

- Arithmetic Mean (AM)
- Median
- Mode
- Harmonic Mean (HM)
- Geometric Mean (GM)

Example

- After one month fasting blood glucose level observed was

Drug A (mg/dl) : 112, 115, 117, 125, 148, 114, 117, 135, 143, 111

Drug B (mg/dl) : 118, 117, 121, 130, 128, 129, 130, 119, 122, 123

Mean Blood glucose level

Drug A - 123.7mg/dl

Drug B - 123.7mg/dl

- Which drug is good?

Example

- After three months fasting blood glucose level observed was

Drug A (mg/dl) : 112, 115, 115, 125, 148, 114, 117, 115, 143, 111

Drug B (mg/dl) : 118, 117, 121, 130, 128, 129, 130, 119, 122, 123

Mean Blood glucose level

Drug A - 121.5mg/dl

Drug B - 123.7mg/dl

- Which drug is good?

Measures of Dispersion

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Absolute measure of dispersion

- Range
- Quartile Deviation (QD)
- Mean Deviation (MD)
- Variance
- Standard Deviation (SD)

Example

Two drugs (A & B) available in the market for the treatment of DM. One clinician prescribed drug A to 10 patients and drug B to 10 patients randomly to know the effect of drugs. The mean fasting blood glucose level was 150mg/dl in two groups before the treatment.

Example

Drug A (mg/dl) : 112, 115, 115, 125, 148, 114, 117, 115, 143, 111

Drug B (mg/dl) : 118, 117, 121, 130, 128, 129, 130, 119, 122, 123

Drug	Mean	SD
A	121.5 mg/dl	13.25 mg/dl
B	123 mg/dl	5.1 mg/dl

- Drug A or Drug B better?

Measures of Dispersion

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Relative measure of dispersion

- Coefficient of range
- Coefficient of quartile deviation
- Coefficient of mean deviation
- Coefficient of Variation (CV)

Example

Drug A (mg/dl) : 112, 115, 115, 125, 148, 114, 117, 115, 143, 111

Drug B (mg/dl) : 118, 117, 121, 130, 128, 129, 130, 119, 122, 123

Drug	Mean	SD	CV
A	121.5 mg/dl	13.25 mg/dl	10.9%
B	123 mg/dl	5.1 mg/dl	4.1%

- Drug A or Drug B better?

Normal (Gaussian) Distribution

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

- If the mean, median and mode values are same, the curve is symmetrical.

- This frequency distribution is known as normal distribution

- Area under the curve in a given interval gives the probability of the value lying in that interval

'X' follows a Normal distribution, with mean μ and standard deviation σ, we write $X \sim N(\mu, \sigma^2)$

Inferential statistics

Inferential statistics

Make inferences about the population using the sample

- Hypothesis testing
- Estimation

Parametric and non-parametric tests

Two means

More than two means

Example

Drug A (mg/dl) : 112, 115, 115, 125, 148, 114, 117, 115, 143, 111

Drug B (mg/dl) : 118, 117, 121, 130, 128, 129, 130, 119, 122, 123

Drug	Mean	SD	CV
A	121.5 mg/dl	13.25 mg/dl	10.9%
B	123 mg/dl	5.1 mg/dl	4.1%

- Which test we can use here?

Original Article

DOI: 10.7860/JCDR/2013/4968.2730

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Physiology Section

The Effect of Smoking on the Hearing Status –A Hospital Based Study

ADESH KUMAR, RAJIV GULATI, SANGEETA SINGHAL, ABRAR HASAN, ASIF KHAN

ABSTRACT

Background: Tobacco smoking has been known to affect the human physiology and among the various damaging effects of tobacco, it has been linked with its effect on the sense of hearing.

Aim and Objective: This study was designed with the aim of finding the relationship between smoking and hearing loss in various age groups.

non smokers. The smoking history of all the subjects whose ages ranged from 20 to 60 years was taken in detail and their audiometric thresholds were recorded in a sound proof room by a professional audiometrist. The data was analyzed by using appropriate statistical tests.

Observation and Results: Smoking was found to be significantly associated with hearing loss. Also, the hearing loss was mainly of the sensorineural type, with the mild type (26-40 dB)

[Table/Fig-1]: ARPHI'S series 500 Portable Audiometer.

Expected cell value

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Smoking Status	Age (in years)	Affected Subjects		Hearing loss (in dB)			No Hearing loss	Total
		No	%	26-40	41-60	>60		
Smoker	**20-30**	6	26.1	6	0	0	17	23
	31-40	29	67.4	27	1	1	14	43
	41-50	21	77.8	19	1	1	6	27
	51-60	15	100.0	9	5	1	0	15
	Total	**71**	**65.7**	**61**	**7**	**3**	**37**	**108**
Non smoker	**20-30**	0	.0	0	0	0	8	8
	31-40	0	.0	0	0	0	10	10
	41-50	1	8.3	1	0	0	11	12
	51-60	5	50.0	5	0	0	5	10
	Total	**6**	**15.0**	**6**	**0**	**0**	**34**	**40**

[Table/Fig-4]: Hearing loss in relationship with age
X^2 test, p value<0.05

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Number of Cigarette/bidi Per day	Hearing Impairment						No Hearing Impairment		Total	
	26-40 db		41-60 db		>60 db					
	No	%	No	%	No	%	No	%	No	%
1-12	2	14.3	0	0.0	0	0.0	12	85.7	14	100
13-24	15	51.7	0	0.0	0	0.0	14	48.3	29	100
25-36	25	65.8	2	5.3	1	2.6	10	26.3	38	100
>36	19	70.4	5	18.5	2	7.4	1	3.7	27	25.0
Total	**61**	**56.5**	**7**	**6.5**	**3**	**2.8**	**37**	**34.3**	**108**	**100**

[Table/Fig-8]: Hearing loss in relation to number of bidi/cigarette consumed
X^2 test, $p<0.05$

Aimed to evaluate year one students' self-reported empathy levels following a 2-hour empathy workshop at a large medical school in Malaysia

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Instruments used

The study used a standardized self-reporting scale, ie, the Jefferson Scale of Physician Empathy (Student Version) JSPE-S, which is a self-report measure of medical students' attitudes towards empathy. The JSPE-S is presented to students in English and does not require any translation. It is a 20-item instrument using a seven-point Likert scale (1, strongly disagree; 7, strongly agree), with ten items that are reverse-scored. When originally developed, the JSPE-S was found to be valid and reliable.[3] A short demographic questionnaire was included.

Data analysis

SPSS version 20.0 (SPSS Inc., Chicago, IL, USA) was used for data storage, tabulation, and generation of descriptive and inferential statistics. Descriptive statistics including means, standard deviations and confidence intervals were used to summarize the demographic data. A paired repeated-measures *t*-test was used to compare before and after results, and chi-square (χ^2) tests were used to explore demographic relationships and empathy levels. All tests were two-tailed, with the results considered to be statistically significant if the *P*-value was <0.05. Effect sizes (*d*) were also calculated to quantify the differences between mean scores.

Wrong statistical test

Why Confidence Interval and p-value

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

- Confidence intervals provide information about statistical significance, as well as the direction and strength of the effect
- When a point estimate is used, the conclusions is on the basis of single value for the sample.
- CI provide a range of possible plausible values for the target population, as well as the probability with which this range covers the real value.

	OR	P	95% CI
Study -1	1.8	<0.05	1.5 – 2.1
Study -2	1.8	<0.05	1.3 – 5.7
Study-3	1.8	<0.05	1.1 – 23.5

Sample size

Multiple case control studies were conducted to determine whether there is any association between alcohol consumption and oesophageal cancer.

Studies with sample sizes 400, 500, 600, 700 and 800

Small and large sample size

A case control study was conducted to determine whether there is any association between Alcohol consumption and oesophageal cancer.

- Sample size is calculated using the formula for case control study.
- Minimum required sample size for this study was 100 cases and 300 controls.

Study was repeated with sample larger sample sizes

500 (125 cases and 375 controls)

600 (150 cases and 450 controls

700 (175 cases and 525 controls)

800 (200 cases and 600 controls)

Example

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Alcohol use	Controls		Cases		Total
	No.	%	No.	%	
No	220	73.3	65	65	285
Yes	80	26.7	35	35	115
Total	300		100		400

The crude Odds Ratio is 1.48

Odds Ratio and CI

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

	B	SE	Sig.	Exp (B)	95% CI. for EXP(B)	
					Lower	Upper
Alcohol	0.39	0.25	0.11	1.48	0.91	2.40
Constant	-1.22	0.14	0.00	0.30		

- Confidence interval for the OR contains 1, we can conclude that the observed OR is not significant (due to chance)

Sample size - 500

Alcohol	Controls		Cases		Total
	No.	%	No.	%	
No	274	73.1	81	64.8	355
Yes	101	26.9	44	35.2	145
Total	375		125		500

Odds Ratio =1.47

	B	SE.	Sig.	Exp (B)	95% CI. for EXP(B)	
					Lower	Upper
Alcohol	0.39	0.22	0.079	1.47	0.96	2.27
Constant	-1.22	0.13	0.000	0.30		

Sample size - 600

Alcohol	Controls		Cases		Total
	No.	%	No.	%	
No	328	72.9	97	64.7	425
Yes	122	27.1	53	35.3	175
Total	450		150		600

Odds Ratio =1.47

	B	S.E.	Sig.	Exp (B)	95% C.I. for EXP(B)	
					Lower	Upper
Alcohol	0.39	0.20	0.056	1.47	0.99	2.18
Constant	-1.22	0.12	0.000	0.30		

Sample size - 700

جامعة الخليج الطبية

Alcohol	Controls		Cases		Total
	No.	%	No.	%	
No	383	73.0	113	64.6	496
Yes	142	27.0	62	35.4	204
Total	525		175		700

Odds Ratio =1.47

	B	SE.	Sig.	Exp (B)	95% CI. for EXP(B)	
					Lower	Upper
Alcohol	0.39	0.19	0.035	1.47	1.03	2.13
Constant	-1.221	0.107	0.000	0.30		

Sample size - 800

جامعة الخليج الطبية

Alcohol	Controls		Cases		Total
	No.	%	No.	%	
No	438	73.0	130	65.0	568
Yes	162	27.0	70	35.0	232
Total	600		200		800

Odds Ratio =1.46

	B	SE.	Sig.	Exp (B)	95% CI. for EXP(B)	
					Lower	Upper
Alcohol	0.38	0.17	0.031	1.46	1.04	2.05
Constant	-1.22	0.10	0.000	0.297		

Significance fallacy

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

Studies were conducted to compare the effect of two drugs (A and B) on the reduction of Fasting Blood Glucose (FBG).

One clinician included 20 patients in each group and another clinician included 100 patients in each group.

Hypothesis testing

	Drug	N	Mean	SD	SE
Investigator A	Drug A	20	124.25	12.43	2.78
	Drug B	20	115.25	18.36	4.11

		t	df	Sig	Mean difference
Investigator A	Equal variances assumed	1.82	38	0.077	9.0
	Equal variances not assumed	1.82	33.38	0.078	9.0

Drug A is better than drug B

Drug B is better then drug A

The effect is same for both drugs

Hypothesis testing

	Drug	N	Mean	SD	SE
Investigator B	Drug A	100	124.25	12.17	1.22
	Drug B	100	118.10	15.07	1.51

	Variance	t	df	Sig	Mean difference
Investigator B	Equal variances assumed	3.17	198	<0.01	6.15
	Equal variances not assumed		189.60	<0.01	6.15

Drug A is better than drug B

Drug B is better then drug A

The effect is same for both drugs

Interpretation

- Not only the value of statistical significance, but also the effect size of the study.

Misconceptions

  - Significance means big.
  - Non- significance means no effect.

- Sample size large, results will be statistically significant.
- Effect size is large; sample size small –Difference will also be significant.

جامعة الخليج الطبية
GULF MEDICAL UNIVERSITY
EDUCATION • HEALTHCARE • RESEARCH

THUMBAY
Growth Through Innovation

DISCLAMER

The contents of this presentation, can be used only for the purpose of a Lecture, Scientific meeting or Research presentation at Gulf Medical University, Ajman.