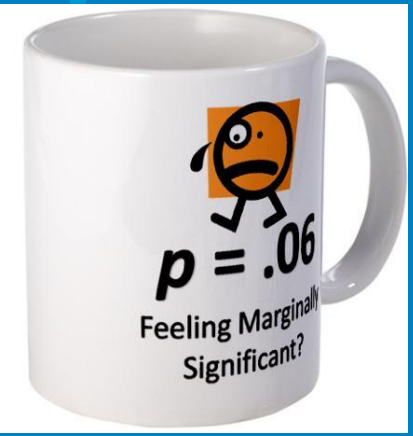
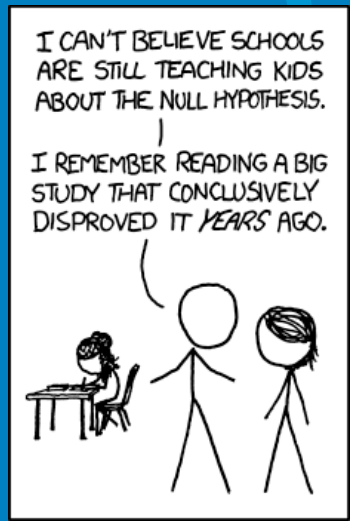
P value, Confidence Intervals, Statistical significance, Clinical significance



Vivek Pravin Dave
MD, DNB, FRCS
Consultant VitreoRetina
L V Prasad Eye Institute
India




Confidence intervals
P value
Significance
The Clinician!!!

Poll Question 1

What is your position?

1. Ophthalmologist
2. Ophthalmologist-in-training
3. Nurse
4. Ophthalmic technician/allied health
5. Medical student

"Tried to take a photo of a grasshopper on my windshield, but it looks like it's giant and destroying the town."

For today…

- P value
- Confidence intervals
- Clinical & statistical significance

Paracetamol or Morphine

- Better pain relief?
- How do you know?

Morphine relieves pain better than paracetamol at a probability that can be assumed to be more than just by chance

P value

- Measure of probability
- The probability of any observed difference having happened by chance
- P = 0.5 → 50% probability of the observed difference happening by chance
- P= 0.2 → 20% probability of the observed difference happening by chance
- P = 0.05 → 5% probability of the observed difference happening by chance
- P = 0.005 → 0.5% probability of the observed difference occurring by chance

So we interpret as…

- Suppose we find a difference for which the p value is say … 0.0032
- Means that there is a 0.32% probability that the observed difference has occurred by chance
- Hence arguably 99.68% chance that the study has measured a difference that actually exists

Confidence intervals

- Take an example
- Sample mean birth weight of a lot of 30 infants is say 2234.7 gram
- It means the population mean birth weight of the population of infants that this sample represents is also 2234.7 gram
- But +/- some small sampling error

All this leads to one question

- How big an error might we be making when we use the sample mean as the population mean
- The answer → confidence interval
- A numerical expression that quantifies the likely size of the sampling error
- CI depends on the standard error of the mean

Standard error of the mean

- Same example
- Sample mean 2234.7 gram
- Take multiple samples from the population
- Each has a mean
- By statistics if arranged in order → Bell curve
- The average of all possible sample means → coincides with the population mean
- SE in other words is similar to SD
- SE is for spread of mean of samples while SD is the measure for spread of data values

Calculating SE of the mean

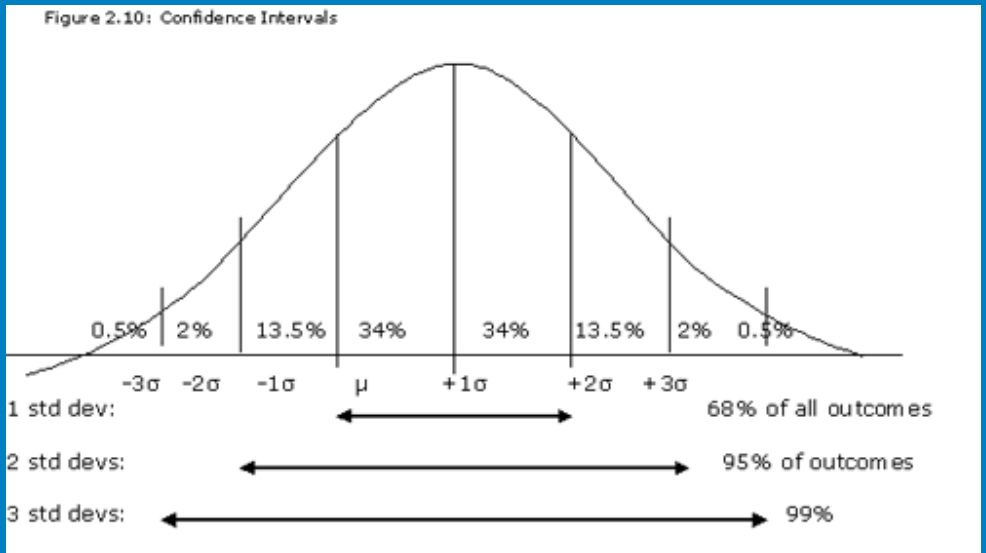
$$S.E. = \frac{S.D.}{\sqrt{n}}$$

- By definition distribution of sample means is "normal"

Confidence Interval

•So population mean confidence interval is calculated as:

Population mean = sample mean +/- [2 x S.E.]



To summarize

- Whatever difference one estimates
- Calculate the p value for the same
- If $p < 0.05 \rightarrow$ results can be considered as statistically significant
- Calculate the confidence interval around the value of the two estimates measured
- See whether they overlap or not
- If they overlap, the significance of the p value is dubious
- If they do not overlap then the significance of the p value is strengthened

To summarize

- Once statistical significance is proven look at the absolute value of the difference
- Ask yourself whether the observed difference is clinically significant practically
- If so then the difference is "statistically and clinically significant"

Poll question 2

If two studies have been conducted. In the first one the p value for the difference in outcome is p=0.056, in the second one it is p=0.047

1. The first has a highly significant difference
2. The second has a highly significant difference
3. The difference though present is not much
4. More information is required to interpret the final results

What is statistically significant

- Arbitrary
- One is not proving that difference exists
- One is estimating to what extent is the difference measured, actually not just by chance
- As usually < 5% by chance is acceptable to be a true difference, $p < 0.05$ is taken as acceptable significance

A non weighted coin

- Null hypothesis → coin is fair
- Can produce equal number of heads:tails
- Outcome variable is number of times you get a heads
- By null hypothesis → p=0.5 (50%)
- To prove it to be a weighted coin → reject the null
- Toss coin 100 times →

45 heads and 55 tails → p=0.31

??????????????????????????????

A non weighted coin

- As p > 0.05 we cannot reject the null
- So we accept that it is a non weighted coin

Heads %	Total sample	Number of heads	% Population C.I. for heads	% Population C.I. for tails	P value for difference
45	100	45	35.25 -54.75	45.25-64.75	0.32
45	200	90	38.11- 51.89	48.11 – 61.89	0.16
45	300	135	39.37 – 50.63	49.37 – 60.63	0.08
45	500	225	40.64 – 49.36	50.64 – 59.36	0.025
45	1000	450	41.92 – 48.08	51.92 – 58.08	0.002

Comparison of ramosetron with ondansetron for the prevention of post-operative nausea and vomiting in high-risk patients

Sandip Agarkar, Aparna S Chatterjee

Department of Anesthesia, Critical Care and Pain, Tata Memorial Hospital, Parel, Mumbai, Maharashtra, India

version 18: Chicago, IL, USA). **Results:** The incidence of PONV was found to be 35% in the ramosetron group as opposed to 43.7% in the ondansetron group ($P = 0.199$). Need for rescue antiemetic was 23.3% in the ramosetron group and 32% in the ondansetron group ($P = 0.156$) in the 24 h following surgery. **Conclusion:** Ramosetron 0.3 mg and ondansetron 8 mg were equally effective in reducing the incidence of PONV in high risk patients.

Table 2: Incidence of nausea, retching, emesis and rescue			
Incidence	*n*=103 each		*P*
	Ramosetron	Ondansetron	
Nausea			
0-6 h	36 (35)	40 (38.8)	0.564
6-24 h	1 (1.0)	4 (3.9)	0.174
Retching			
0-6 h	9 (8.7)	16 (15.5)	0.135
6-24 h	0 (0.0)	3 (2.9)	0.081
Emesis			
0-6 h	15 (14.6)	13 (12.6)	0.684
6-24 h	0 (0.0)	3 (2.9)	0.081
Rescue antiemetic			
0-6 h	24 (23.3)	30 (29.1)	0.342
6-24 h	0 (0.0)	3 (2.9)	0.246

Data represented as the number of patients (%)

Total	Ramo	Ondan	CI ramo	CI ondan	P value
103	35	38.8	25.79-44.21	29.39-48.21	0.67
206	35	38.8	28.49-41.51	32.15-45.45	0.48
309	35	38.8	29.68-40.32	33.37-44.23	0.37
1030	35	38.8	32.09-37.91	35.82-41.78	0.08
2060	35	38.8	32.94-37.06	36.7-40.9	0.01
5000	35	38.8	33.68-36.32	37.45-40.15	0.0001

Intrathecal dexmedetomidine as adjuvant for spinal anaesthesia for perianal ambulatory surgeries: A randomised double-blind controlled study

SS Nethra, M Sathesha[1], Dixit Aanchal[1], Pradeep A Dongare[1], SS Harsoor[1], D Devikarani[1]

Department of Anaesthesiology, Bangalore Medical College and Research Institute, Victoria Hospital,
[1]Department of Anaesthesiology, Bangalore Medical College and Research Institute, Bengaluru, Karnataka, India

Table 2: Sensory and motor parameters			
Parameter (min)	**Group D**	**Group N**	***P***
Duration of sensory block	430.05±89.13	301.10±94.86	<0.001**
Time for first administration of analgesic	459.80±100.9	321.85±95.08	<0.001**
Duration of motor block	323.05±54.58	220.10±63.61	<0.001**
Time to ambulation	329.55±54.06	221.60±63.84	<0.001**
Time to void	422.30±87.59	328.45±113.38	0.007**

**P<0.05 suggests statistically significant difference, Data presented as mean±SD. SD – Standard deviation

20 in each group

Table 2: Sensory and motor parameters			
Parameter (min)	**Group D**	**Group N**	***P***
Duration of sensory block	430.05±89.13	301.10±94.86	<0.001**

Confidence intervals	Group D	Group N
	390.99-469.11	259.53-342.67
Repeat with n = 10		
	358.73-501.37	225.1- 377.5 (p=0.16)

VISUAL AND ANATOMICAL OUTCOMES OF INTRAVITREAL AFLIBERCEPT IN EYES WITH PERSISTENT SUBFOVEAL FLUID DESPITE PREVIOUS TREATMENTS WITH RANIBIZUMAB IN PATIENTS WITH NEOVASCULAR AGE-RELATED MACULAR DEGENERATION

NISHANT KUMAR, FRCOPHTH,*† MARCELA MARSIGLIA, MD, PHD,*‡§ SARAH MREJEN, MD,* ADRIAN TIEN-CHIN FUNG, MBBS, MMED,* JASON SLAKTER, MD,*‡ JOHN SORENSON, MD,*‡ K. BAILEY FREUND, MD*‡§

RETINA 33:1605–1612, 2013

Table 2. Visual and Anatomical Results

	Baseline	After 3 IVA	*P* (Baseline Compared with After Third IVA)	At the Final Follow-up	*P* (Baseline Compared with the Final Follow-up
	Mean ± SD (IQR)	Mean ± SD (IQR)		Mean ± SD (IQR)	
Visual acuity (logMAR)	0.57 ± 0.36 (0.30–1.00)	0.52 ± 0.34 (0.30–0.70)	0.24	0.47 ± 0.32 (0.30–0.60)	0.004
mCFT	416 ± 217 (263–487)	351 ± 172 (224–445)	<0.001	348 ± 171 (235–419)	<0.001
Subfoveal distance between Bruch membrane and RPE	187 ± 158 (85–275)	161 ± 131 (78–225)	0.002	149 ± 125 (64–200)	0.002
Subfoveal distance between RPE and IS/OS junction line	32 ± 48 (0–58)	17 ± 28 (0–31)	0.02	14 ± 27 (0–0)	0.01
Subfoveal distance between IS/OS line and ILM	198 ± 150 (122–202)	173 ± 103 (107–178)	0.11	186 ± 101 (121–207)	0.44
Subfoveal PED height	260 ± 162 (129–368)	228 ± 140 (114–340)	0.001	215 ± 142 (111–305)	<0.001
Subfoveal PED diameter	3,265 ± 1,622 (2,354–4,555)	3,188 ± 1,599 (2,091–4,487)	0.51	2,949 ± 1,634 (1,721–4,484)	0.04

ILM, inner limiting membrane; IS/OS, inner and outer segment junction (also called the ellipsoid layer); IVA, intravitreal aflibercept (2 mg); logMAR, logarithm of minimum angle o
resolution

Author's interpretation

acuity. However, a significant improvement in visual acuity was observed at 6 months. In patients with persistent fluid despite previous treatments with other

Let's analyze

- Pre and post treatment values were Pre treatment → 0.57 logMAR , post treatment → 0.47 logMAR
- $p = 0.004$
- Pre CI 0.45-0.69, post CI 0.36-0.58 (Overlapping)

Let's analyze

Table 1. Available VA measurement systems

Snellen 6m	Snellen 20ft	Decimal	LogMar
6/6	20/20	1.00	0.0
6/9	20/32	0.63	0.2
6/12	20/40	0.50	0.3
6/18	20/60	0.33	0.5
6/24	20/80	0.25	0.6
6/36	20/120	0.17	0.8
6/60	20/200	0.10	1.0
3/60	20/400	0.05	1.3
1/60	20/1200	0.02	1.8
PL+	PL+	PL+	3
NPL	NPL	NPL	4

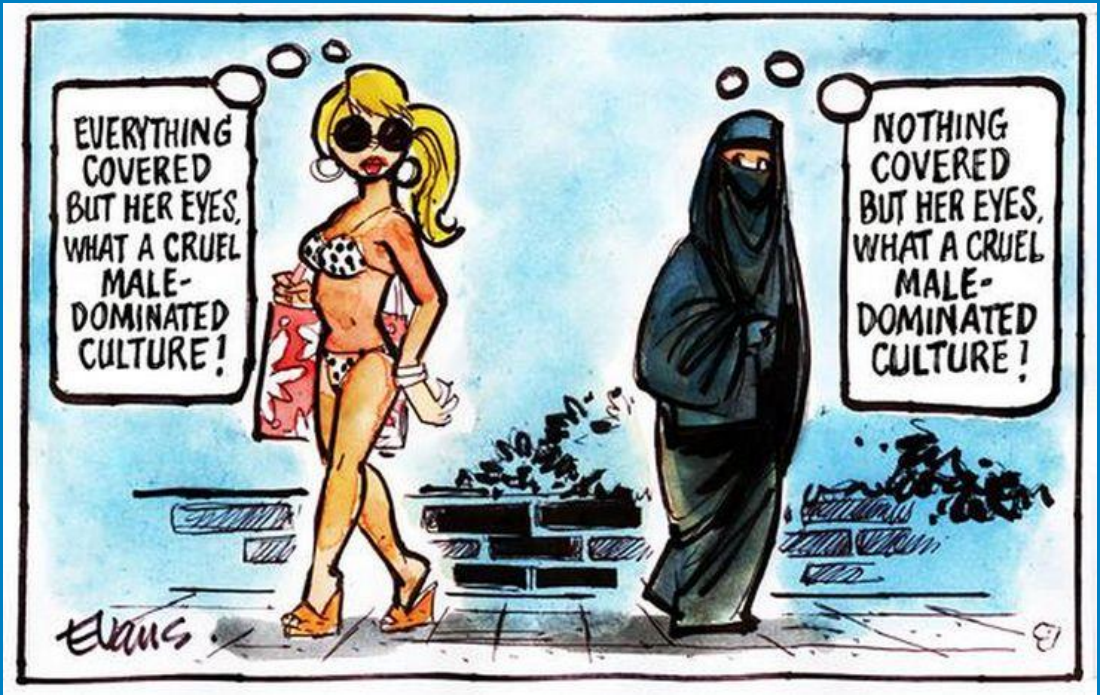
Source: The Epidemiology of Eye Diseases. Johnson GJ. London 2003, Arnold

Trial	Drug	Cost	No. of patients per group	Mean cholesterol (mg/decilitre) in drug group	Mean cholesterol (mg/decilitre) in control group	Reduction (mg/decilitre)
1	A	Cheap	30	140	180	40
2	A	Cheap	3000	140	180	40
3	B	Cheap	40	160	180	20
4	B	Cheap	4000	178	180	2
5	C	Expensive	5000	175	180	5

See the whole picture

- Absolute difference
- Total "n"
- p value
- Confidence intervals
- Derive statistical significance
- Set your own clinical significance

STATISTICS → you see what you want, interpret what you feel!!



Poll question 3

Which is the best drug

1. A
2. B
3. C
4. All are the same

Trial	Drug	Cost	No. of patients per group	Difference in mean cholesterol (mg/decilitre)	s.e. of difference	95% CI for difference	*P*-value
1	A	Cheap	30	−40	40	−118.4 to 38.4	0.32
2	A	Cheap	3000	−40	4	−47.8 to −32.2	< 0.001
3	B	Cheap	40	−20	33	−84.7 to 44.7	0.54
4	B	Cheap	4000	−2	3.3	−8.5 to 4.5	0.54
5	C	Expensive	5000	−5	2	−8.9 to −1.1	0.012

Thank you!

LVPEI
So that all may see

L V Prasad Eye Institute

Excellence • Equity • Efficiency