

Biostatistics in Dentistry

Data Entry (Excel & SPSS)

Session 1

Mahdi Akbarzadeh

akbarzadeh.ms@gmail.com

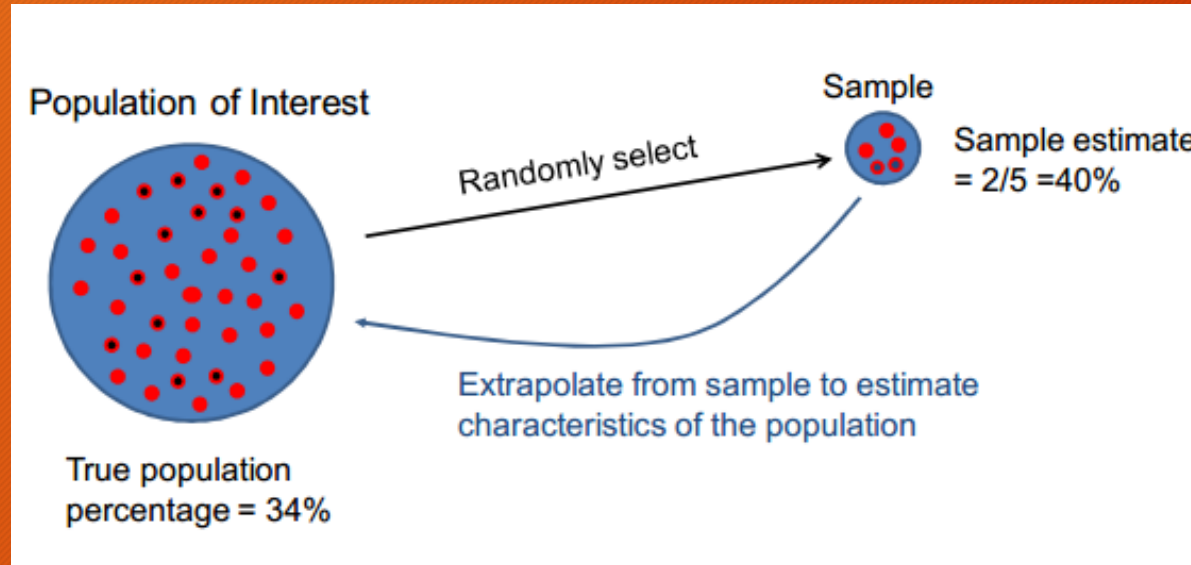
Introduction

2

Biostatistics is statistical methods applied primarily in medical/dental research

- Medical/dental research often involves human subjects, study of highly complex mechanisms
- This leads to highly variable outcomes.
- Statistical methods are used in order to separate the effects of interest from random “noise” in the data.

Statistics - used to describe characteristics of a large population using a small subset (sample)



- Two areas of statistics:

Generally the field of statistics can be divided in two parts:

➤ **Descriptive**

➤ **Inferential**

Example: Estimating ECC prevalence

4

age (months)	number with erupted teeth	number with caries	% with caries	95% confidence interval
6-8	15	0	0%	(0%, 21%)
9-10	48	10	21%	(11%, 34%)
11-14	167	86	51%	(44%, 59%)
15-19	157	130	83%	(76%, 88%)

*Data from Vachirarojpisan *et al* (2004), "Early childhood caries in children aged 6–19 months". Community Dentistry And Oral Epidemiology **32** (2), 133-142.

Example: Estimating ECC prevalence

age (months)	number with erupted teeth	number with caries	% with caries	95% confidence interval
11-14	167	86	51%	(44%, 59%)

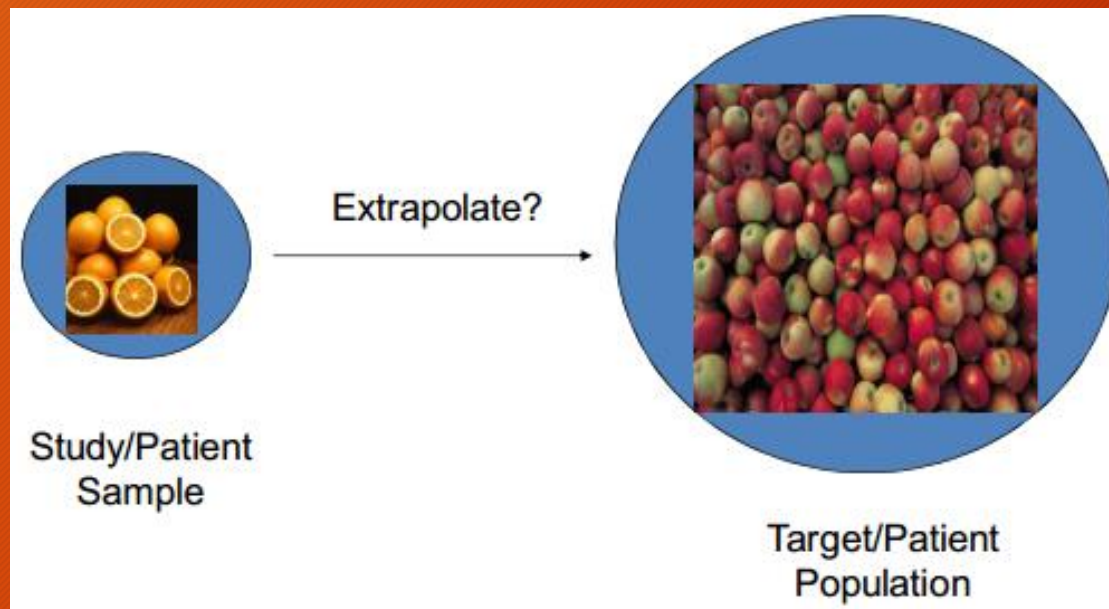
The size of the sample is 167.

- We estimate the prevalence in the **population** with the proportion in the sample:
 $86/167 = 51\%$
- With **careful sampling**, there's a good chance that the prevalence in the sample is close to the prevalence in the population.

Statistics only work if we have a representative, unbiased sample

6

Is the study sample representative of the target population?



Example: Estimating ECC prevalence

age (months)	number with erupted teeth	number with caries	% with caries	95% confidence interval
11-14	167	86	51%	(44%, 59%)

- Because of the random sampling, the “answer” we get is **random**.
- If we repeated this experiment we would likely get a different answer, because we would randomly sample a different group of children.

Statistics quantifies the random variability due to sampling in order to give us an idea of how much we can trust our results.

- While the true prevalence may not be exactly 51%, we can be 95% certain that the population prevalence is between **44% and 59%**.

Data Entry

8

Types of data

- Continuous
 - Differences between values have meaning, and are interpretable independent of the values themselves
 - E.g. difference between 8 and 9 basically the same as difference between 1 and 2.
- Ordinal
 - Values have an order, but differences are not easily interpretable (e.g. good, fair, poor)
- Categorical
 - Values are descriptive but do not have any obvious ordering. E.g. tx A, tx B, tx C.
- Binary, Dichotomous
 - Fancy names for categorical variables with only two possible values.

Data Entry

9

Some useful packages (ways) in data entry:

Excel

Text

SPSS

Stata

•

•

•

•

Data Entry: Chewing Gum Data

10

Please open the excel and enter the following dataset: (Track me)

order	variable	description
*****	*****	*****
1	ID	subject identifier
2	TX	treatment groups (551 = A, 552 = B, 652 = C)
3	SEX	gender (1=male, 2=female)
4	RACE	race (1=GARAFUNA, 2=CREOLE, 3=SPANISH, 4=INDIAN, 5=OTHER)
5	EXAMIN	examiner identifier (1-4)
6	AGE	age of subject at entry (years)
7	DMFS	CHANGE in DMFS scores (Decayed, Missing, or Filled surfaces)
8	NSAB	number of surfaces at baseline
9	DMFSAB	decayed,missing or filled surfaces at baseline

	A	B	C	D	E	F	G	H	I
1	ID	TX	SEX	RACE	EXAM	AGE	DMFS	NSAB	DMFSAB
2	331	652.00	1	2	4	10.24	6.00	71.00	2.0
3	335	652.00	1	3	4	9.68	1.00	52.00	8.0
4	336	652.00	1	3	1	9.93	1.00	58.00	0.0
5	340	652.00	1	2	1	10.01	0.00	108.00	0.0
6	342	652.00	1	2	2	9.58	0.00	44.00	0.0
7	344	652.00	2	1	1	10.08	0.00	89.00	1.0
8	345	652.00	2	2	3	9.70	5.00	52.00	0.0
9	346	652.00	2	2	2	9.90	-1.00	128.00	3.0
10	348	652.00	2	3	1	9.82	0.00	61.00	5.0
11	352	652.00	2	2	1	10.00	1.00	98.00	1.0

Data Entry: Chewing Gum Data

11

Next: an overview of Excel and some commands

- Merge cells
- Sequencing
- Transpose dataset
- Statistical commands
- Freeze panes
- Insert charts (single and multivariate)
- Save as text or ASCII file
- Import to SPSS and other statistical packages

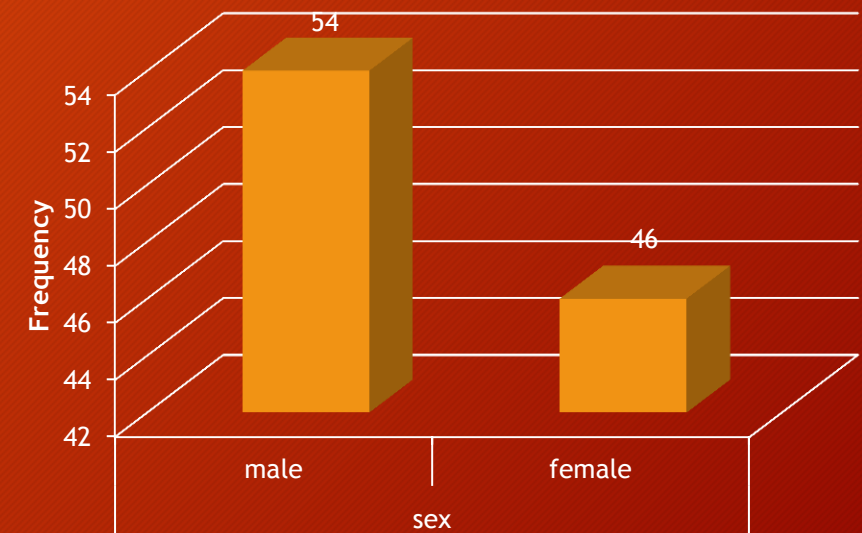


Data Entry: Group working

12

Please enter following data table on into excel and draw bar chart

gender					
		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	1	54	54.0	54.0	54.0
	2	46	46.0	46.0	100.0
	Total	100	100.0	100.0	



Data Entry: Group working

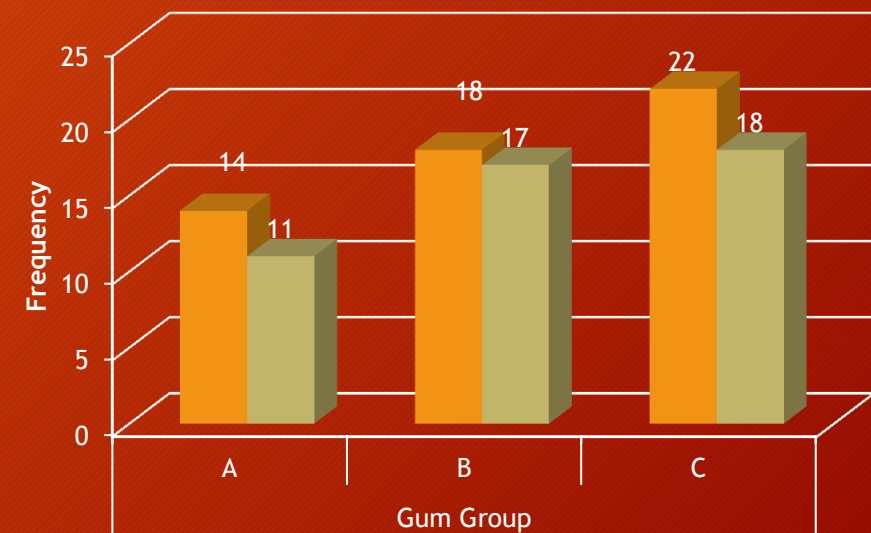
13

Please enter following data table on into excel and draw the cluster bar chart

gum type * gender Crosstabulation

Count

		gender		Total
		1	2	
gum type	A	14	11	25
	B	18	17	35
	C	22	18	40
Total		54	46	100

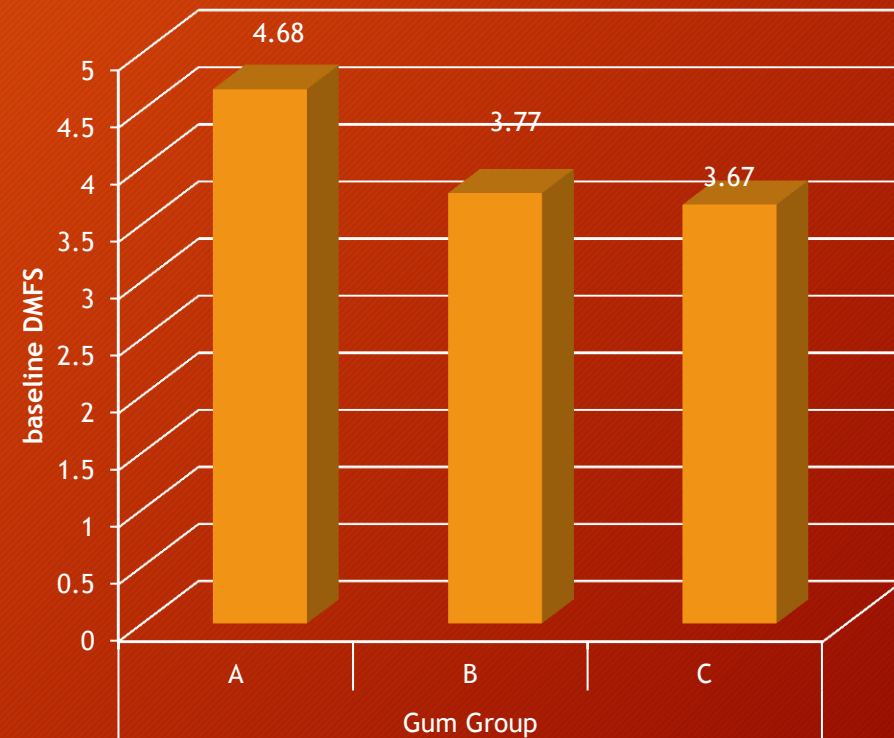


Data Entry: Group working

14

Please enter following data table on into excel and draw bar chart

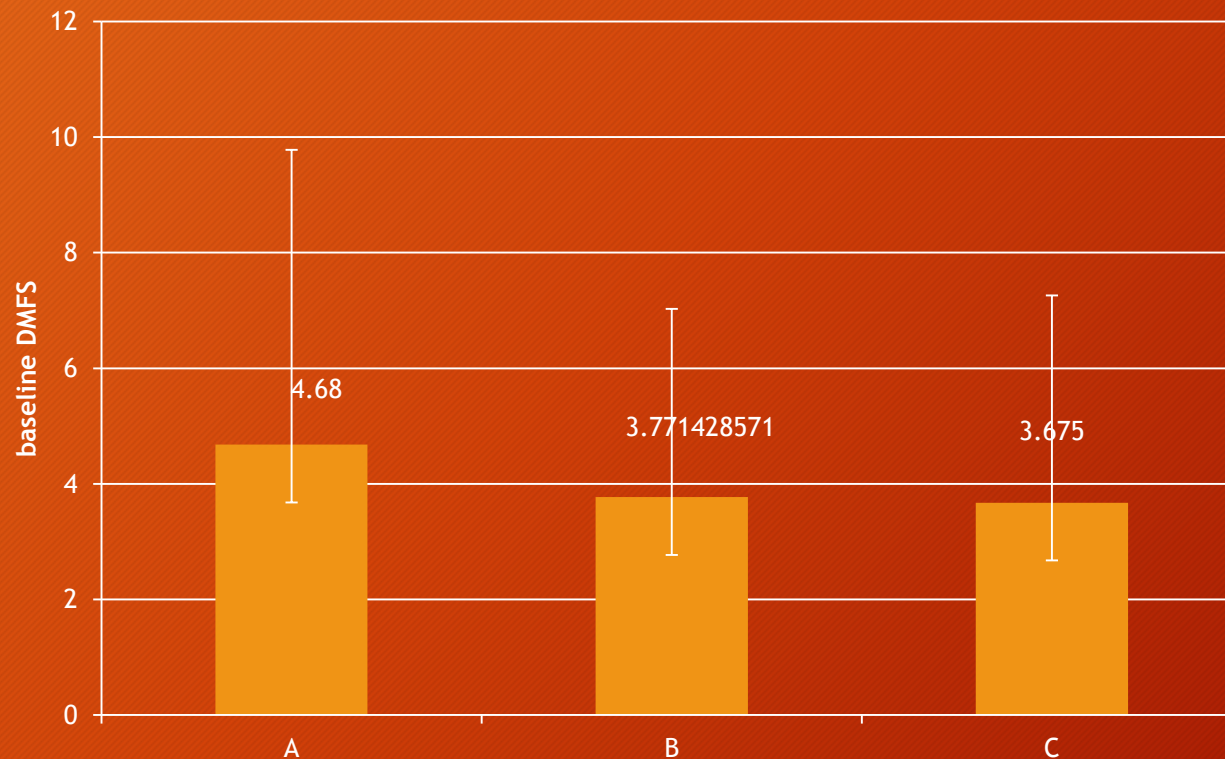
baseline DMFS			
A	N	Valid	25
		Missing	0
	Mean	4.680	
B	N	Valid	35
		Missing	0
	Mean	3.771	
C	N	Valid	40
		Missing	0
	Mean	3.675	



Data Entry: Group working

15

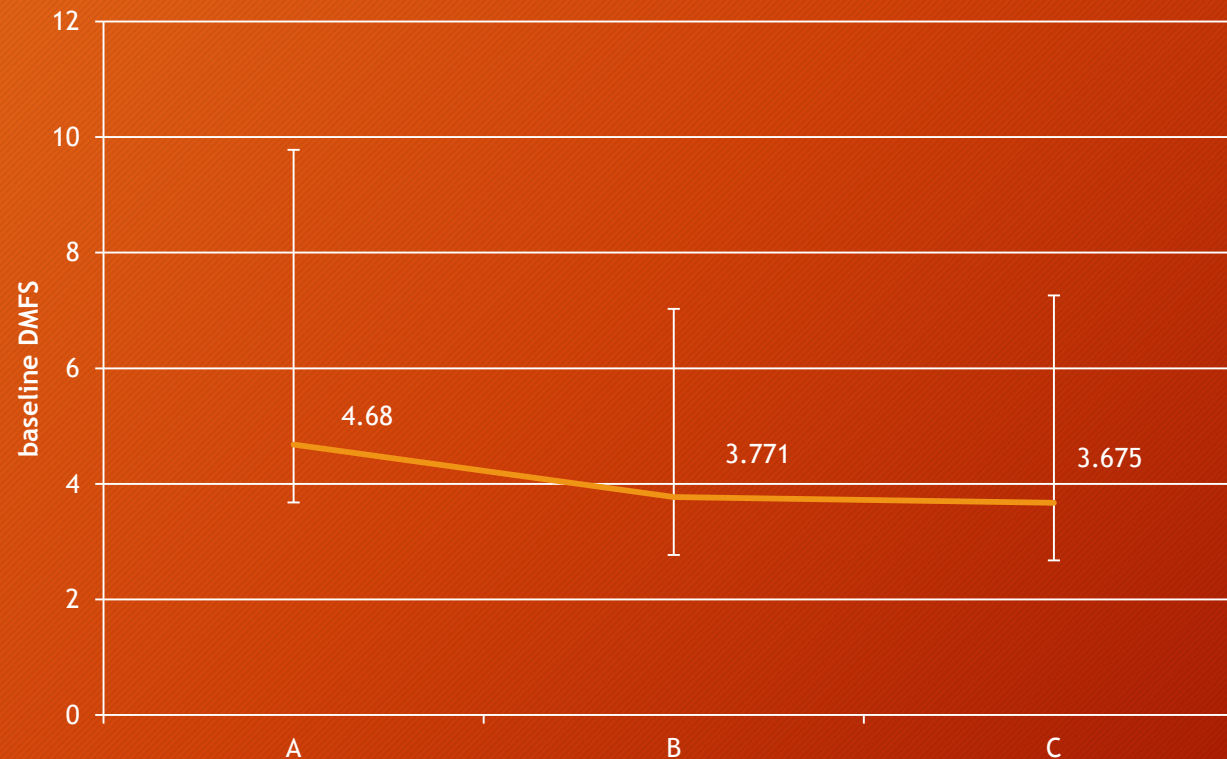
Please enter following data table on into excel and draw bar chart with error bar:



Data Entry: Group working

16

Please enter following data table on into excel and draw bar chart with error bar:



Example: Paper

17

Orthodontic forces increase tumor necrosis factor α in the human gingival sulcus

Jennifer J. Lowney, DMD,^a Louis A. Norton, DMD,^b David M. Shafer, DMD,^c and Edward F. Rossomando, DDS, PhD^d

Farmington, Conn.

The production of cytokines has been associated with the biology of tooth movement in animal populations. The purpose of this study was to measure tumor necrosis factor- α (TNF) directly in the human gingival sulcus before and after the application of an orthodontic force. To recover TNF from the sulcus, paramagnetic beads, coated with monoclonal antibodies for TNF, were introduced into the gingival sulcus of 50 teeth undergoing orthodontic tooth movement (by two force systems) in 20 patients. Retrieval was performed by a permanent magnetic device designed to fit the periodontal sulcus. The samples were taken before force application (controls); and at a fixed time after force application. The amount of immunoabsorbed TNF was quantified with an immunochemical assay. There was a greater than twofold increase in TNF recoverable from the gingival sulcus after application of orthodontic forces (mean of 12.9 ng vs 30.5 ng). A Student's *t* test for paired samples demonstrated statistical significance at $p < 0.01$. We conclude that the quantity of paradental TNF, found in human gingival sulcus, is elevated during tooth movement. The source may be from the adjacent gingiva, but more likely the compressed periodontal ligament and the resorbing bone adjacent to the root surface. (AM J ORTHOD DENTOFAC ORTHOP 1995;108:519-24.)

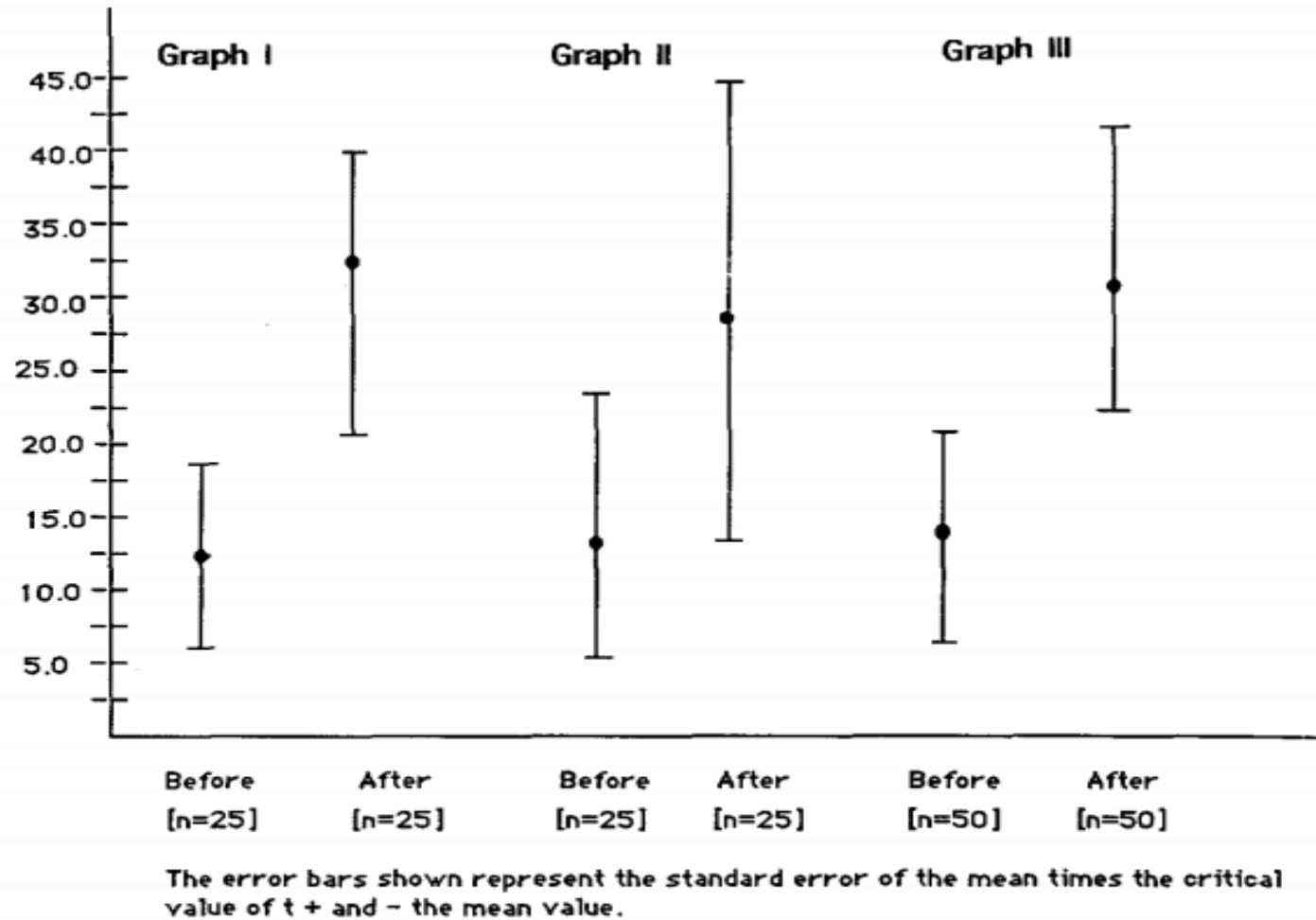


Fig. 4. Graph I, II, and III. Graph I shows mean concentrations of TNF before and after application of force for 25 sample subset in which T-springs were used and graph II shows mean concentration of TNF for 25 sample subset with chain elastic. Graph III show mean concentrations for all 50 samples. Error bars shown represent standard error of mean \times critical value of $t +$ and $-$ mean value.