منهجية البحث العلمي والوبائيات والاحصاء الطبي

السنة الثالثة كلية الطب البشري

البحث العلمي

أهمية البحوث:

- أسباب المشاكل الصحية وعلاجها.... محددات الصحة
- انتاج المعارف الجديدة وبالتالي تحسين الممارسات الصحية والطبية
- بناء سياسات صحية
- بحوث عملياتية، بحوث النظم الصحية، بحوث الخدمات الصحية....

البحث العلمي

تعريف:

- البحث العلمي بالتعريف هو عمل منهجي منظم يتضمن حملة من الاجراءات والتقنيات الجوهرية بقصد الحصول على معرفة جديدة من خلال الاستقصاء او التحقيق او التجريب للمساهمة في حل مشكلة معينة أو الاجابة على تساؤل ما
- أساليب البحث العلمي هي اساس تحسين فهم محددات الصحة وسبل تعزيز الصحة والوقاية من المرض

أنماط البحوث

تقسم البحوث بأشكال عدة فهناك التنميط على أساس النهج الفلسفي للبحث وهناك التنميط على أساس وظيفة البحث وهناك التنميط على أساس أهداف البحث وغير ذلك

البحوث التجريبية والبحوث النظرية
البحوث الأساسية والبحوث التطبيقية

النمط حسب مجالات البحوث

دوائية-سلوكية- خدمات صحية....

أساسيات البحوث الصحية

1- الترتيب

البحوث تعتمد على الملاحظة والقياس، وتتصف الملاحظة في البحوث بأنها منهجية ومرتبة، وهذا أساس في الاستقصاء العلمي.

2- الاستدلال

الاستدلال inference هو الأساس المنطقي للبحث العلمي، وقد يكون استقرائياً أو استنباطياً inductive or deductive.

أساسيات البحوث الصحية

3-الاحتمال

وهو المكونة الحاسمة والهامة في أساسيات البحوث ففيها يتم الاعتماد على تقييم الاحتمالات عند الاستدلال من الخاص إلى العام حيث أن دراسة على عينة من الجمهرة ستعتمد في إمكانية تعميم نتائجها على الجمهرة كاملة على الاحتمال. والاحتمال يحسب بالطرائق الإحصائية المعروفة في البحث العلمي وستكون مجالاً للمزيد من التفصيل لاحقاً في هذا الكتاب.

4- الفرضية

الفرضية هي عبارة مبناة بشكل دقيق حول الظاهرة قيد الدراسة. وهي تأتي من طرائق الاستدلال المختلفة لتصاغ بشكل يمكن التحقق منها. ورغم أنه لا يمكن أحياناً التأكد التام من الفرضية إلا أنه من الممكن رفضها وصياغة فرضية أكثر معقولية منها.

أساسيات البحوث الصحية

الاستنتاج

البحث العلمي هو من أهم التحديات التي تواجه البشر فهو أساس لاستقدام المعارف والحقائق والتي يجب أن تعتمد في استقدامها على طرائق ذات صدقية وعلى منهجيات بحثية سليمة وأخلاقية.

مخطط بحث علمي

د. عبير قدسي

قسم طب الأسرة والمجتمع كلية الطب-جامعة دمشق

- تختلف كتابة مشروع بحث علمي من بحث إلى آخر ولكن هناك **خطوط عريضة أساسية** لأي مشروع/مقترح

- يجب على خطة البحث أن تشمل **جميع العناصر الأساسية** المطلوبة **لإجراء بحث ناجح**، وجميع المعلومات الضرورية التي تسمح **بتقييم البحث**.

Element	Purpose
Research question	What questions will the study address?
Significance (background)	Why are these questions important?
Design	How you organize and carry out the study?
Hypothesis	
Time frame	
Methods	
Subjects	Who are the subjects and how will you selected them?
Variables	What will you measure and what will your data look like?
Predictor variables	
Confounding variables	
Outcome variables	
Statistical analysis	What will you do with the data?
Sample size calculation	
Data analysis method	

يشمل مخطط البحث العلمي

- العنوان
- الملخص والكلمات المفتاحية
- أهمية البحث
- مقدمة تشمل مراجعة الأدب الطبي
- سؤال البحث

يشمل مخطط البحث العلمي

- الهدف من البحث
- المواد والطرائق
  - تصميم الدراسة
  - الزمان-المكان
  - جمهرة الدراسة- العينة –الاعتيان- حجم العينة
  - التحليل الإحصائي

يشمل مخطط البحث العلمي

- خطة العمل
- الجوانب الأخلاقية
- مدة التنفيذ (الجدول الزمني للتنفيذ)
- الميزانية والتمويل
- المراجع

اختيار الموضوع

- البحث العلمي وسيلة للحصول على **معرفة جديدة**

اختيار الموضوع

- **فكرة مبتكرة؟؟**
- توفر العوامل المختلفة الضرورية لتنفيذ هذه الفكرة ..... الأفراد – الزمن – الأجهزة والمواد - التمويل

اختيار الموضوع

- لاتتوفر دائما أفكاراً مبتكرة !!
- الاستفادة من خبرات باحثين آخرين لتطوير بعض جوانب أفكار موجودة مسبقا
- تصاميم وبينات اقوى... إعادة بعض الدراسات مع تحسين شروط الدراسة كزيادة عينة الدراسة او إطالة فترة المتابعة
- تكرار دراسة سابقة ولكن على عينة مختلفة.. من المرضى.....

اختيار الموضوع

- من الهام عند اختيار الموضوع التوجه للمواضيع التي تهم البلد/الوضع/المجتمع/التنمية
- يجب أخذ خصوصيات مجتمعنا وعاداتنا والاهتمام بها عند اختيار المواضيع وطرائق البحث.
- مشاريع بحثية واسعة مقررة مسبقا وممولة ومنجز بعض حلقاتها

المعلومات العامة - الباحث

- الصفحة الأولى بمثابة غلاف يشمل المعلومات التالية: عنوان البحث، اسم الباحث، المؤسسة التي يتم فيها البحث، عنوان الباحث الالكتروني ، ورقم الهاتف، ويكتب الغرض من البحث (ماجستير، بحث، دكتوراه.....).

العنوان

- الوضوح .
- الشمول .
- الإيجاز .
- لا يحوي نتائج أو أحكام .
- التجديد .

الملخص

- مختصر واضح وغني بالمعلومات لمشكلة البحث وطريقة حلها.
- فكرة واضحة للقارئ عن الهدف من البحث وطريقة إنجازه
- عدد محدود من الكلمات
- يكتب الملخص عادة بعد إنجاز كتابة مشروع البحث.

الأساس المنطقي - التبرير

- مقدمة الخطة يبيّن فيها الباحث أهمية بحثه بالنسبة للبحوث والكتابات السابقة في ذات المجال، كما يوضح الدافع وراء اختياره لموضوع البحث

- السياق
- الدافع
- الأهمية

المراجعة -الأدب

- الأبحاث المجراة
- عددها
- نوعها
- البينة الحالية
- تقييم نقدي....

سؤال البحث

- قابل للتطبيق
- مثير للاهتمام
- جديد
- أخلاقي

الغرض-الهدف

- أكاديمي واستراتيجي.....
- عام وخاص
- متسلسل ...

المواد والطرائق

تصميم الدراسة

1. الرقابية
2. التجريبية
3. المراجعة المنهجية؟

المواد والطرائق

تصميم الدراسة

- البيئة الحالية
- طبيعة الأهداف الموضوعة في البحث
- الإمكانات المتاحة للحصول على المعلومات
- الجانب الأخلاقي الذي قد لايسمح بإجراء تجربة

المواد والطرائق

المكان

المواد والطرائق

الزمان

المواد والطرائق

أفراد الدراسة

- الجمهرة
- العينة
- حجم العينة
- الاعتيان – معايير الادخال والاستبعاد...

الأخلاقيات

- سلامة الأفراد – احترام الحقوق
- الموافقة المستنيرة
- معلومات
- السرية واحترام الخصوصية
- التدريب الكافي للقيام بالإجراءات المعنية

المواد والطرائق

العمل الميداني – الجهات المشاركة – التجهيزات - المخابر

المواد والطرائق

المواد-الأداة-الاستبيان

- بالتفصيل
- التبرير
- مصدوقية؟
- قابلية التطبيق

المواد والطرائق

التحليل الإحصائي

المواد والطرائق

خطة العمل

- مخطط واضح
- خطوات العمل

المواد والطرائق

البرنامج الزمني

- بالتفصيل
- الالتزام بالمدة المسموح بها

المواد والطرائق

الميزانية- التمويل

- المواد بالتفصيل
- سعر الوحدة
- الإجمالي....
- أجور النقل....
- أجور المساعدين الفنيين؟؟

المراجع

- التوثيق
- السرقة؟؟؟؟

- Breen KJ . Misconduct in medical research: whose responsibility? Intern Med J 2003;33:186-91.

الفوائد من
كتابة مشروع بحث علمي

- صياغة هدف البحث بدقة ووضوح.
- وضع خطة غنية بالتفاصيل حول ما يتوجب عمله، وأخذ ملاحظات واقتراحات الآخرين على هذه الخطة.
- رؤية العملية البحثية بشكل متكامل.

الفوائد من
كتابة مشروع بحث علمي

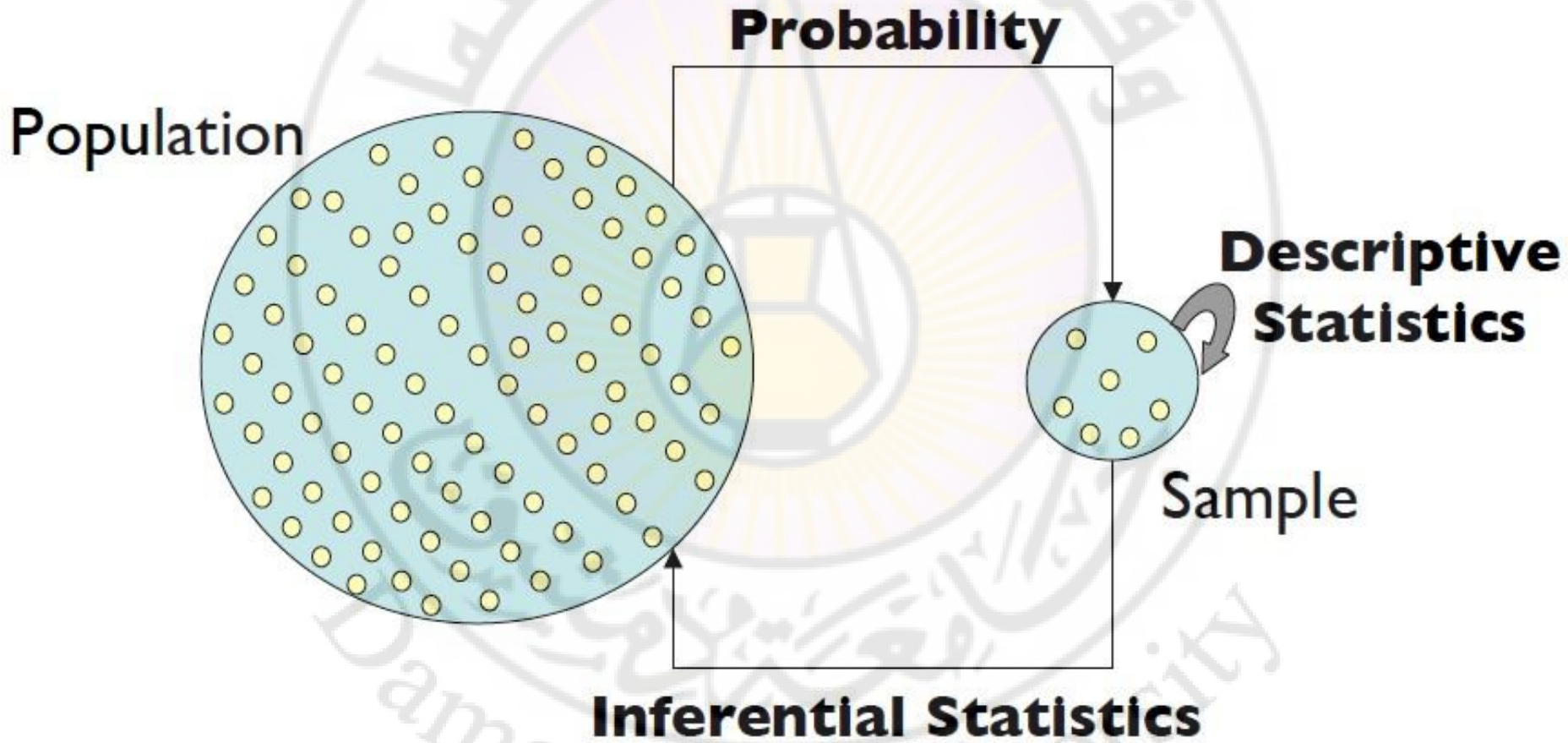
- دليل لمجموع العاملين في البحث وينسّق نشاطهم.
- وسيلة للتذكير بخطوات يمكن أن ننساها.
- مراقبة مقدار التقدم في إنجاز البحث .
- الحصول على التمويل–الدعم المادي لتنفيذ البحث.

المراجع

- Research proposal guide, Developing and submitting a research proposal. University of the Western Cape: http://www.uwc.ac.za/usrfiles/users/270084/RESEARCH_PROPOSAL_2.pdf.[Accessed 24 January 2011].

MEDICAL STATISTICS

Basis of Statistics



http://www.gs.washington.edu/

DATA AND VARIABLES

DATA: the answers to questions or measurements from the experiment

VARIABLE = measurement which varies between subjects e.g. height or gender

One variable per column

One row per subject

	A	B	C	D
1	Subject ID	Gender	Year of study	Height
2	1	Male	1	170
3	2	Female	2	160
4	3	Female	3	165
5	4	Male	PG	175
6	5	Female	3	168

Variables

Numerical

Continuous
Measurements
takes any value

Discrete:
Counts/ integers

Categorical

Ordinal:
obvious order

Nominal:
no meaningful order

Nominal Variable: A qualitative variable that categorizes (or describes, or names) an element of a population.

Ordinal Variable: A qualitative variable that incorporates an ordered position, or ranking.

Discrete Variable: A quantitative variable that can assume a countable number of values. Intuitively, a discrete variable can assume values corresponding to isolated points along a line interval. That is, there is a gap between any two values.

Continuous Variable: A quantitative variable that can assume an uncountable number of values. Intuitively, a continuous variable can assume any value along a line interval, including every possible value between any two values.

QUESTIONNAIRE FOR ……

What data types relate to following questions?

- Q1: What is your favourite subject?

Maths	English	Science	Art	French

- Q2: Gender:

Male	Female

- Q3: I consider myself to be good at mathematics:

Strongly Disagree	Disagree	Not Sure	Agree	Strongly Agree

- Q4: Score in a recent mock maths exam:

Score between 0% and 100%

QUESTIONNAIRE FOR ……

What data types relate to following questions?

- Q1: What is your favourite subject?

Nominal

Maths	English	Science	Art	French

- Q2: Gender:

Male	Female

Binary/ Nominal

- Q3: I consider myself to be good at mathematics:

Ordinal

Strongly Disagree	Disagree	Not Sure	Agree	Strongly Agree

- Q4: Score in a recent mock maths exam:

Numerical

Score between 0% and 100%

Example: Identify each of the following as examples of (1) nominal, (2) ordinal, (3) discrete, or (4) continuous variables:

1. The length of time until a pain reliever begins to work.
2. The number of chocolate chips in a cookie.
3. The number of colors used in a statistics textbook.
4. The brand of refrigerator in a home.
5. The overall satisfaction rating of a new car.
6. The number of files on a computer's hard disk.
7. The pH level of the water in a swimming pool.
8. The number of staples in a stapler.

Variable: A characteristic about each individual element of a population or sample.

Data (singular): The value of the variable associated with one element of a population or sample. This value may be a number, a word, or a symbol.

Data (plural): The set of values collected for the variable from each of the elements belonging to the sample.

Experiment: A planned activity whose results yield a set of data.

Parameter: A numerical value summarizing all the data of an entire population.

Statistic: A numerical value summarizing the sample data.

WHAT IS STATISTICS?

Statistics: The science of collecting, describing, interpreting data.

Two areas of statistics:

Descriptive Statistics: collection, presentation, description of sample data.

Analytic(Inferential) Statistics: making decisions and drawing conclusions about populations.

Descriptive Statistics

- collection, presentation, and description of sample data.

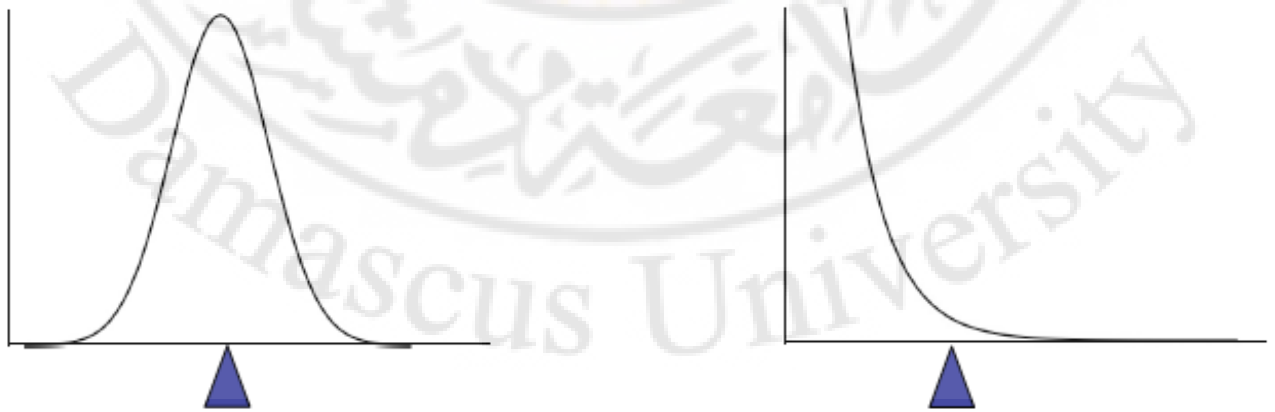
Descriptive Statistics: Quantitative Variables

- Central Tendency Determinants
  - Mean
  - Median
  - Mode (most prevalent)
- Dispersion Determinants
  - Range
  - Variance
  - SD
  - SE

Location: Mean

- The average of a set of observations
- Add values and divide by the number of observations

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \ldots + x_n}{n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$



How can exam score data be summarised?

Exam marks for 60 students (marked out of 65)

48	37	1	33	26	22	15	22	40	30	12	36
21	20	29	13	44	52	28	39	16	48	56	27
47	12	35	24	10	36	18	34	9	25	31	42
31	27	64	25	58	17	26	38	28	43	33	5
25	55	7	32	39	23	49	43	11	37	22	54

Mean = 30.3 SD = 14.46

Summary statistics

- Mean = $\frac{\sum_{i=1}^{n} x}{n} = \bar{x}$

Standard deviation (s) is a measure of how much the individuals differ from the mean

$$s = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1}}$$

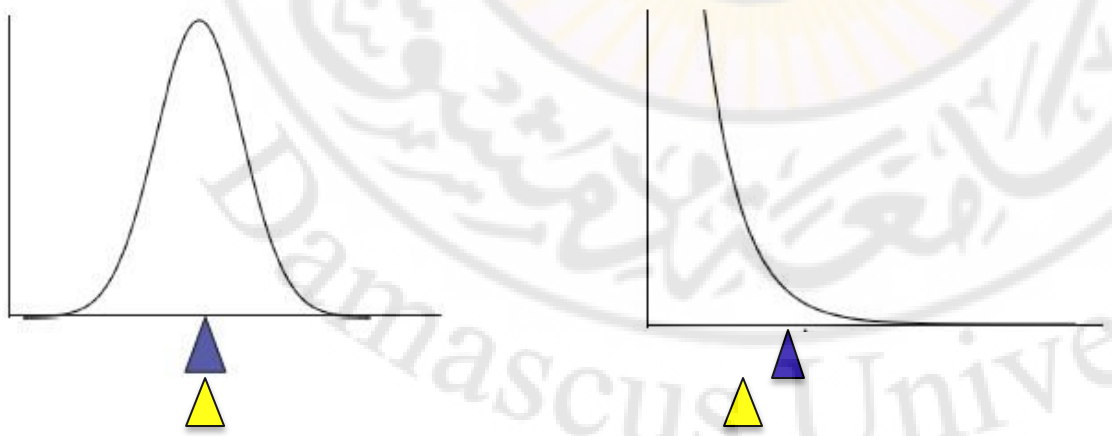
Large SD = very spread out data

Small SD = there is little variation from the mean

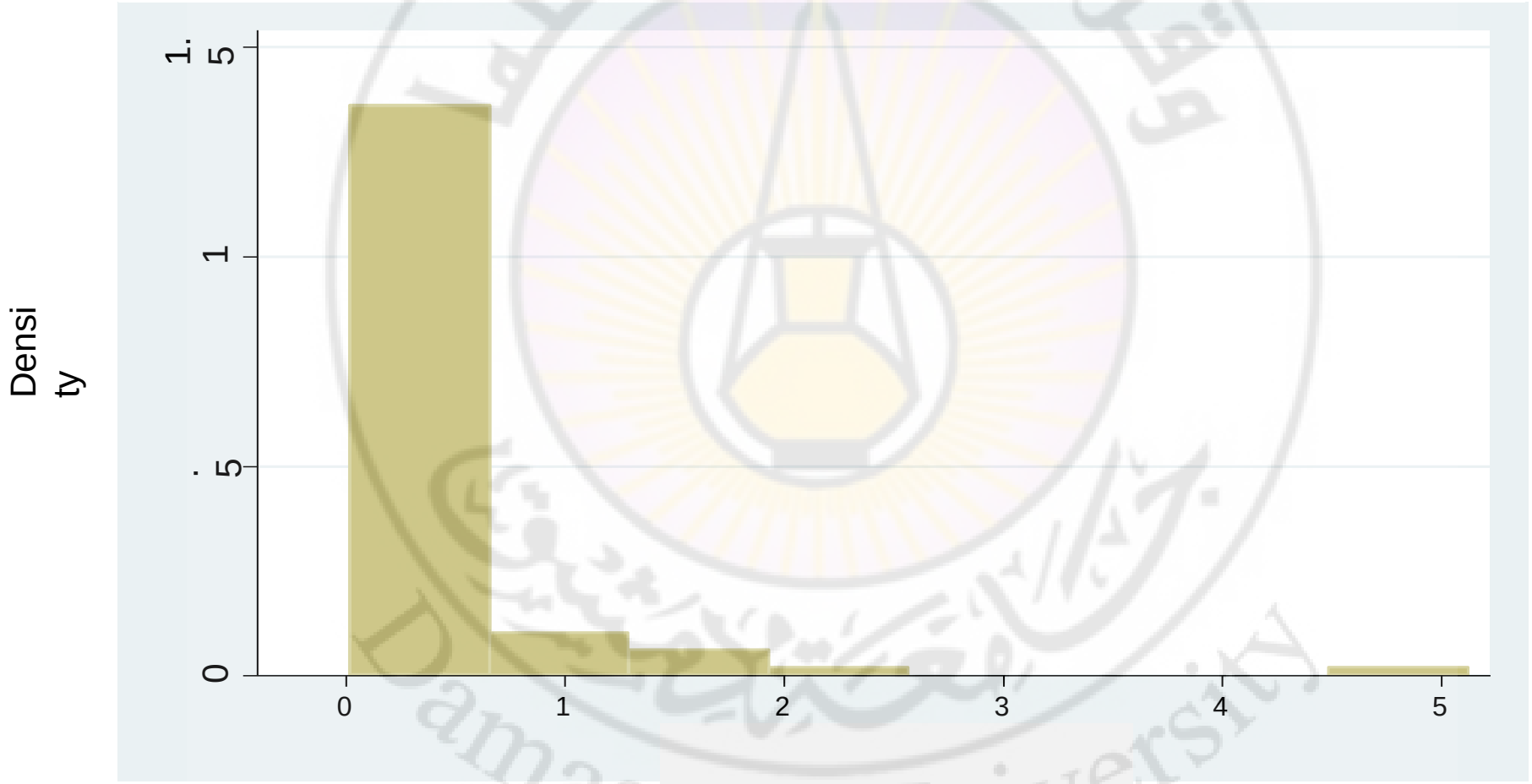
For exam scores, mean = 30.5, SD = 14.46

Which Measure is Best?

- Mean
  - best for symmetric (or normal) distributions
- Median
  - Useful for skewed distributions or data with outliers



Biomarker – one time point

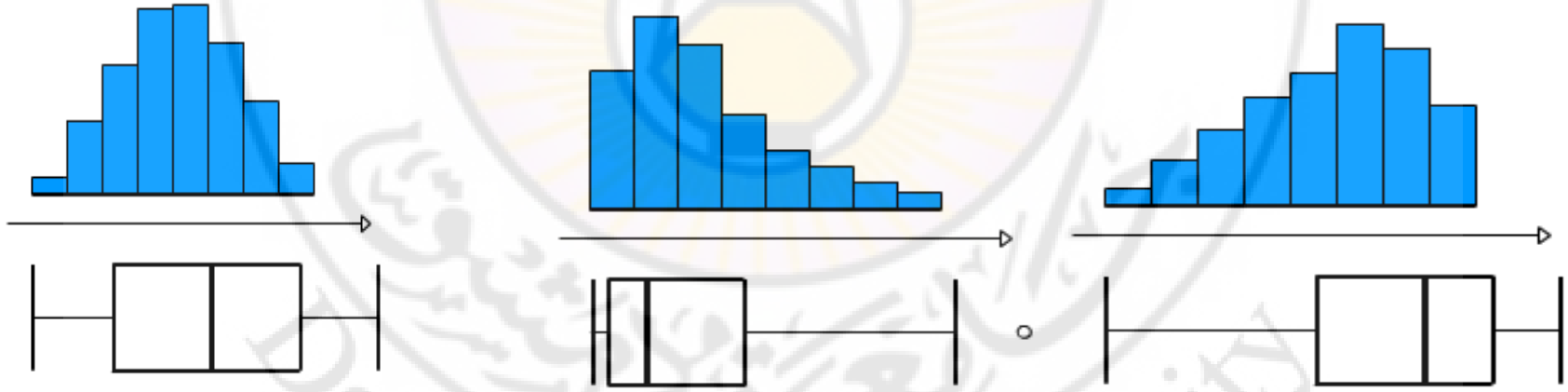


Assessing Normality

Charts can be used to **informally** assess whether data is:

Normally distributed

Or….Skewed



The mean and median are very different for skewed data.

Scale: Variance

- Average of the squared deviations of values from the mean
- Example, sample variance

$$\hat{\sigma}^2 = \frac{\sum_{i}^{n} (x_i - \bar{x})^2}{n-1}$$

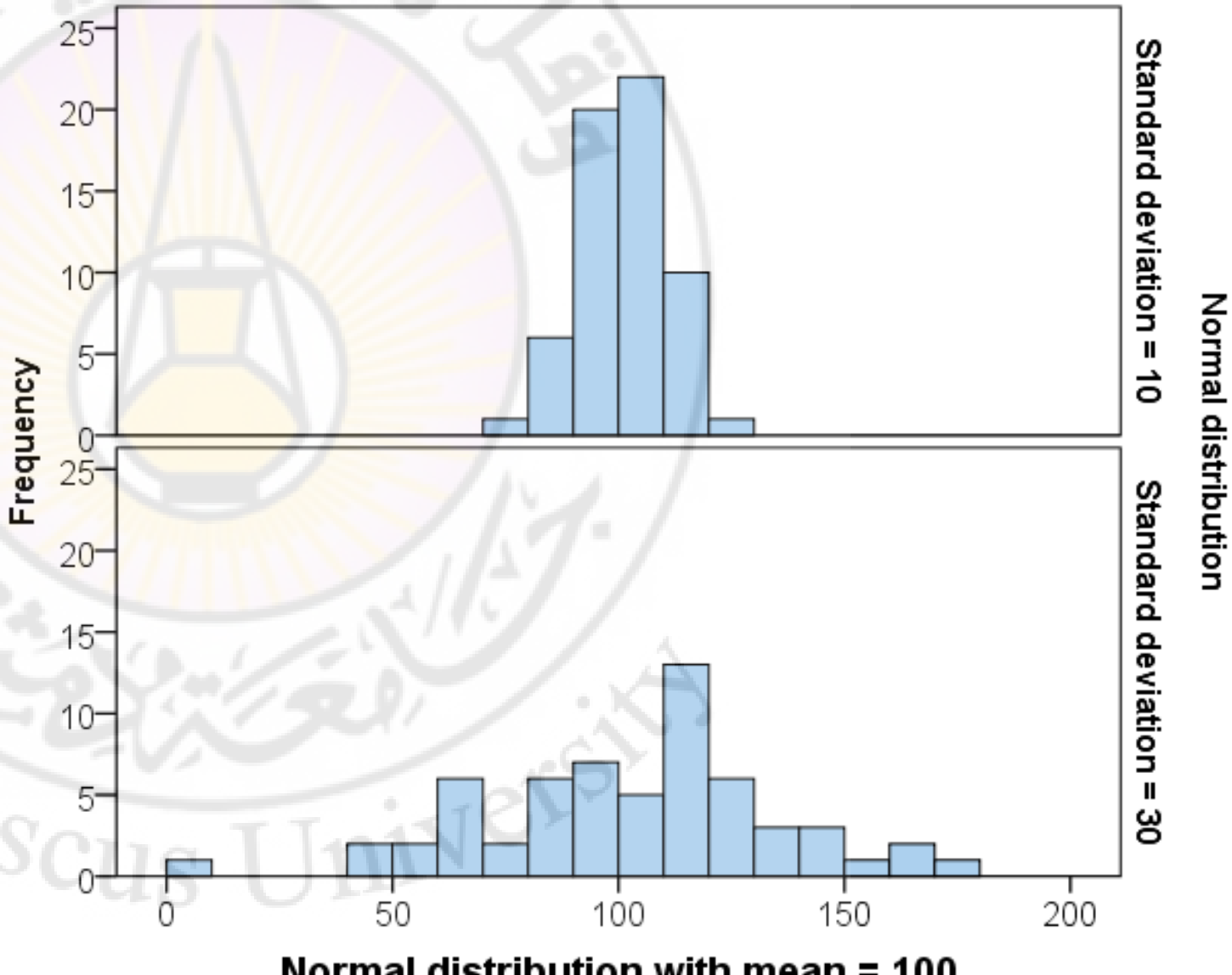
Scale: Standard Deviation

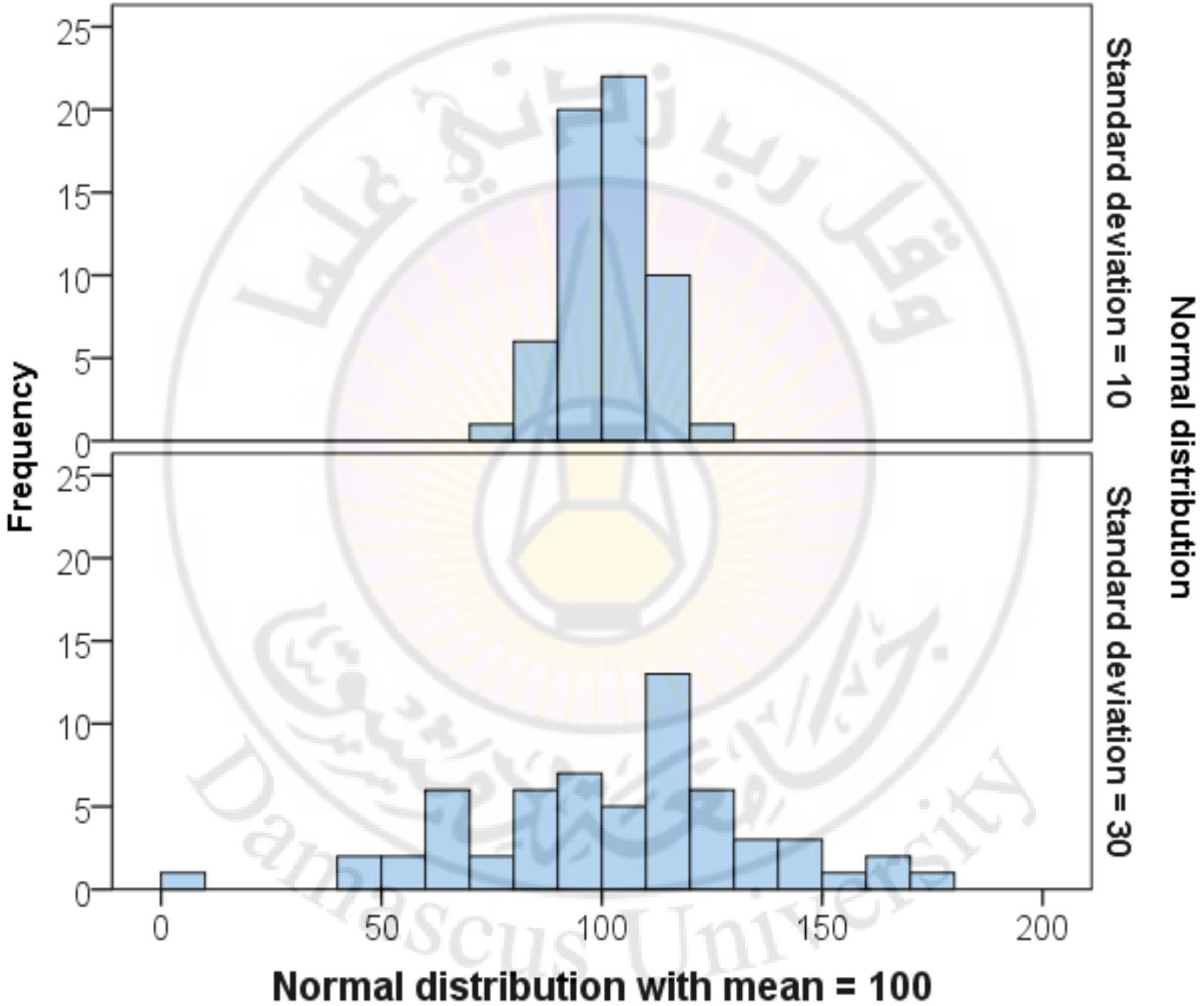
- Variance is somewhat arbitrary
- Standardizing helps to bring meaning to deviation from the mean
- Standard deviations are simply the square root of the variance
- Example, sample SD

$$\hat{\sigma} = \sqrt{\frac{\sum_{i}^{n}(x_i - \bar{x})^2}{n-1}}$$

Interpretation of standard deviation

- **The larger the standard deviation, the more spread out the data is.**




Frequency
25
20
15
10
5
0
Standard deviation = 10
Standard deviation = 30
Normal distribution
0
50
100
150
200
Normal distribution with mean = 100

Scale: Quartiles and Inter Quartile Range (IQR)

- Quartiles or percentiles (order data first)
  - Q_1 (1st quartile) or 25th percentile is the value for which 25% of the observations are smaller and 75% are greater
  - Q_2 is the median or the value where 50% of the observations are smaller and 50% are greater
  - Q_3 is the value where 75% of the observations are smaller and 25% are greater

Graphical Summaries of Data: Box Plots and Histograms

- Box plot (i.e. box-and-whisker plots)
  - Shows frequency or proportion of data in categories, i.e categorical data
  - Visual of frequency tables
- Histogram
  - Shows the distribution (shape, center, range, variation) of continuous variables
  - Bin size is important

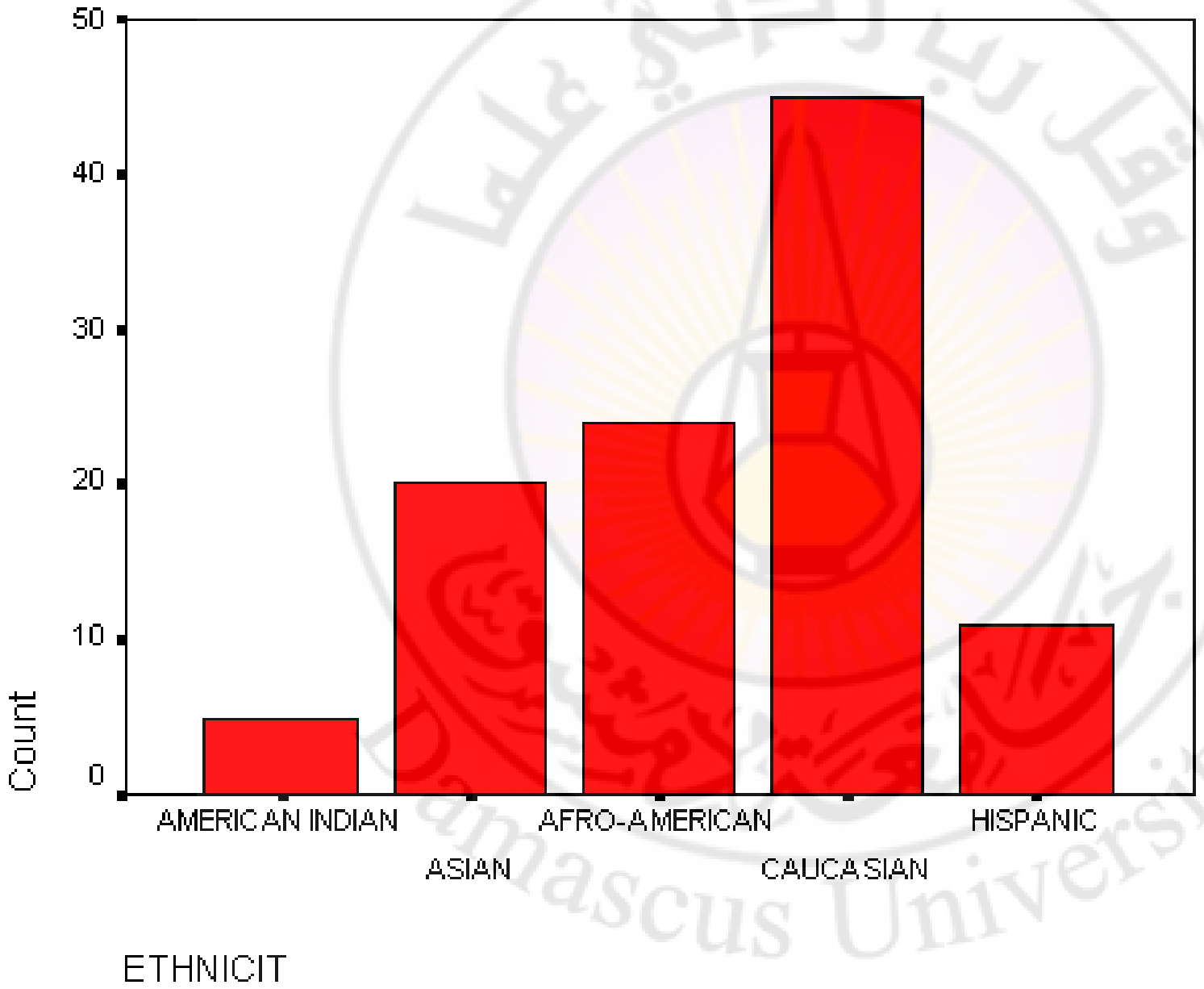
Introduction to Plots

- A **plot**(graphs) is a graphical technique for representing a data set.
- Graphs are a visual representation of the variables and relationship between variables.
- Plots are very useful for humans who can quickly derive an understanding which would not come from lists of values.

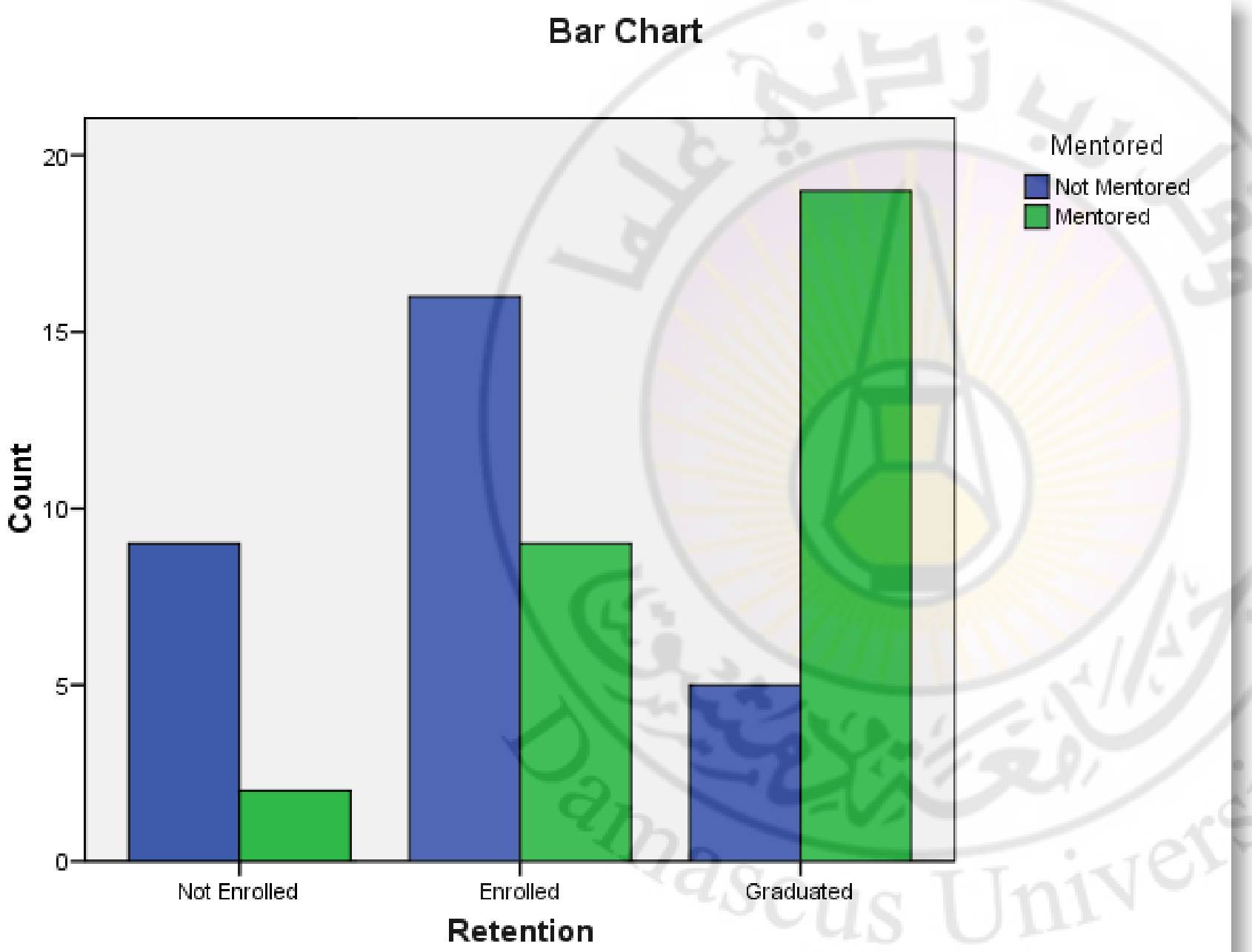
Dos and Do Nots of Graphing

- Goal of graphing
  - To portray data accurately and clearly
- Rules of graphing
  - Label and appropriately scale axis
  - Simplify, display only the necessary information
  - Stay away from pie charts

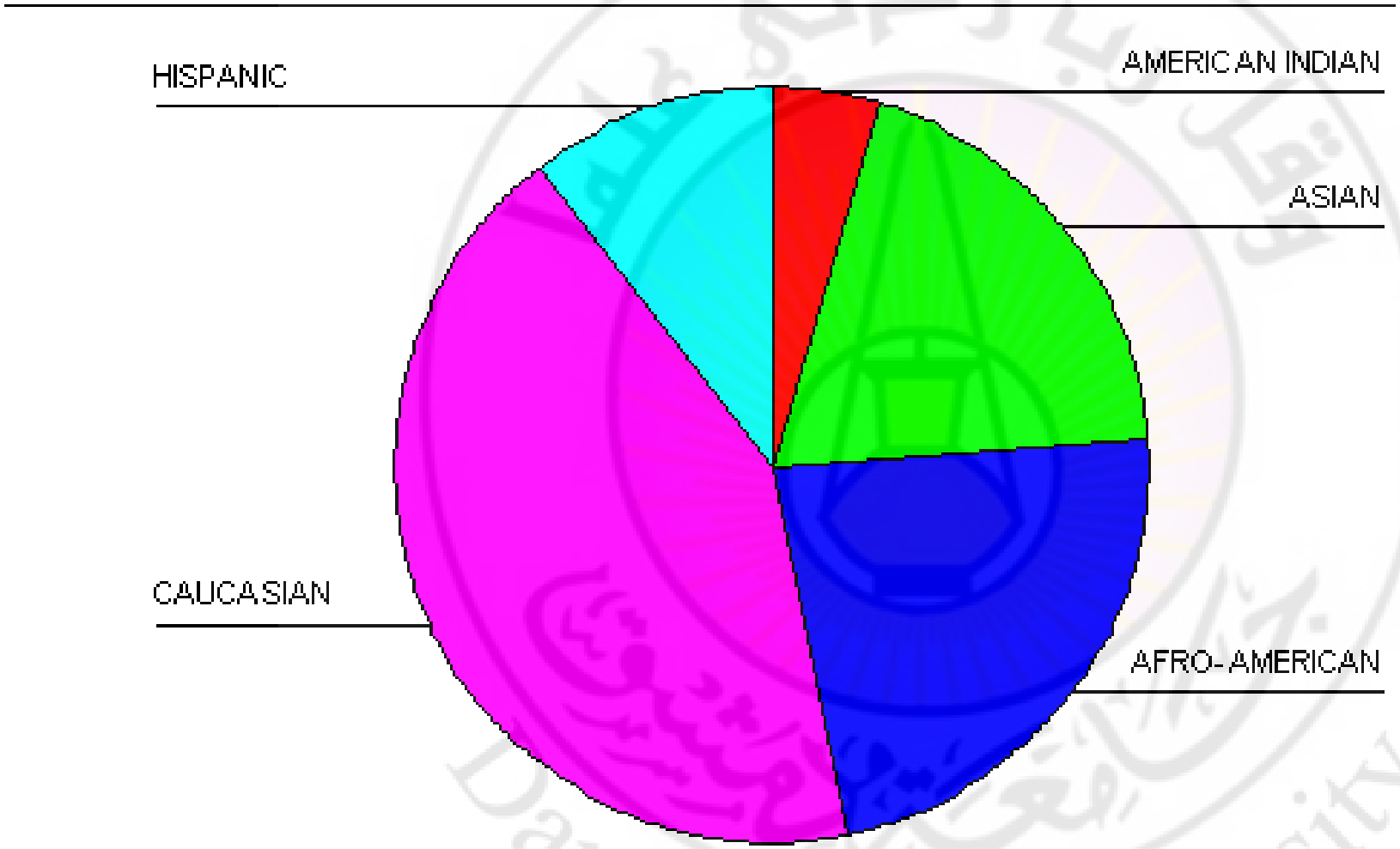
Bar Chart


50
40
30
20
10
0
Count
AMERICAN INDIAN
ASIAN
AFRO-AMERICAN
CAUCASIAN
HISPANIC
ETHNICIT

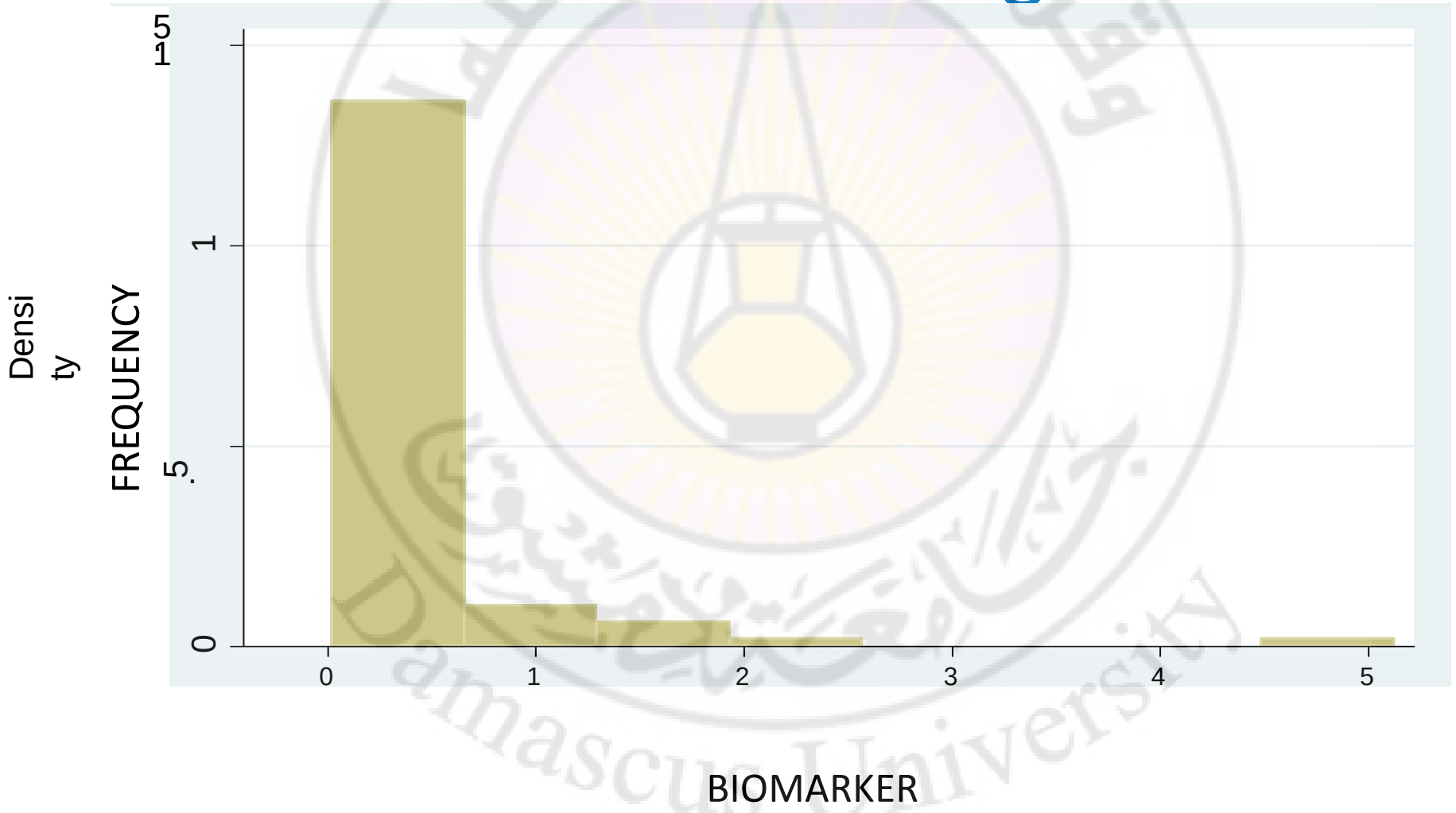
Clustered Bar Chart



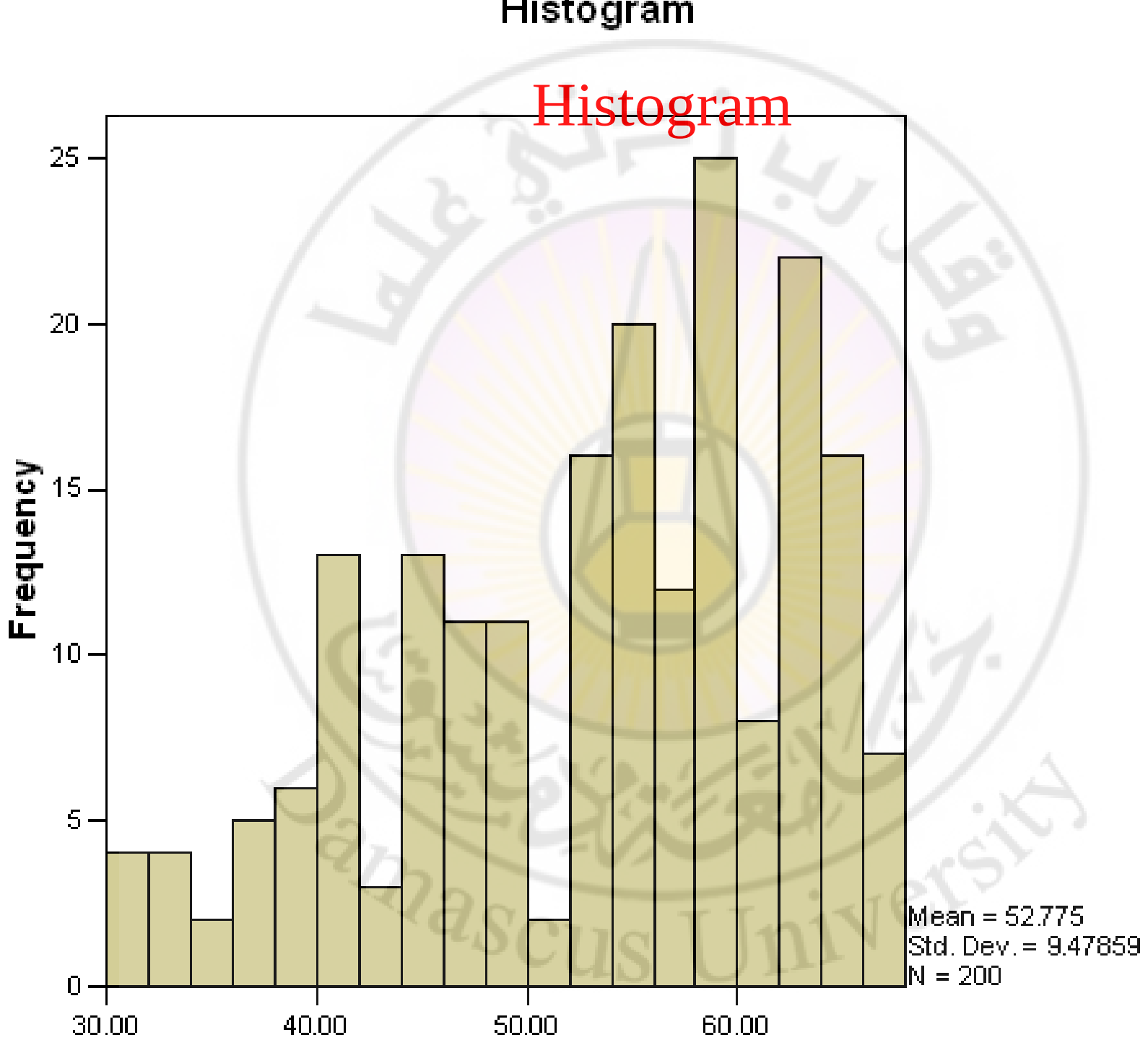
Pie Chart



Histogram

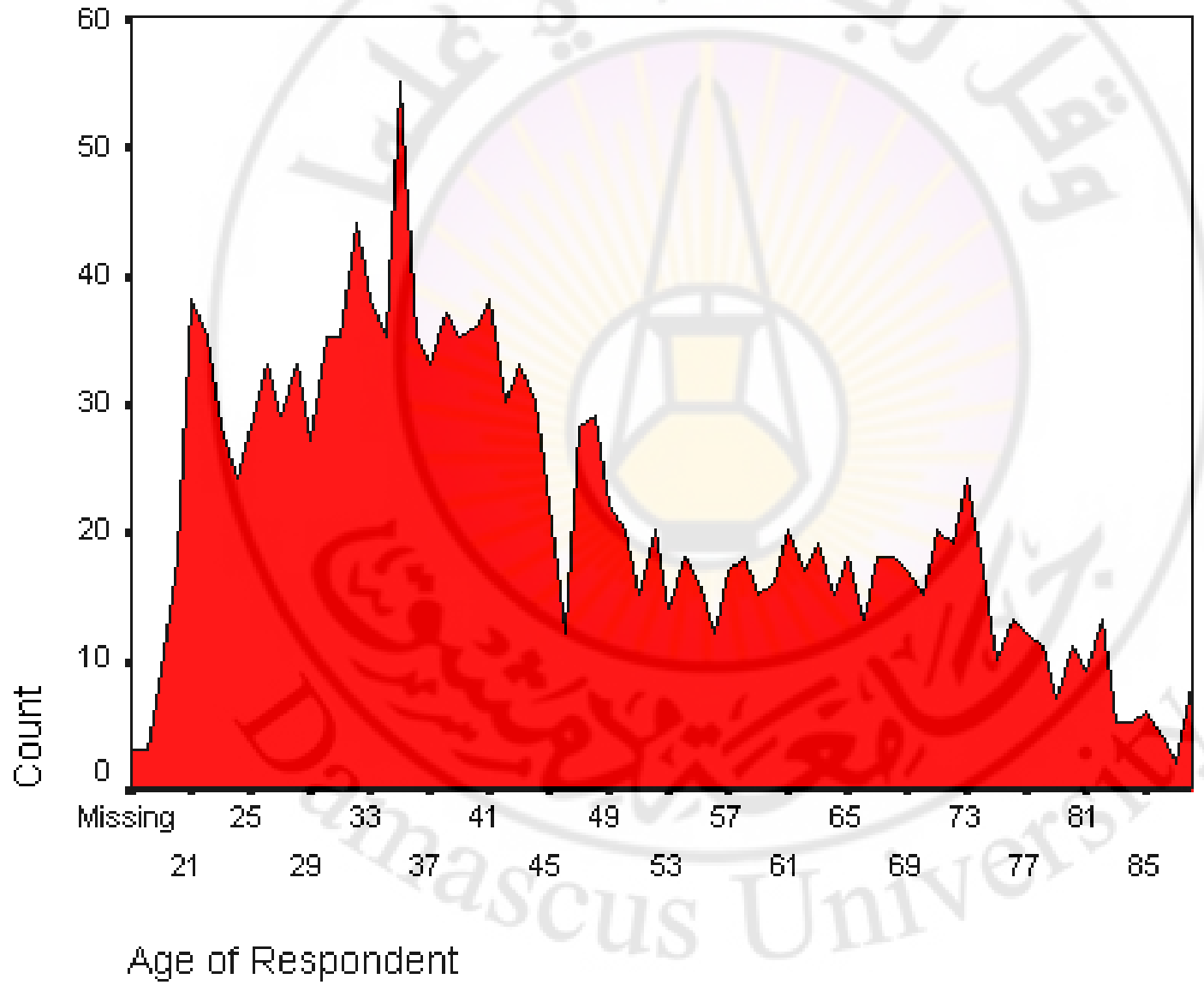


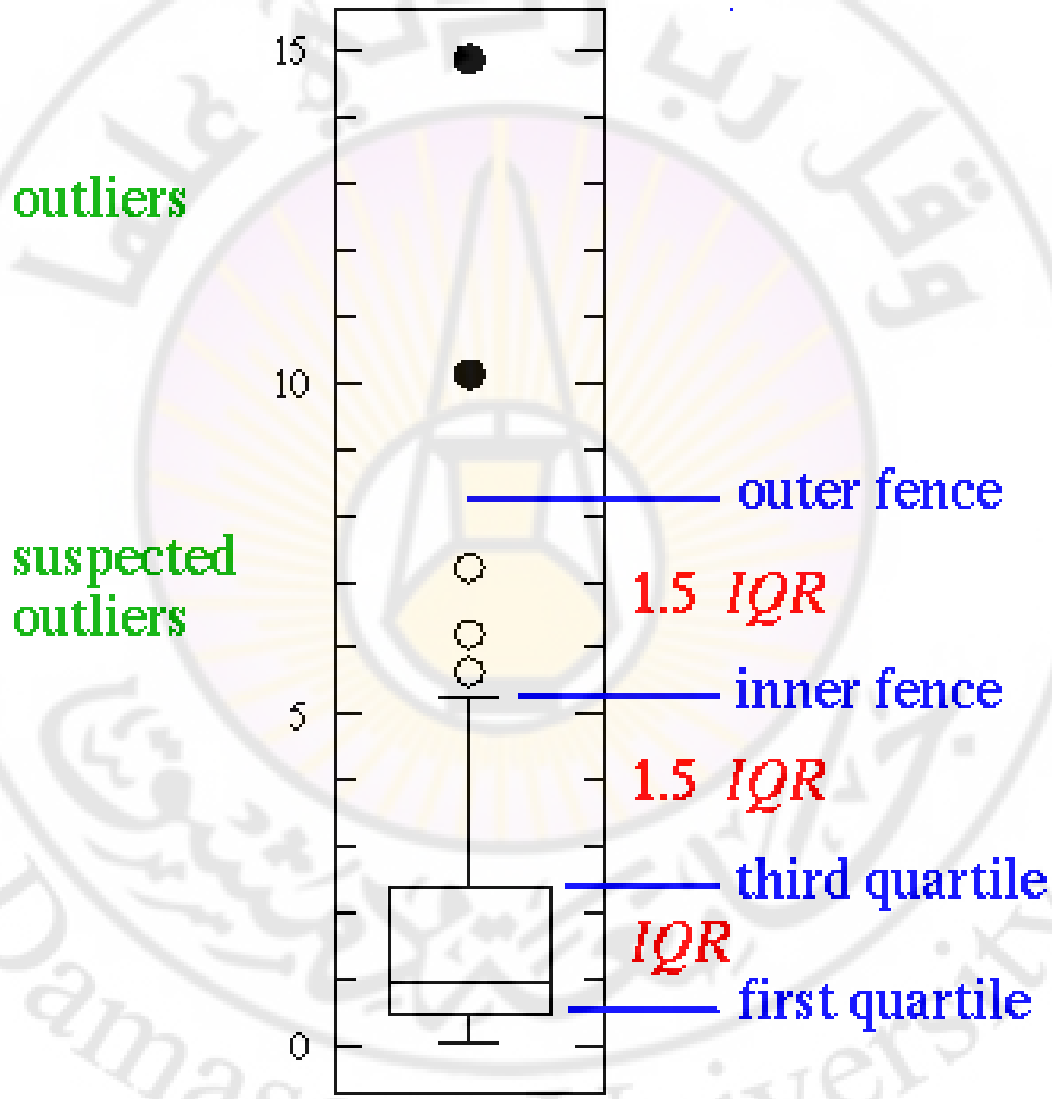
Histogram


Histogram
Frequency
25
20
15
10
5
0
30.00
40.00
50.00
60.00
Mean = 52.775
Std. Dev. = 9.47859
N = 200

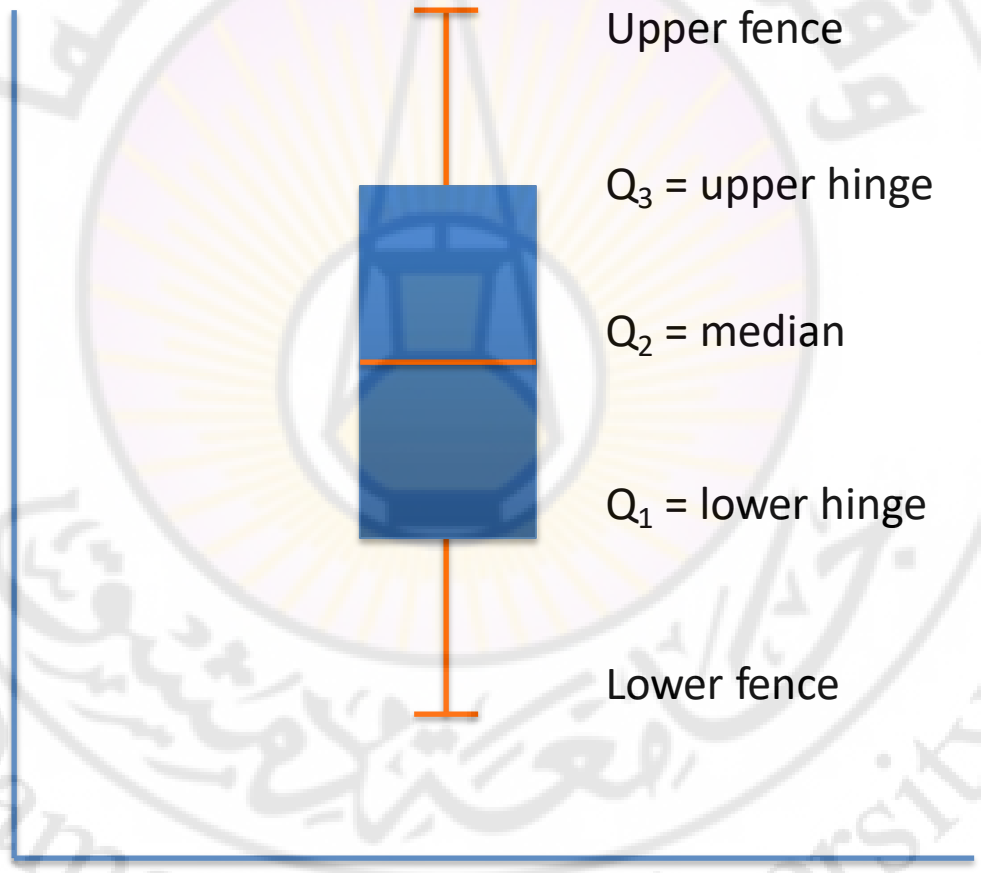
Graph

Area Chart

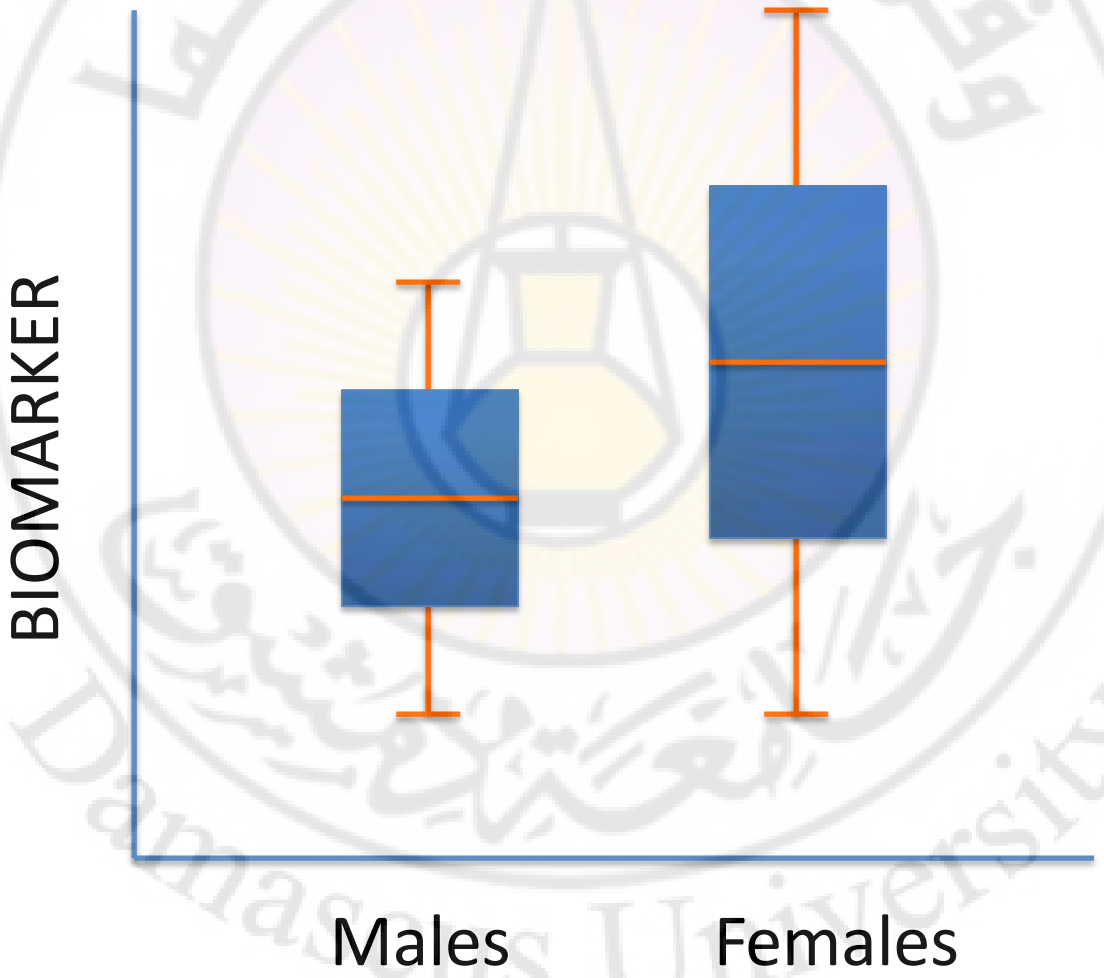



15
10
5
0
outliers
suspected outliers
outer fence
1.5 IQR
inner fence
1.5 IQR
third quartile
IQR
first quartile

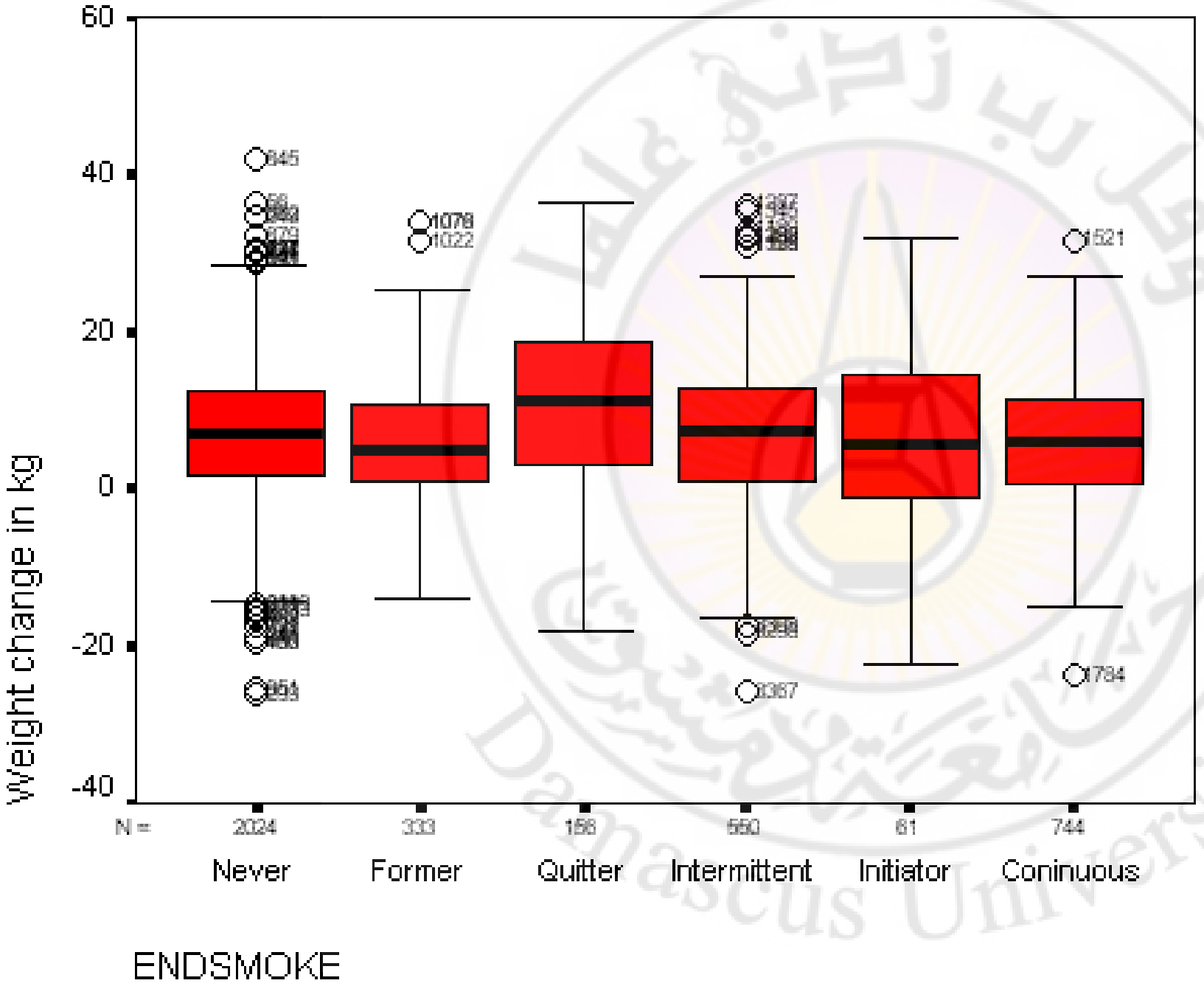
Box Plot



Side-by-Side Box Plot



Boxplot


60
40
20
0
-20
-40
Weight change in kg
N =
2024
333
158
550
61
744
Never
Former
Quitter
Intermittent
Initiator
Coninuous
ENDSMOKE

HYPOTHESIS TESTING

HYPOTHESIS TESTING

- An **objective** method of making decisions or **inferences** from sample data (evidence)
- Sample data used to choose between two choices i.e. **hypotheses** or statements about a population
- We typically do this by comparing what we have observed to what we expected if one of the statements (**Null Hypothesis**) was true

HYPOTHESIS TESTING FRAMEWORK WHAT THE TEXT BOOKS MIGHT SAY!

- Always two hypotheses:

  H_A: Research (Alternative) Hypothesis

  - What we aim to gather evidence of
  - Typically that there **is** a difference/effect/relationship etc.

  H_0: Null Hypothesis

  - What we assume is true to begin with
  - Typically that there is **no** difference/effect/relationship etc.

DISCUSSION

- Understanding what hypothesis testing is and why they need to use it?

COULD TRY EXPLAINING THINGS IN THE CONTEXT OF "THE COURT CASE"?

- Members of a jury have to decide whether a person is guilty or innocent based on evidence

  Null: The person is innocent

  Alternative: The person is not innocent (i.e. guilty)

- The null can only be rejected if there is enough evidence to doubt it
- i.e. the jury can only convict if there is beyond reasonable doubt for the null of innocence
- They do not know whether the person is really guilty or innocent so they may make a mistake

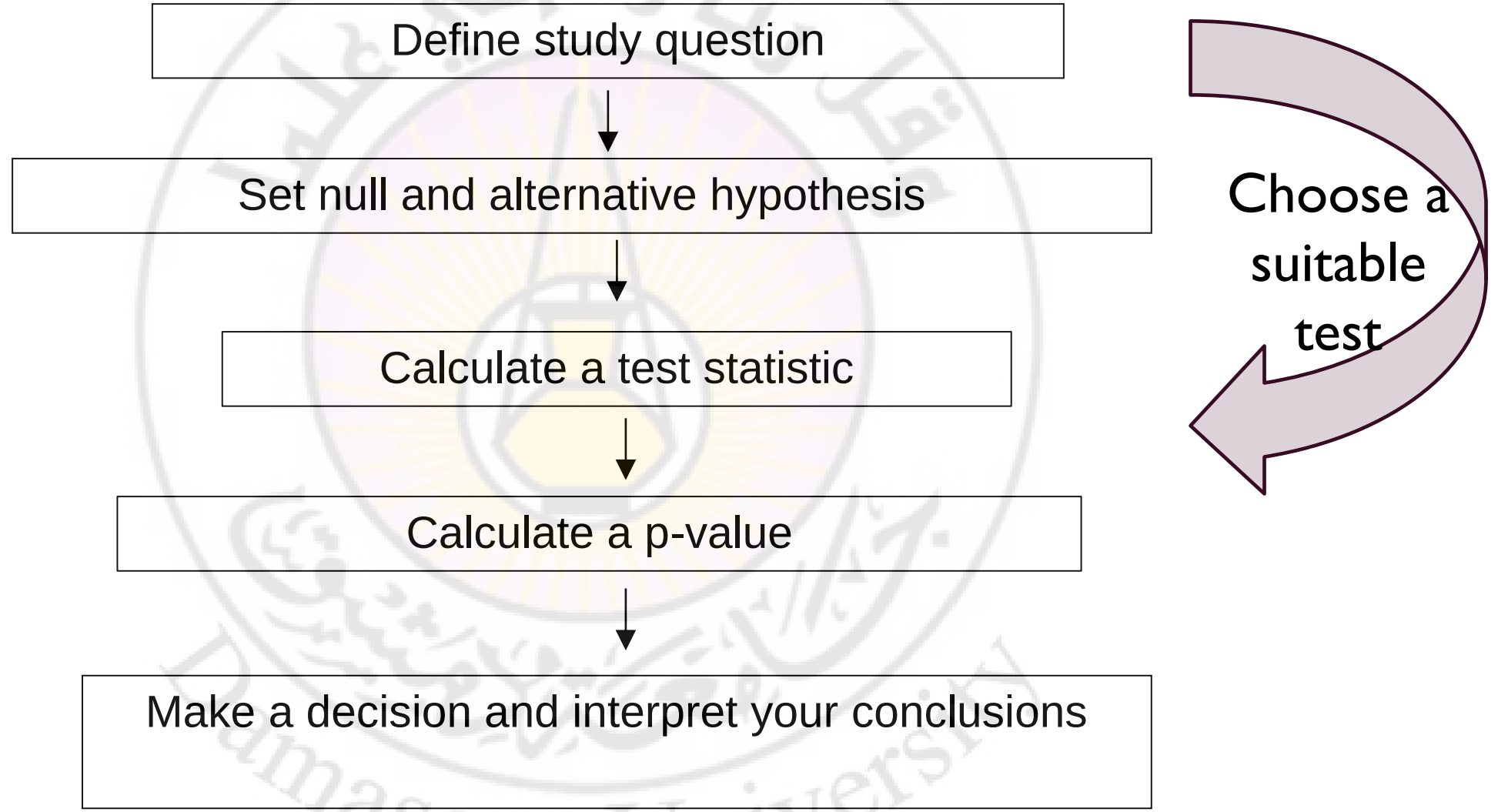
TYPES OF ERRORS

Controlled via sample size (=1-Power of test)

Typically restrict to a 5% Risk = level of significance

	Study reports **NO** difference (Do not reject H_0)	Study reports **IS** a difference (Reject H_0)
H_0 is true Difference Does **NOT** exist in population	✓	X Type I Error
H_A is true Difference **DOES** exist in population	X Type II Error	✓

Prob of this = Power of test



Elements of Testing hypothesis

- Null Hypothesis
- Alternative hypothesis
- Level of significance
- Test statistics
- P-value
- Conclusion

Case Control Study of Drinking and Lung Cancer

- Null Hypothesis There is no association between Drinking and Lung cancer
- Alternative Hypothesis There is some kind of association between Drinking and Lung cancer

Z Test:

- Numerial
- Sample to Population

T Student Test:

- Numerial
- Sample to Sample

Chi SquareTest:

- Categorial
- Sample to Sample…..

T-tests
Paired or Independent (Unpaired) Data?

T-tests are used to compare two population means

- **Paired data:** same individuals studied at two different times or under two conditions
  PAIRED T-TEST
- **Independent:** data collected from two separate groups INDEPENDENT SAMPLES T-TEST

CHI-SQUARED TEST STATISTIC

- The chi-squared test is used when we want to see if two categorical variables are related
- The test statistic for the Chi-squared test uses the sum of the squared differences between each pair of observed (O) and expected values (E)

$$\chi^2 = \sum_{i=1}^{n} \frac{(O_i - E_i)^2}{E_i}$$

Epidemiology

The STUDY OF 5Ws

- What
- Who
- Where
- When
- Why

SO WHAT !!

EPIDEMIOLOGY!!!

EPI + DEMOS + LOGY

EPIDEMIOLOGY!!!

Upon + People + Study

Epidemiology

- *The study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the prevention and control of health problems*

John Last, Dictionary of Epidemiology

The study *of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the prevention and control of health problems*

- **Surveillance**
- **Observation**
- **Hypothesis testing**
- **Analytic research and experiments**

The study of the ***distribution*** *and determinants of health-related states or events in specified populations, and the application of this study to the prevention and control of health problems*

Refers to analysis of: times, persons, places and classes of people affected

The study of the distribution and ***determinants*** *of health-related states or events in specified populations, and the application of this study to the prevention and control of health problems*

Factors that influence health: biological, chemical, physical, social, cultural, economic, genetic and behavioural

BEINGS MODEL

B	Biologic
B	Behavioural
E	Environmental
I	Immunologic
N	Nutritional
G	Genetic
S	Services, Social
S	Spiritual

The study of the distribution and determinants of ***health-related states or events*** *in specified populations, and the application of this study to the prevention and control of health problems*

diseases, causes of death, behaviours such as use of tobacco, positive health states, reactions to preventive regimes and provision and use of health services

The study of the distribution and determinants of health-related states or events in specified populations, and the ***application*** *of this study* ***to the prevention and control of health problems***

the aims of public health—to promote, protect, and restore health

So

EPIDEMIOLOGY IS:

The study of the distribution and determinants of health-related states or events in specified populations, and the application of this study to the prevention and control of health problems

EPIDEMIOLOGY!!!

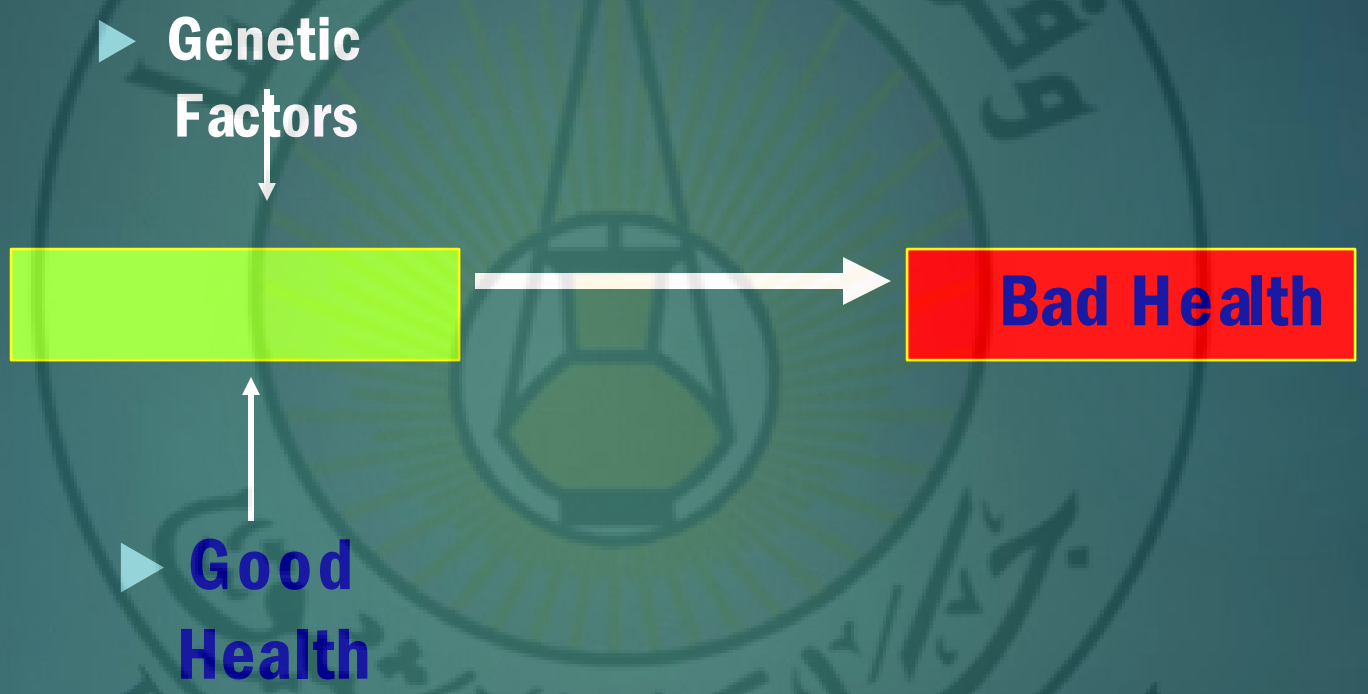
One of the tools for improving

public health

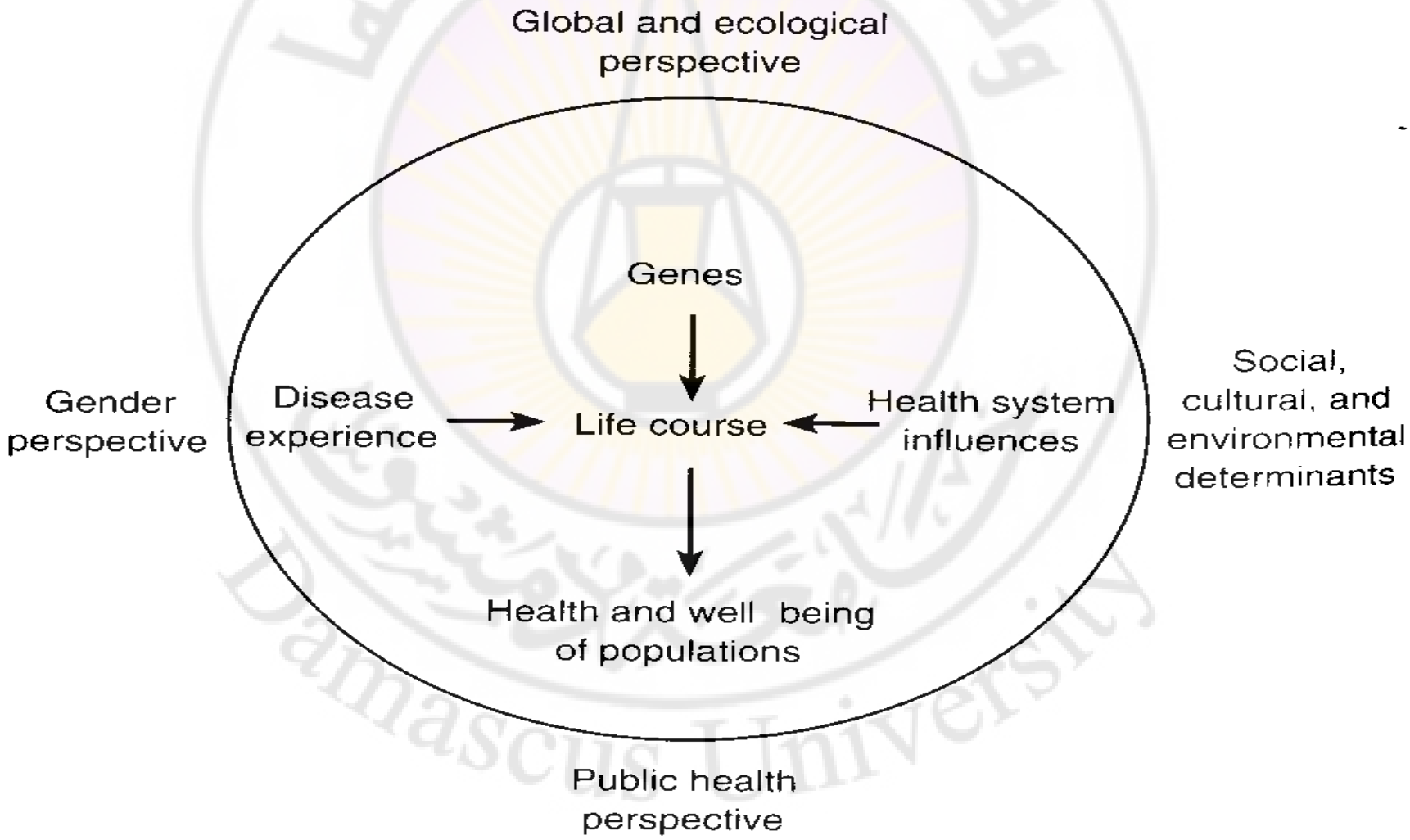
Uses of Epidemiology in Public Health

- Causation of disease
- Natural history of disease
- Health status of populations
- Evaluating interventions

Causation



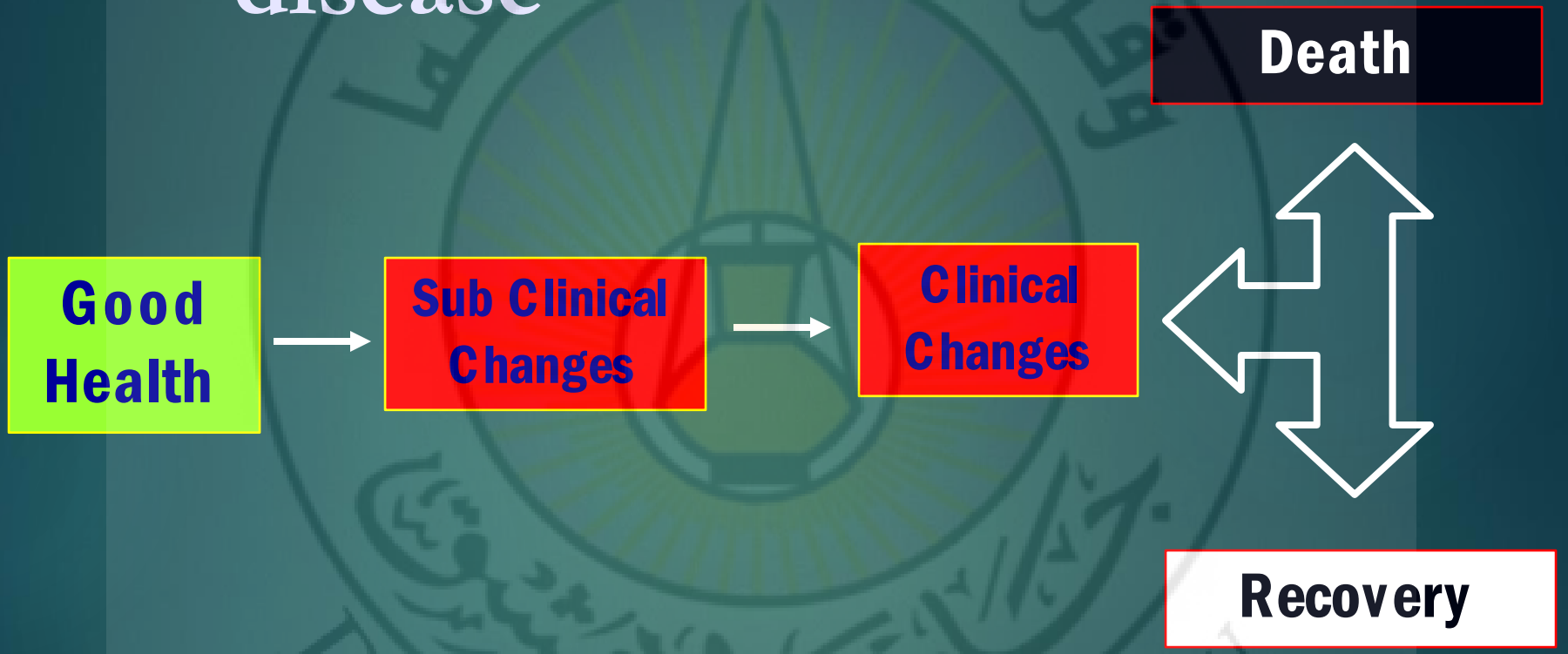
Determinants of Health



Uses of Epidemiology in Public Health

- Causation of disease
- Natural history of disease
- Health status of populations
- Evaluating interventions

Natural history of disease



So

Epidemiology is concerned with the course and outcome (natural history) of diseases in individuals and groups

Uses of Epidemiology in Public Health

- Causation of disease
- Natural history of disease
- Health status of populations
- Evaluating interventions

Health status of populations

Figure 1.5. Describing the health status of populations

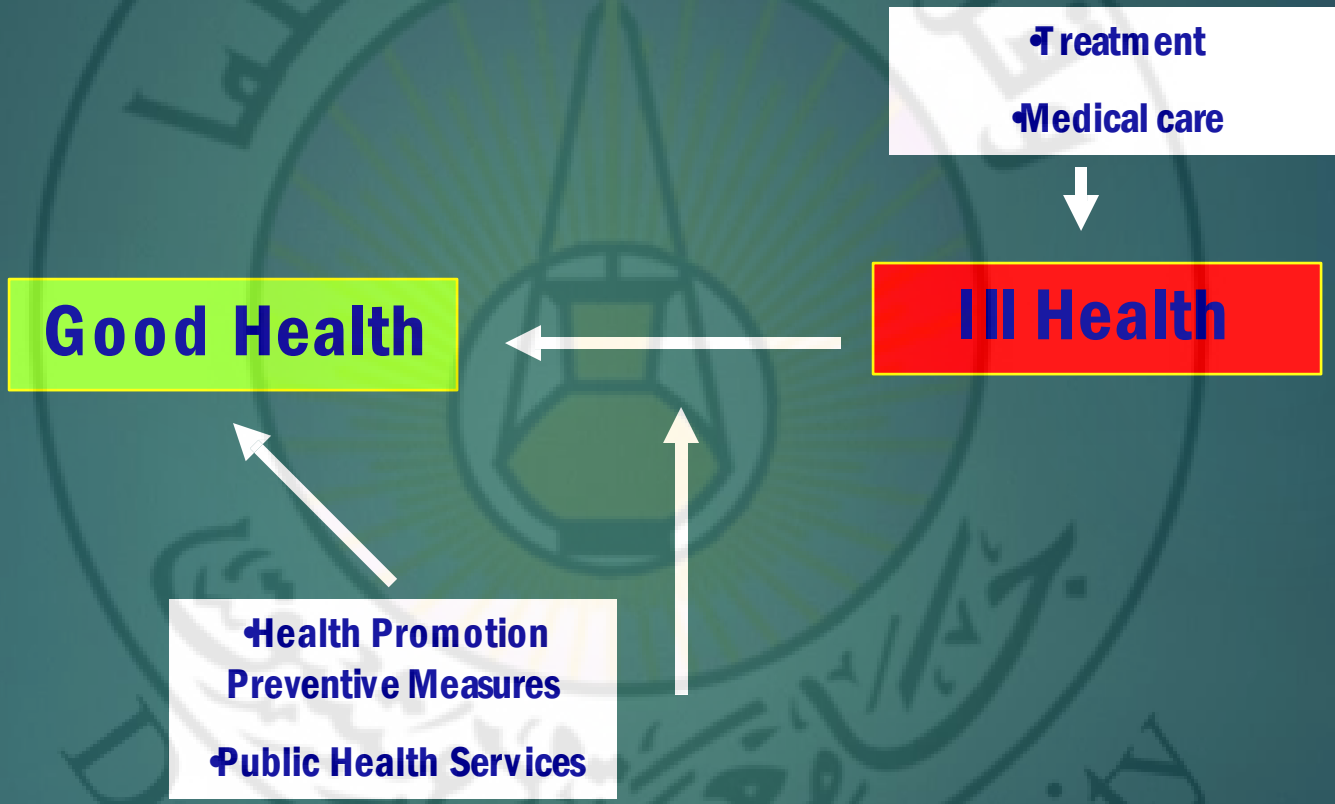
- Epidemiology is often used to describe the health status of population groups

What is Population??

Uses of Epidemiology in Public Health

- Causation of disease
- Natural history of disease
- Health status of populations
- Evaluating interventions

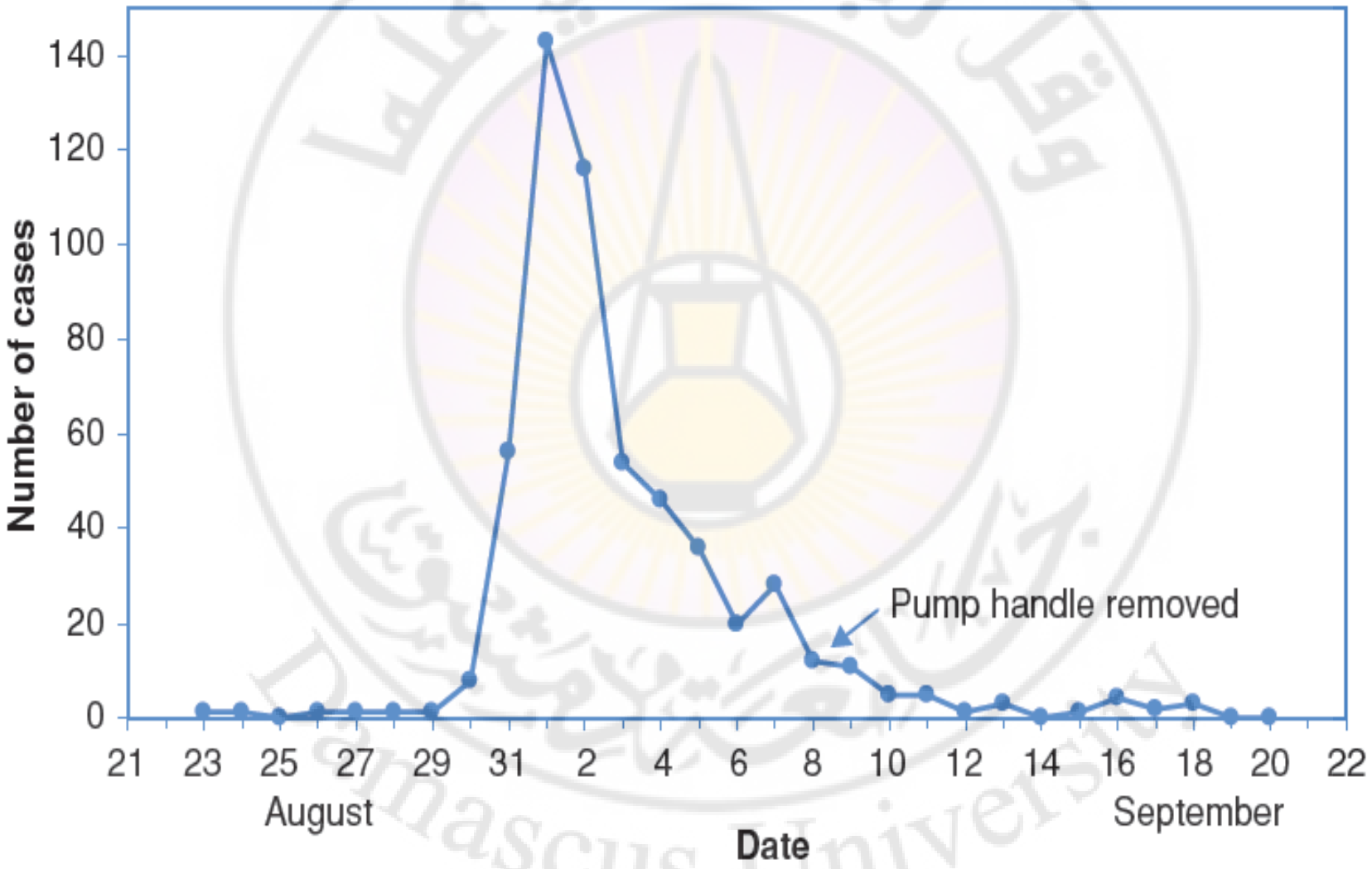
Evaluating interventions



Epidemiology

- John Snow

John Snow


Number of cases
140
120
100
80
60
40
20
0
Pump handle removed
21
23
25
27
29
31
2
4
6
8
10
12
14
16
18
20
22
August
September
Date

What does epidemiology offer?

- Description of health status of populations
- Causation
- Evaluation of interventions
- Natural history and prognosis

Food	People who ate the food		People who didn't eat the food	
	Total	Number ill	Total	Number ill
Friday dinner:				
Hot chicken	343	156	231	74
Peas	390	175	184	55
Potato fries	422	184	152	46
Saturday lunch:				
Cold chicken	202	155	372	75
Salad	385	171	189	59
Saturday dinner:				
Fruit salad	324	146	250	84

$$156 \div 343 = 0.45 = 45\%$$

- 45% of people who ate hot chicken became sick. This is known as the **attack rate** for hot chicken

Food	People who ate the food			People who didn't eat the food		
	Total	Number ill	Attack rate	Total	Number ill	Attack rate
Friday dinner:						
Hot chicken	343	156	45%	231	74	32%
Peas	390	175	45%	184	55	30%
Potato fries	422	184	44%	152	46	30%
Saturday lunch:						
Cold chicken	202	155	77%	372	75	20%
Salad	385	171	44%	189	59	31%
Saturday dinner:						
Fruit salad	324	146	45%	250	84	34%

	People who ate the food			People who didn't eat the food			
Food	Total	Number ill	Attack rate	Total	Number ill	Attack rate	Relative ris
Friday dinner:							
Hot chicken	343	156	45%	231	74	32%	1.4
Peas	390	175	45%	184	55	30%	1.5
Potato fries	422	184	44%	152	46	30%	1.4
Saturday lunch:							
Cold chicken	202	155	77%	372	75	20%	3.8
Salad	385	171	44%	189	59	31%	1.4
Saturday dinner:							
Fruit salad	324	146	45%	250	84	34%	1.3

Subdisciplines of epidemiology

- *Public health epidemiology*
- *Infectious disease epidemiology*
- *Nutritional epidemiology*
- *Social epidemiology*
- *Environmental epidemiology*

Subdisciplines of epidemiology

- *Cancer epidemiology*
- *Injury epidemiology*
- *Perinatal epidemiology*
- *Occupational epidemiology*

Subdisciplines of epidemiology

- ***Molecular epidemiology***
- ***Clinical epidemiology***

Molecular Epidemiology

Measures exposure to specific substances and early

biological response, by:

- **evaluating host characteristics mediating response to external agents, and**
- **using biochemical markers of a specific effect to refine disease categories.**

Genetic Epidemiology

- **D eals with the etiology, distribution, and control of disease in groups of relatives, and with inherited causes of disease in populations.**

Clinical epidemiology

- **Applying epidemiological principles and methods to problems encountered in the practice of medicine.**

“**Determinants**” is defined as a collective or individual risk factor (or set of factors) that is causally related to a health condition, outcome, or other defined characteristic.”

- Related to determinants are **exposures**, which pertain either to contact with a disease-causing factor or to the amount of the factor that impinges upon a group or individuals.
- Epidemiology searches for associations between exposures and health outcomes.

- The term ***risk factor*** (an exposure that increases the probability of a disease or adverse health outcome)

Outcomes is all the possible results that may stem from exposure to a causal factor….”

Epidemiology

Who, what, where and when?
Descriptive Epidemiology

Case reports and case series

- The identification of a new or recurring health problem often begins with a **case report** or **case series**.
- Detailed descriptions….. doctor or group of doctors….. of one or more cases of a disease that are unusual for some reason.

Case reports and case series

- Disease might has not been seen before or the cases may have occurred in unusual individuals or area where the disease had not previously been reported or was thought to have been controlled.

Case reports and case series

- **The classic description of a series of infants born with congenital cataracts, some with additional cardiac abnormalities in Sydney to postulate a causal link between a severe epidemic of rubella (German measles) that had occurred six to nine months before the children were born and the subsequent abnormalities (Gregg, 1941).**

Case reports and case series

- **A case report published in the UK in 1961 described the development of a pulmonary embolism in a 40-year-old pre-menopausal woman, five weeks after she had started using an oral contraceptive (OC) to treat endometriosis (Jordan, 1961).**

Case reports and case series

- Provide little evidence of causality
- Cannot say much about patterns of disease occurrence.
- Help identify potential health problems such as the acute outbreaks of severe acute respiratory syndrome (SARS
- They may also stimulate interest in an area, leading on to more detailed studies, and in this regard some have been seminal in advancing knowledge

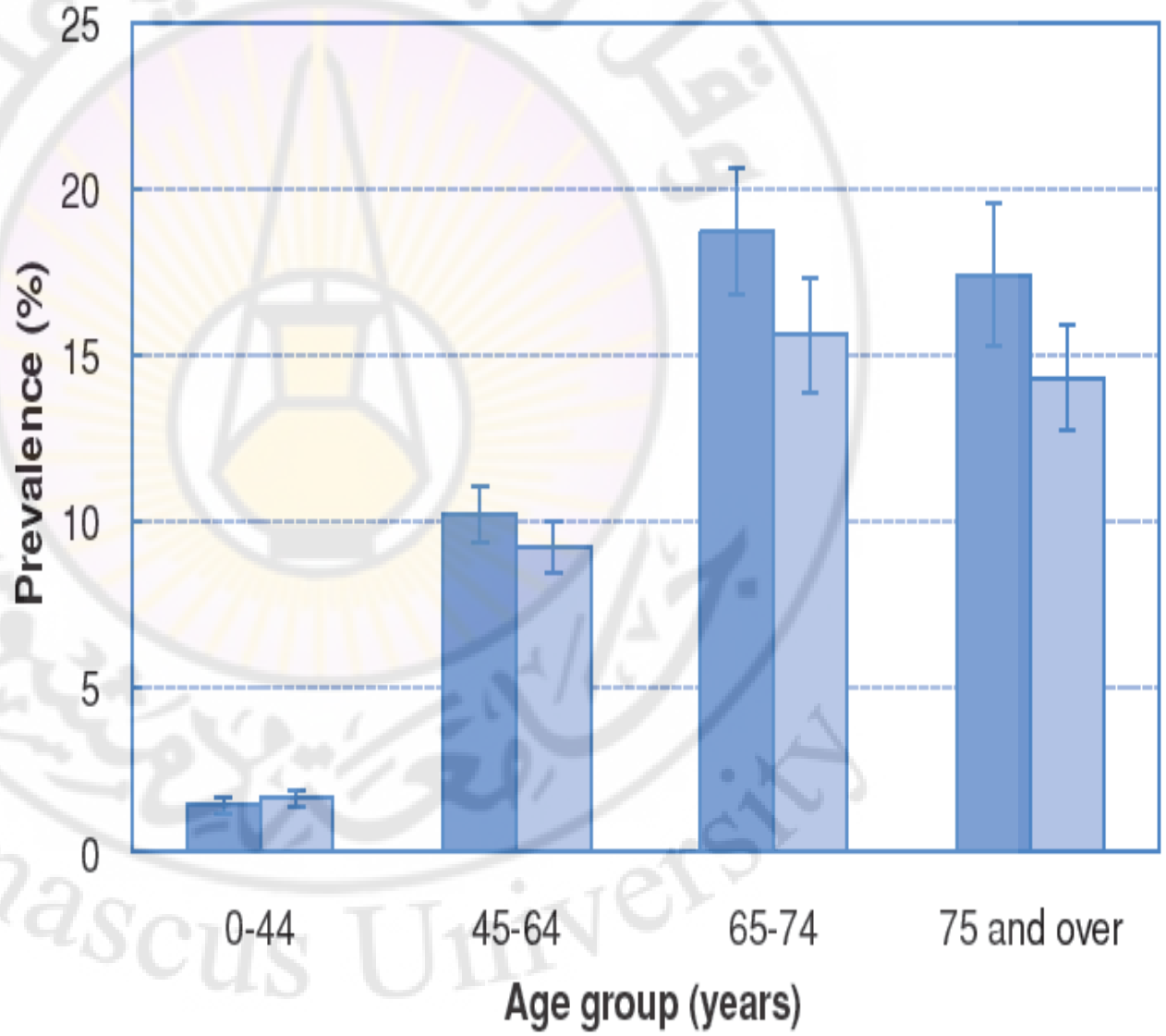


Figure 3.1 Prevalence of diagnosed diabetes (self-reported) in the United States, 2006, by age group and sex (dark bars, males; light bars, females). (*Source:* Centers for Disease Control http://www.cdc.gov/diabetes/statistics/prev/national/fig2004.htm, accessed on 16 January 2010.)

Prevalence surveys

- Surveys are conducted to measure the prevalence of a wide variety of aspects of health, including diseases that are not captured by other routine statistics; conditions such as obesity; health-related behaviors such as smoking, sun-exposure and diet; and use of health services.

Prevalence surveys

- These spot checks on the health of a nation or region are crucial to expanding our understanding of health burdens, needs and services beyond the hospital sector. In recent decades they have become a feature of broad-based community diagnosis and health planning, using a wide range of sampling and data-capture designs.

Prevalence surveys

- Telephone and face-to-face interviews, and sometimes very detailed physical examinations, as in the US National Health and Nutrition Examination Surveys (NHANES)

Routine data collections

- Governments, healthcare providers and statistical agencies routinely collect vast amounts of information that we will collectively describe as ‘routine’ data.

Table 3.2 Some common health data collections and reporting systems.

Data collection or reporting system	Summary data often published	Individual level data available	Source of raw data
Vital statistics	Mortality rates	Date and cause of death, demographics[a]	Death certificates
Disease registries (e.g. cancer registries, injury registers)	Incidence, mortality and survival rates, prevalence	Diagnosis, date and demographics	Pathology reports, testing laboratories, hospital and medical records
Notifiable diseases (e.g. AIDS, SARS, TB, other infectious diseases)	Numbers of cases, incidence	Diagnosis, date and demographics	Laboratories, medical practitioners and hospitals

[a] Basic demographic information such as age, sex and last known address.

Table 3.2 Some common health data collections and reporting systems.

Data collection or reporting system	Summary data often published	Individual level data available	Source of raw data
Hospital administrative systems	None	Diagnosis, date and demographics	Hospital discharge sheets, medical records
Health surveys (morbidity, risk factors, needs, service use etc.)	Special reports	Self-reported health states	Special surveys (often whole population)
Special surveillance systems	Varied	Varied	e.g. 'sentinel' primary care practices or disease registers (UK GP data base), MONICA (international CHD)
Rapid community assessments (health, nutrition)	Varied	Varied	Special surveys (sometimes of targeted groups)

[a] Basic demographic information such as age, sex and last known address.

Routine data collections

Mortality data

- Death certificates
- Verbal autopsy

Medical Certificate of the Cause of Death

To the Registrar-General

I hereby certify that _____ (name in full)

aged _____ years, date of birth __/__/__ who usually resided at

_____ Postcode was attended and last

seen by me on __/__/__ (or by* Dr. _____ on __/__/__

*If not attended by certifying medical practitioner within 3 months prior to death, insert name of medical practitioner who last attended deceased and date)

and I am informed that he/she died on __/__/__ at _____

(town, place etc of death)

Cause of Death (print clearly and do not abbreviate)

		Duration of last illness
Disease or condition directly leading to death (This means the disease, injury or complication which caused death – NOT ONLY for example, the mode of dying such as heart failure', 'asphyxia', etc)	1a	
due to, or as a consequence of		
	1b	
due to, or as a consequence of		
Antecedent causes - morbid conditions, if any giving rise to the above cause, stating the underlying condition last	1c	
due to, or as a consequence of		
	1d	
Other significant conditions Contributing to the death, but not related to the disease or condition causing it	2	

Date and type of operation in the last 4 weeks __/__/__ _____

Was a Coroner consulted before issuing this certificate?

No, death not subject to the provisions of the Coroners Act ☐

Yes, issue of this certificate agreed to by _____, Coroner

Signature of Medical Practitioner _____ Date __/__/__

Initials and Surname (BLOCK Letters) _____

Professional Qualifications _____

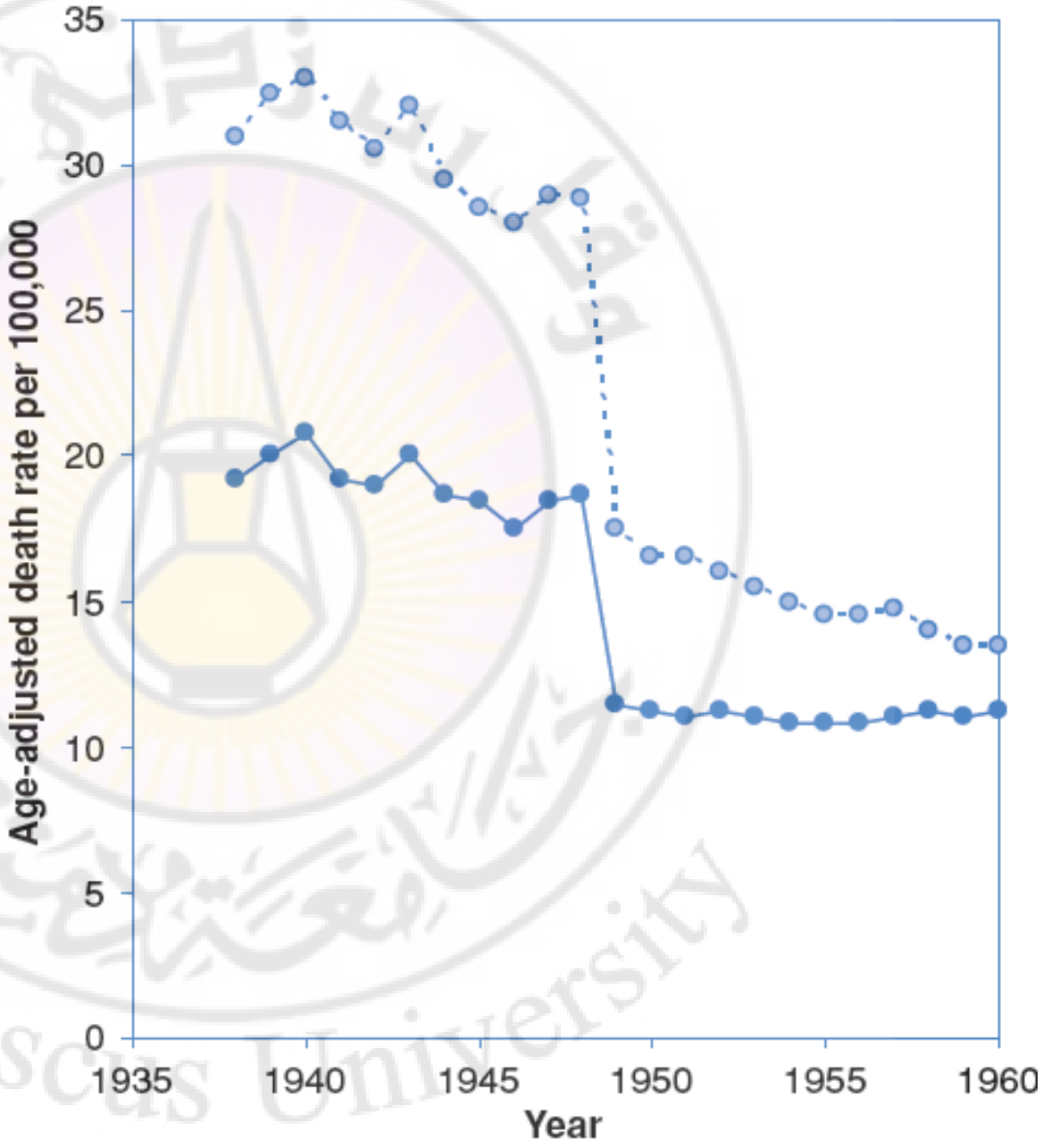


Figure 3.3 Age-adjusted mortality rates for diabetes by gender in the USA, 1938–1960, Whites only. Figure 3–2, p. 51, from *Methods in Observational Epidemiology*, 2nd edition, by Jennifer L. Kelsey, W. Douglas Thompson and Alfred S. Evans, copyright 1996 by Oxford University Press. Used by permission of Oxford University Press.

Routine data collections

Morbidity data

- Disease registries
- Hospital records

Sources of summary data

- International sources
- National data sources
- Disease-based resources
- Some academic and other idiosyncratic sites

Creative use of existing data

- Migrant studies
- Ecological or correlation studies

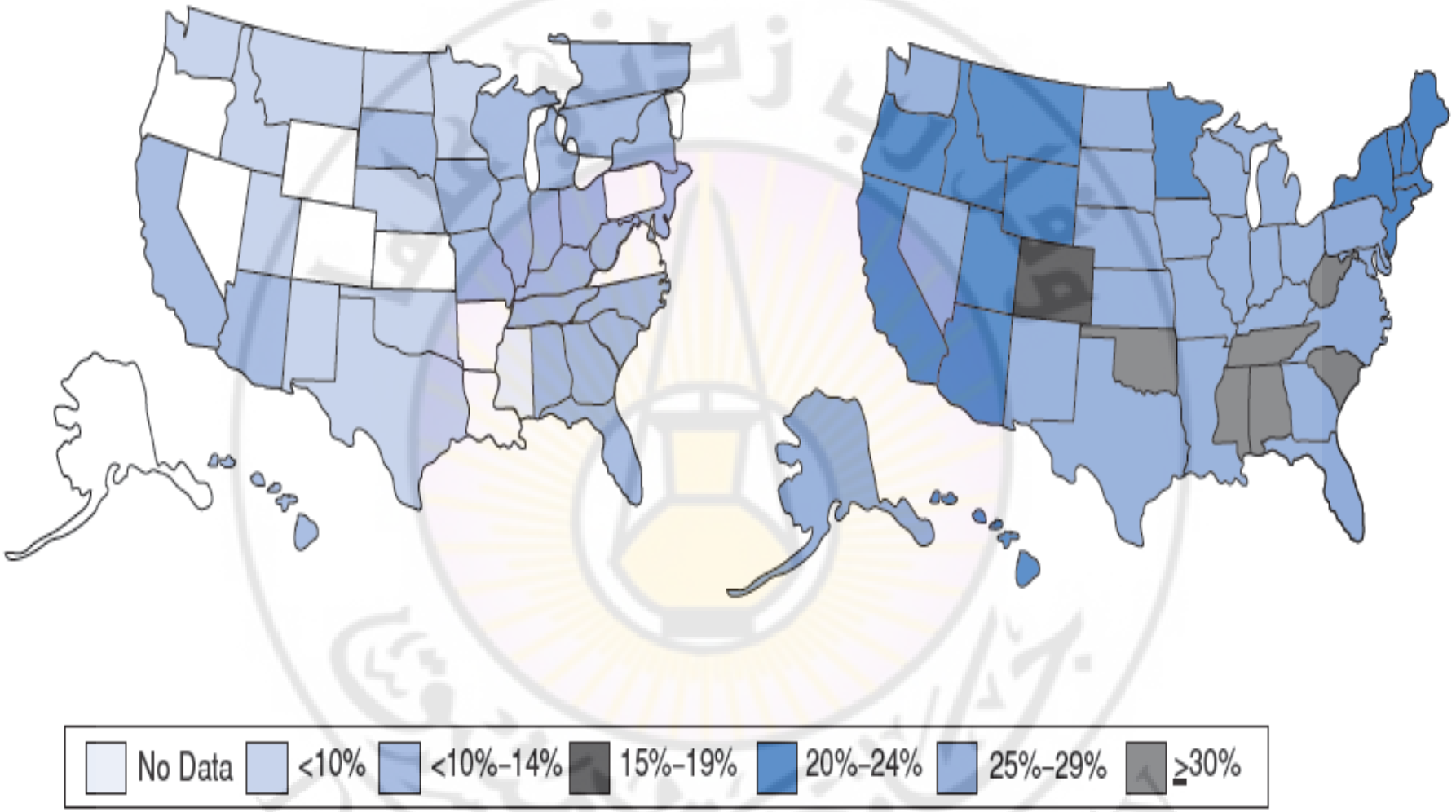


Figure 3.4 The prevalence of obesity (body-mass index ≥ 30 kg/m^2) by state in the USA in 1988 and 2008. (Data from the CDC Behavioral Risk Factor Surveillance System, BRFSS, accessed via http://www.cdc.gov/obesity/data/trends.html on 16 January 2010.)

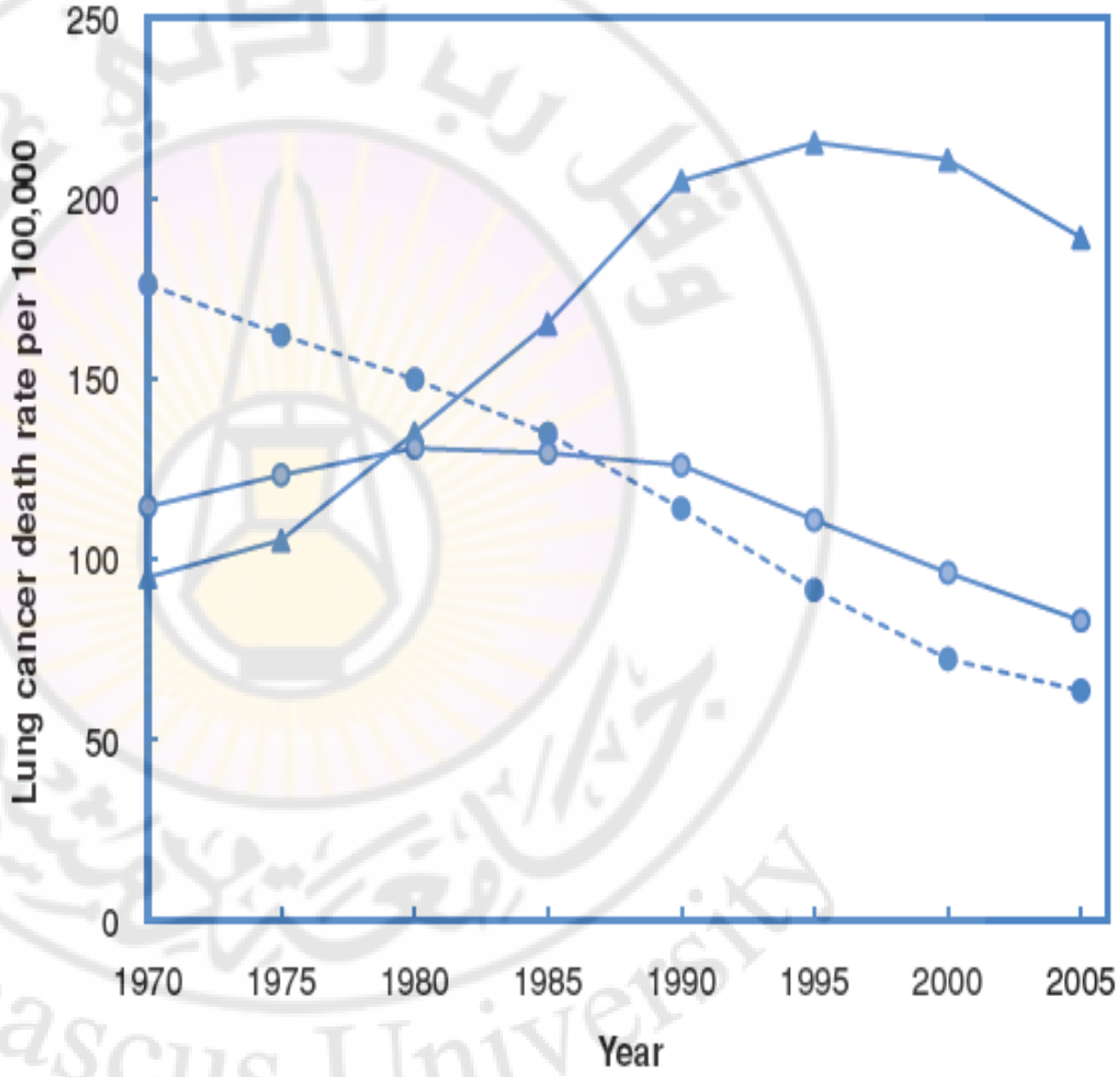


Figure 3.5 Age-standardised annual death rates from lung cancer among men aged 40–69 years in Hungary (▲), the USA (○) and the UK (●). (*Data source*: WHO Mortality Database, accessed via http://www-depdb.iarc.fr/ on 16 January 2010.)

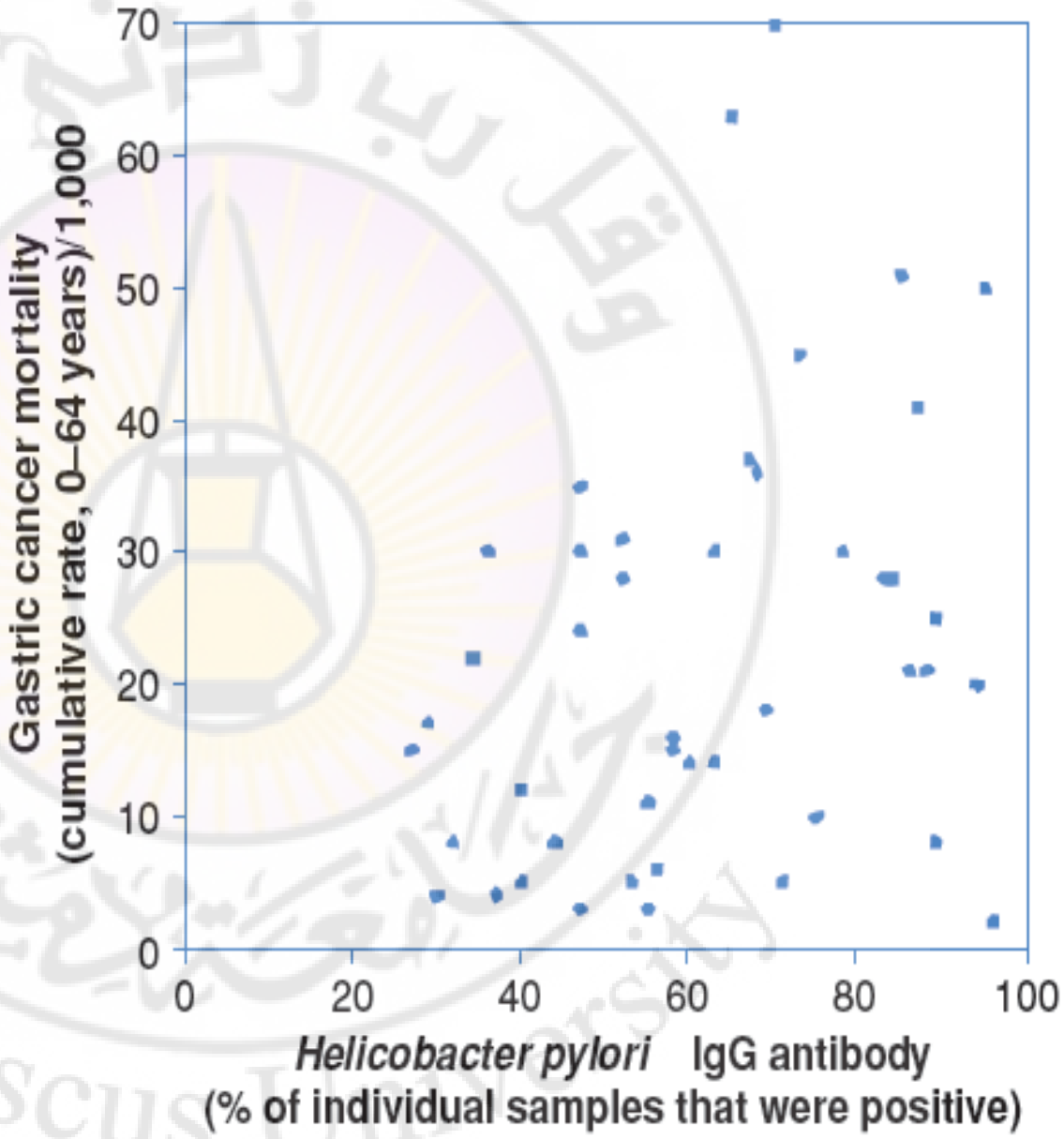
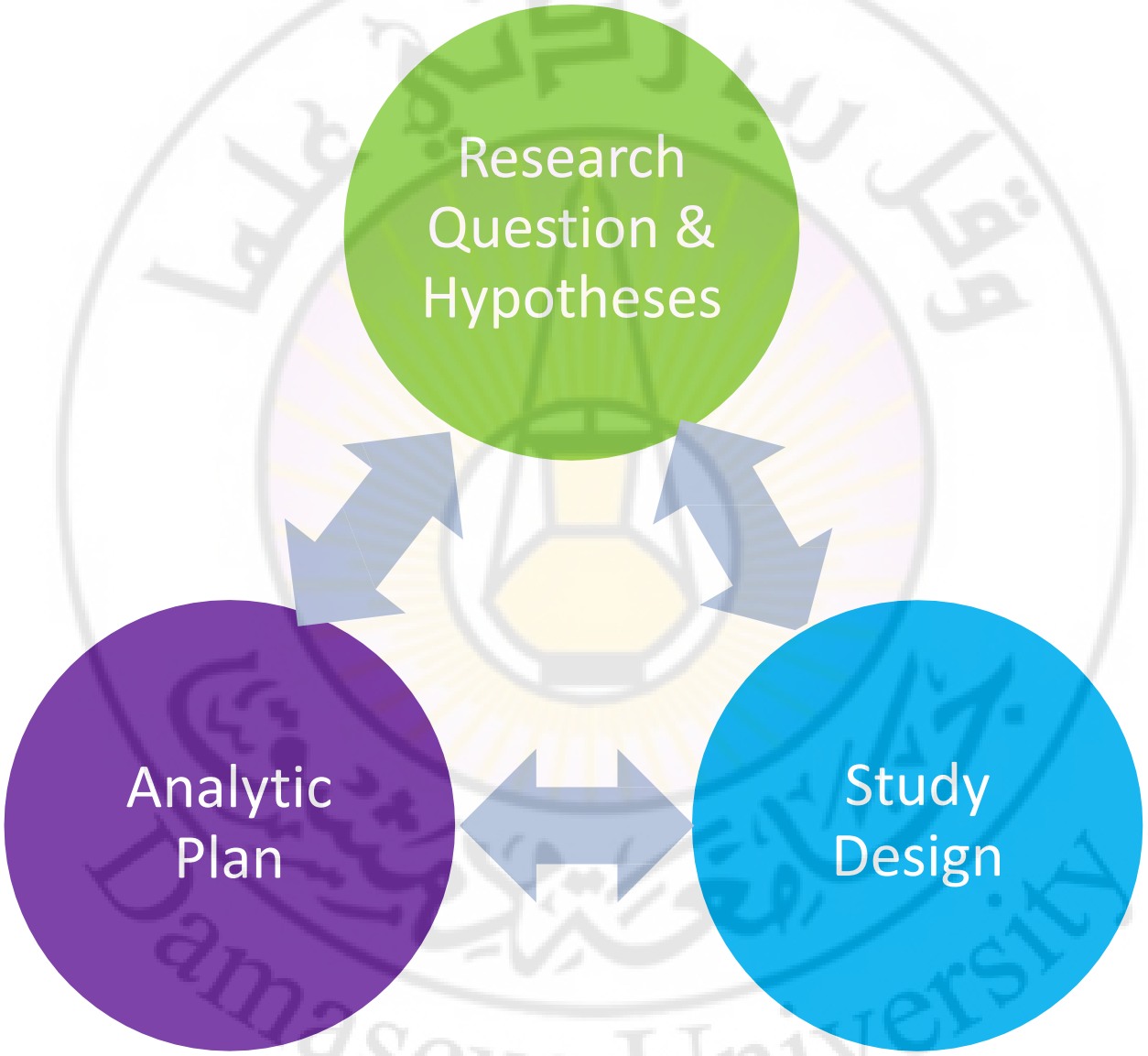


Figure 3.7 An ecological study comparing the prevalence of serum antibodies to *H. pylori* (a gastric infection) and gastric cancer mortality rates in 46 rural Chinese counties. (From Forman *et al.*, Geographic association of *Helicobacter pylori* antibody prevalence and gastric cancer mortality in rural China. *Int. J. Cancer*, 1990; 46: 608–611, reprinted by permission of John Wiley & Sons.)

Epidemiologic Study Designs


Research Question & Hypotheses
Analytic Plan
Study Design

Basic Study Designs and their Hierarchy

Study Design in Epidemiology

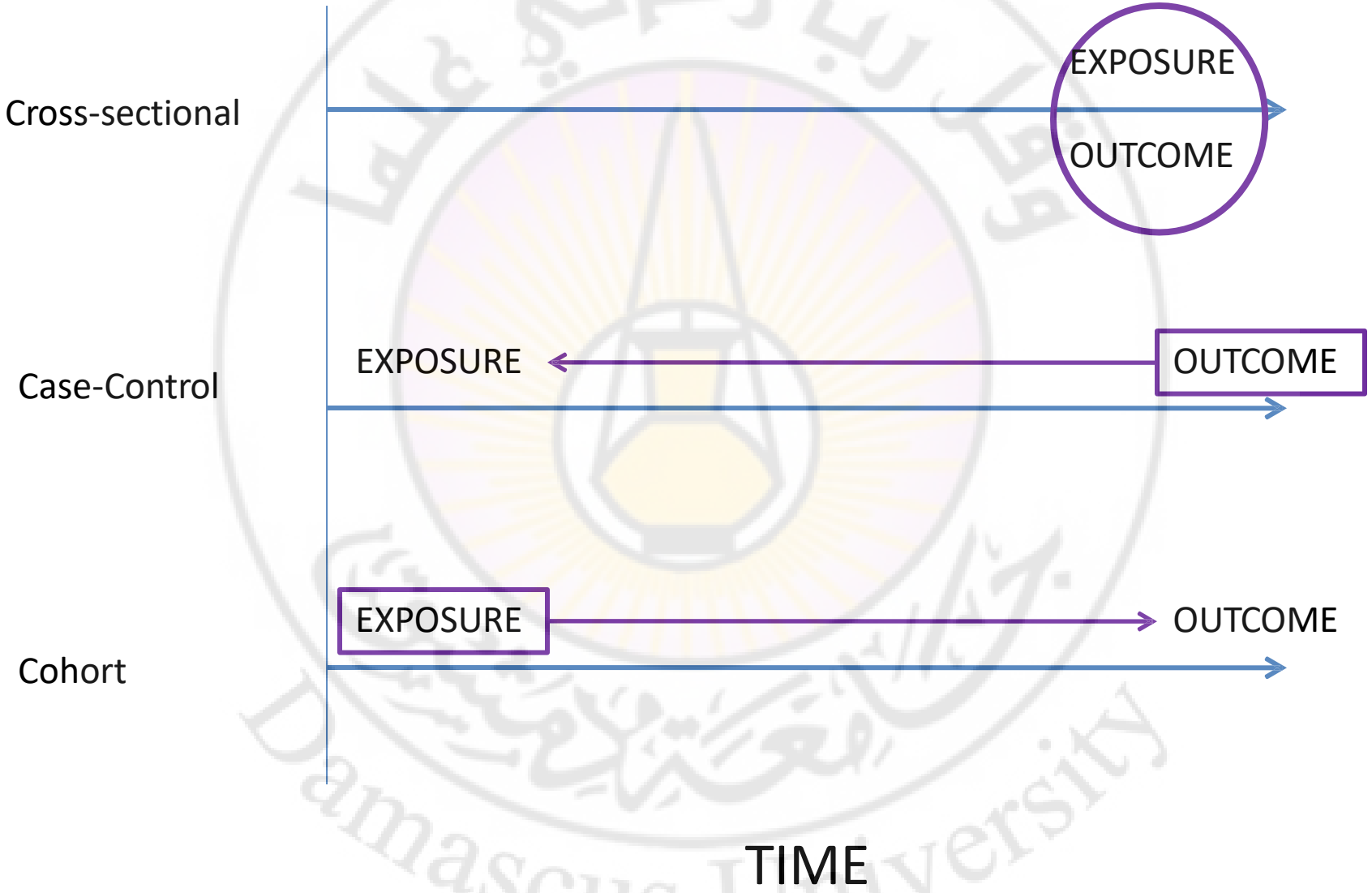
- Depends on:
  - **The research question and hypotheses**
  - Resources and time available for the study
  - Type of outcome of interest
  - Type of exposure of interest
  - Ethics

Study Design in Epidemiology

- Includes:
  - The research question and hypotheses
  - Measures and data quality
  - Time
  - Study population
    - Inclusion/exclusion criteria
    - Internal/external validity

Epidemiologic Study Designs

- Descriptive studies
  - Seeks to measure the frequency of disease and/or collect descriptive data on risk factors
- Analytic studies
  - Tests a causal hypothesis about the etiology of disease
- Experimental studies
  - Compares, for example, treatments


Cross-sectional
EXPOSURE
OUTCOME
Case-Control
EXPOSURE
OUTCOME
Cohort
EXPOSURE
OUTCOME
TIME

Epidemiology

Epidemiology as a Population Science
Basic Epidemiology Measures

- Understanding epidemiology as a science focused on populations
- Samples of the population are taken to assess health issues
- Health outcomes data can be expressed through multiple measures
- These measures can be expressed as differing metrics

What is Epidemiology?

Epidemiology

- Epidemiology is the study of the determinants, distribution, and frequency of health-related states or events in specified populations, and the application of this study to the control of health problems

Epidemiology

- Epidemiology ……

- Who gets disease and why
- Epidemiologists study sick and well people to determine the crucial difference between those who get disease and those who are spared

Purpose of Epidemiology

- To provide a basis for developing disease control and prevention measures for groups at risk. This translates into developing public health measures to prevent or control disease.

Population Focus

- The focus of epidemiology is on the occurrence of health and disease in the population.
  - What happens to many
- The population approach contrasts with clinical medicine's primary concern with health and disease in the individual.
  - What happens to one

What is a Population?

- The common definition of a **population** is "All the inhabitants of a given country or area considered together;…"
- A "population" can also be groups of individuals that share a common thread
  - Clinical populations
  - Subgroups of the population by age, race, etc…

Epidemiology is…

- The study of disease and its treatment, control, and prevention in a population of individuals.
- Whole populations may be examined, but…
- More frequently, samples of the population may be examined. Samples that are studied must be representative of the population for the results to be generalized to the total population.

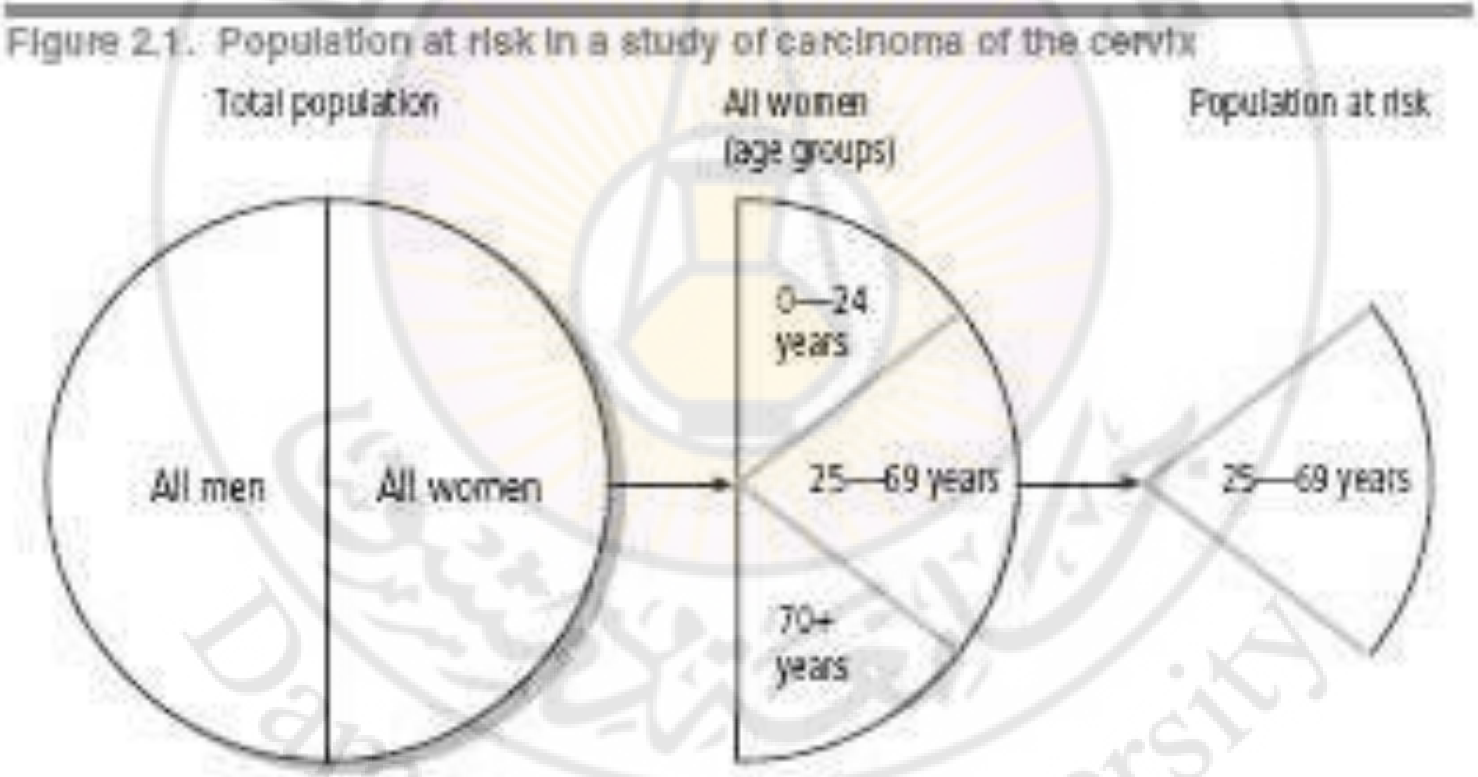
Background

- Different types of activities and practices are undertaken in epidemiology to develop disease control and prevention measures for groups at risk.

- Population at risk
- Incidence and Prevalence
- Case Fatality

- Who is at RISK??
  - Common Cold!!
  - Smoking!!
  - Cardiovascuslar disease!!
  - Cancer of prostate!!
  - Cancer of Cervix!!
  - AIDS
  - Brucellosis

Population AT RISK



Population AT RISK

- Demographic groups (Age, sex, …)
- Occupational groups
- Genetic
- Environmental
- Geographical
- All at risk

Broad Characterizations of Epidemiology Practices

- Descriptive Epidemiology
  - Examining, identifying, and reporting on the frequency and distribution of disease in a population. Learning the basic features of its distribution.
- Analytic Epidemiology
  - Identifying factors underlying disease or health events. Testing a hypothesis by studying how exposures relate to outcomes

Broad Characterizations of Epidemiology Practices

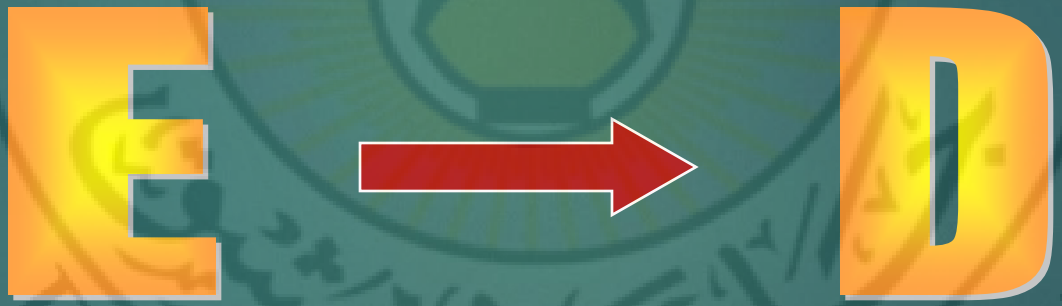
- Developing interventions to reduce disease or improve health in the community
  - Using information from analytical studies, develop strategies centered around an important exposure factor. Test these strategies with clinical trials.
- Program Evaluation
  - Examining the effectiveness of programs for disease control in the community

There is a logical sequence to the practice of epidemiology in disease prevention

Basic Question in Research

Are exposure and disease/outcome linked?

Is there an association between them?



Exposure

Disease / Health Outcome

Health Outcomes

Related to Prognosis or the Evaluation of Health Care Interventions

- Death
- Recovery
- Ongoing Disease
  - Stable disease with treatment
  - Progressive disease
    - disability
  - HRQOL; health related quality of life
- Re-infection, Recurrence

Health Outcomes

- May be intermediate in the clinical course of a disease or treatment
  - Short-term events
- May be the end-result of the clinical course of a disease or treatment
  - Longer-term events

Health outcomes in research studies may be expressed through multiple types of measures

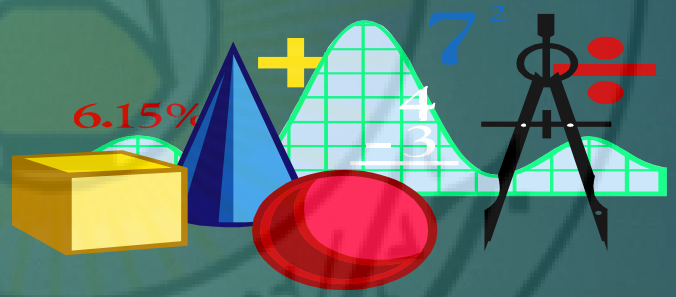
Basic Measurements of Disease or Health Outcome Frequency in Epidemiology

- Measurement of Mortality (death)
- Measurement of Morbidity (incidence, prevalence)

Outcome measures in research studies may be expressed in differing units, though most often as rates

Tools of Measurement

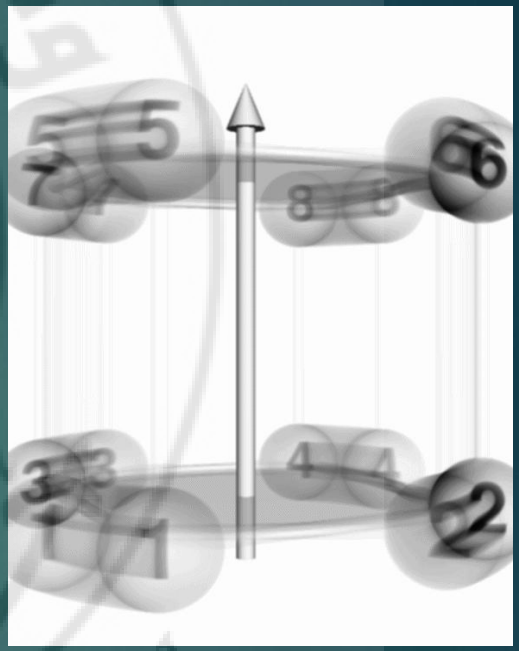
- Counts
- Proportions
- Ratios
- Rates



Case Counts

- Measuring disease or health or health care frequency ***starts*** with counting cases

- Simplest and most frequently gathered measure in epidemiology



Counts

- Refers to the number of cases of a disease or other health phenomenon being studied

  i.e. Number of cases of influenza in Uhan in January 2021

- Can be useful for allocation of health resources
- Limited usefulness for epidemiologic purposes without knowing size of the source population

Counts – Limited Interpretation

Location	New Cases of Disease	Year	Population
City A	20	2008	100
City B	100	2008	1000

Annual Rate of Occurrence

City A: 20 / 100 = 1 / 5

City B: 100 / 1000 = 1 / 10

Proportions

- Persons included in the numerator are <u>always included</u> in the denominator:

$$\text{Proportion:} \quad \frac{A}{A + B}$$

- Indicates the magnitude of a part, related to the total. In epidemiology, tells us the fraction of the population that is affected.

Proportions - Example

A	B	Total (A + B)
# persons with hypertension	# persons without hypertension	Total study population
1,400	9,650	11,050

$$P = A / (A + B) = (1{,}400 / 11{,}050) = 0.127$$

Ratios

- Like a proportion, is a fraction, BUT without a specified relationship between the numerator and denominator

- Example: Occurrence of Major Depression

Female cases = 240 / Male cases = 120 = 240 / 120 = 2:1 female to male

Rates

- A ratio in which TIME forms part of the denominator

- Epidemiologic rates contain the following elements:

  - health issue frequency (in the numerator)
  - unit size of population
  - time period during which an event occurs

- Rate: a measure of the occurrence of a health event in a population group at a specified time period

$$\frac{\text{numerator}}{\text{denominator}} : \frac{\text{Number of events in time period}}{\text{Number at risk for the event in period}}$$

Rates are the basic tool of epidemiologic practice

- Why are rates important?
- because they provide more complete information to describe or assess the impact of a health issue in a community or population

Incidence and Prevalence

Incidence

- represents the rate of occurrence of

new cases

arising in a given period in a specified population

Prevalence

- is the frequency of

existing

cases in a defined population at a given point in time

Incidence

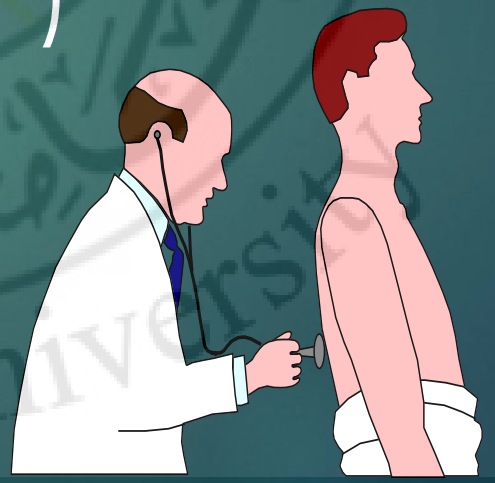
- New Cases/Specified population in a given period

Prevalence

- Existing Cases/defined population in a given
time

Incidence

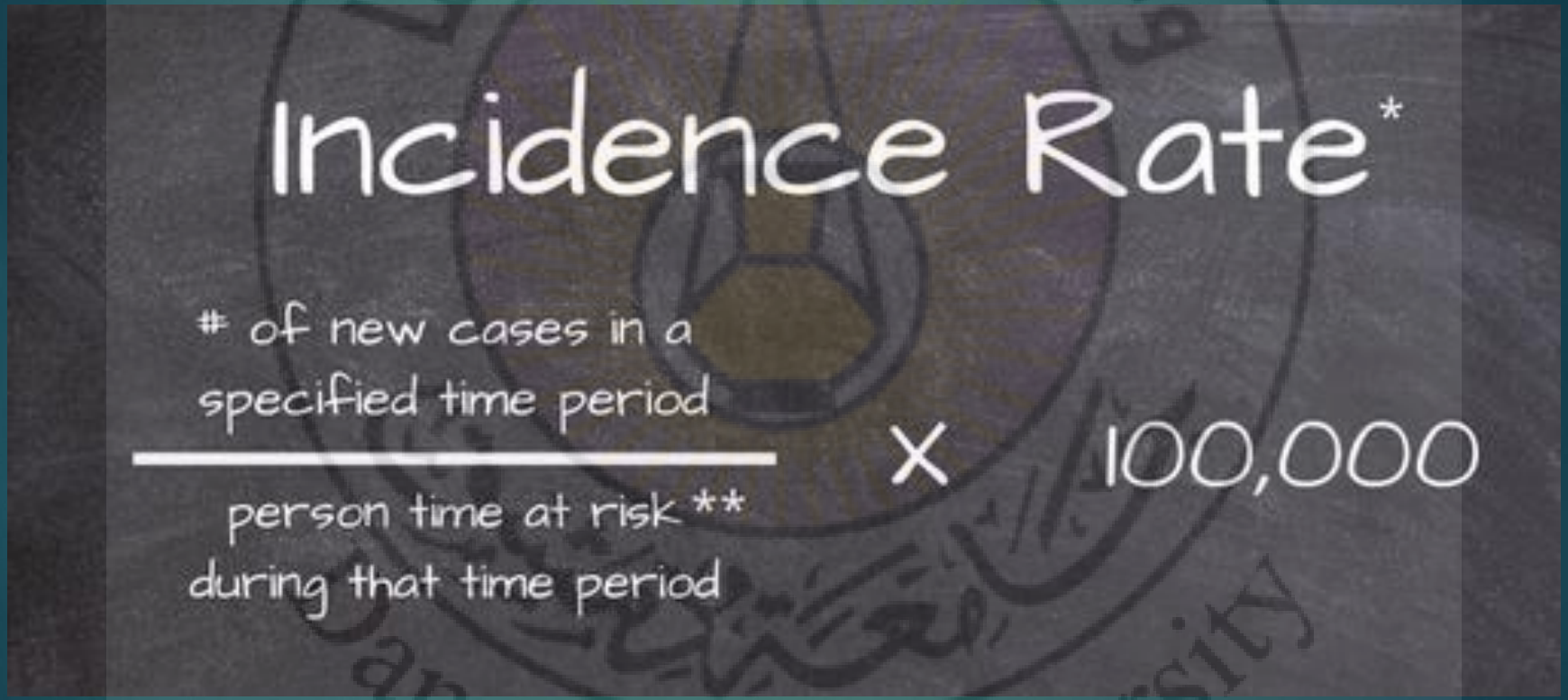
The development of new cases of a disease that occur during a specified period of time in previously disease-free or condition-free ("at risk") individuals.



Incidence

- Incidence rate :
- Cumulative Incidence :

Incidence


Incidence Rate*
# of new cases in a specified time period
person time at risk** during that time period
X
100,000

Cumulative Incidence

Cumulative Incidence (CI) = Incidence

$$CI = \frac{\text{No. of individuals who get the disease during a certain period}}{\text{No. of individuals in the population at the beginning of the period}}$$

If we count all new cases of influenza occurring in MSU undergraduates from September 1, 1997 - August 31, 1998, and we take as the denominator all undergraduates enrolled in September 1, 1997, we would be describing the **cumulative incidence rate** of influenza.

The numerator does not differ between the two types of incidence

However, the denominator can differ in incidence density from cumulative incidence because it takes account of (in the example):

- Students who left school during the year
- Students who died
- Students who had influenza once and will not have it again the same season
- Students who entered school later in the year

COMPARISON OF INCIDENCE DENSITY AND CUMULATIVE INCIDENCE ON A MONTHLY BASIS IN PROBLEM #3, EXERCISE 2.2

(NOTE – ALL INCIDENCES ARE PER MONTH, NOT PER YEAR)

MONTH	POPULATION AT RISK	CUMULATIVE CASES OF CHICKENPOX	INCIDENCE DENSITY		CUMULATIVE MONTHLY INCIDENCE	
			FOR THE MONTH	FROM THE START	FOR THE MONTH	FROM THE START
JANUARY	1000	25	.0253	.0253	.0250	.025
FEBRUARY	975	50	.0260	.0256	.0256	.025
MARCH	950	75	.0267	.0260	.0263	.025
APRIL	925	100	.0274	.0263	.0270	.025
MAY	900	125	.0282	.0267	.0278	.025
JUNE	875	150	.0290	.0270	.0286	.025
JULY	850	175	.0299	.0274	.0294	.025
AUGUST	825	200	.0308	.0278	.0303	.025
SEPTEMBER	800	225	.0317	.0282	.0312	.025
OCTOBER	775	250	.0328	.0286	.0323	.025
NOVEMBER	750	275	.0339	.0290	.0333	.025
DECEMBER	725	300	.0351	.0294	.0345	.025
TOTAL FOR YEAR	**700**	**300**	**.0294**	**.0294**	**.025**	**.025**

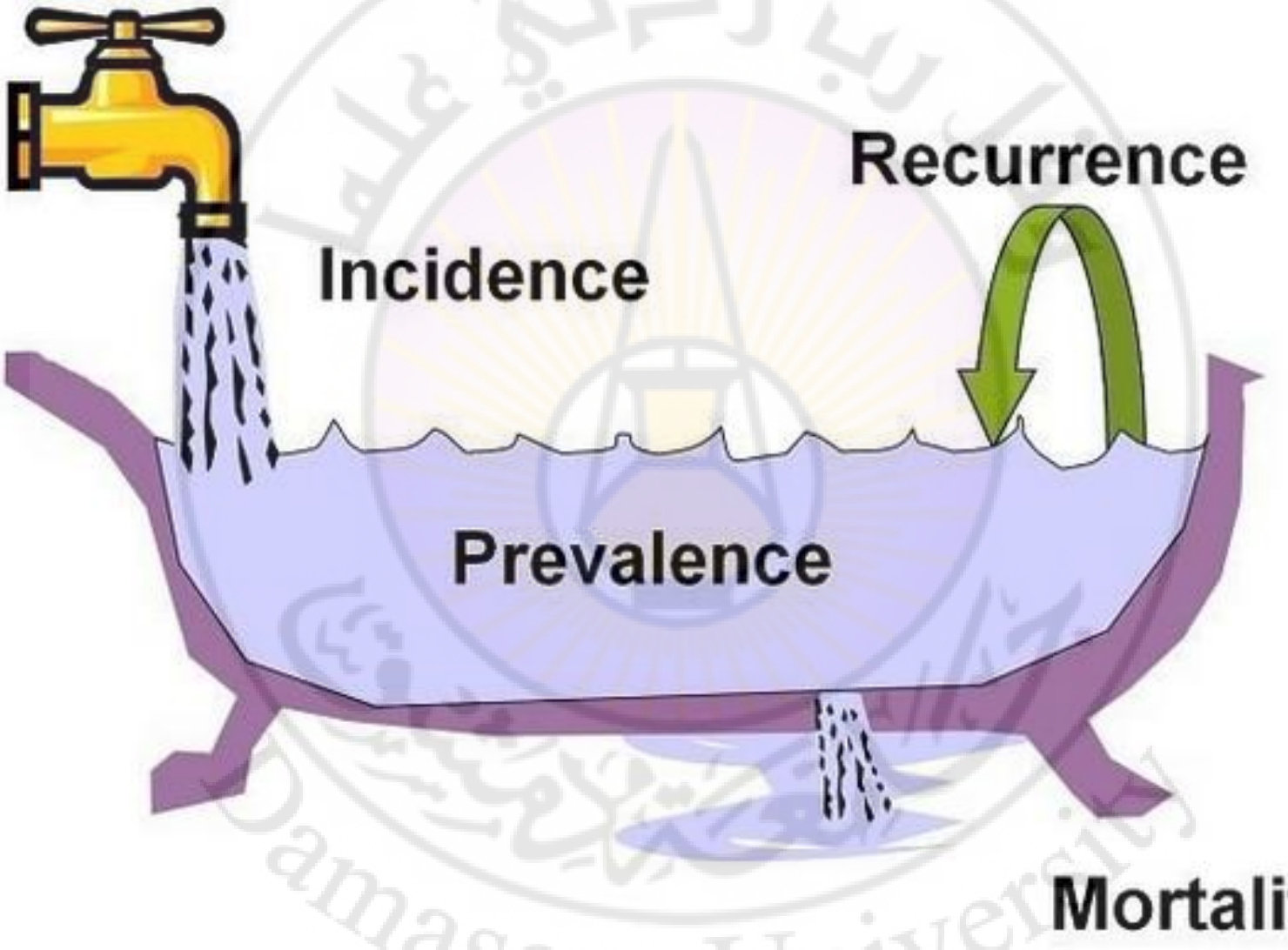
- Investigators enrolled 2,100 women in a study and followed them annually for four years to determine the incidence rate of heart disease. After one year, none had a new diagnosis of heart disease, but 100 had been lost to follow-up. After two years, one had a new diagnosis of heart disease, and another 99 had been lost to follow-up. After three years, another seven had new diagnoses of heart disease, and 793 had been lost to follow-up. After four years, another 8 had new diagnoses with heart disease, and 392 more had been lost to follow-up.

- Prevalence: is another major measure of disease in the population and it quantifies the "burden" of disease

$$\text{Prevalence Rate} = \frac{\text{Number of existing cases of disease in population in time period}}{\text{Persons in population in same time period}}$$

Prevalence

- **Point Prevalence** : At a certain point of time
- **Period Prevalence** : In a given period of time


Recurrence
Incidence
Prevalence
Mortality

Factors influencing prevalence

- Shorter duration of the disease
- High case fatality rate from disease
- Decrease in new cases (incidence)
- In-migration of healthy people
- Out-migration of cases
- Improved cure rate of cases

Factors influencing prevalence

- Longer duration of disease
- Prolongation of life of patients without cure
- Increase in new cases
- In-migration of cases
- Out-migration of healthy people
- In-migration of susceptible people
- Improved diagnostic facilities

Relationship between prevalence + incidence

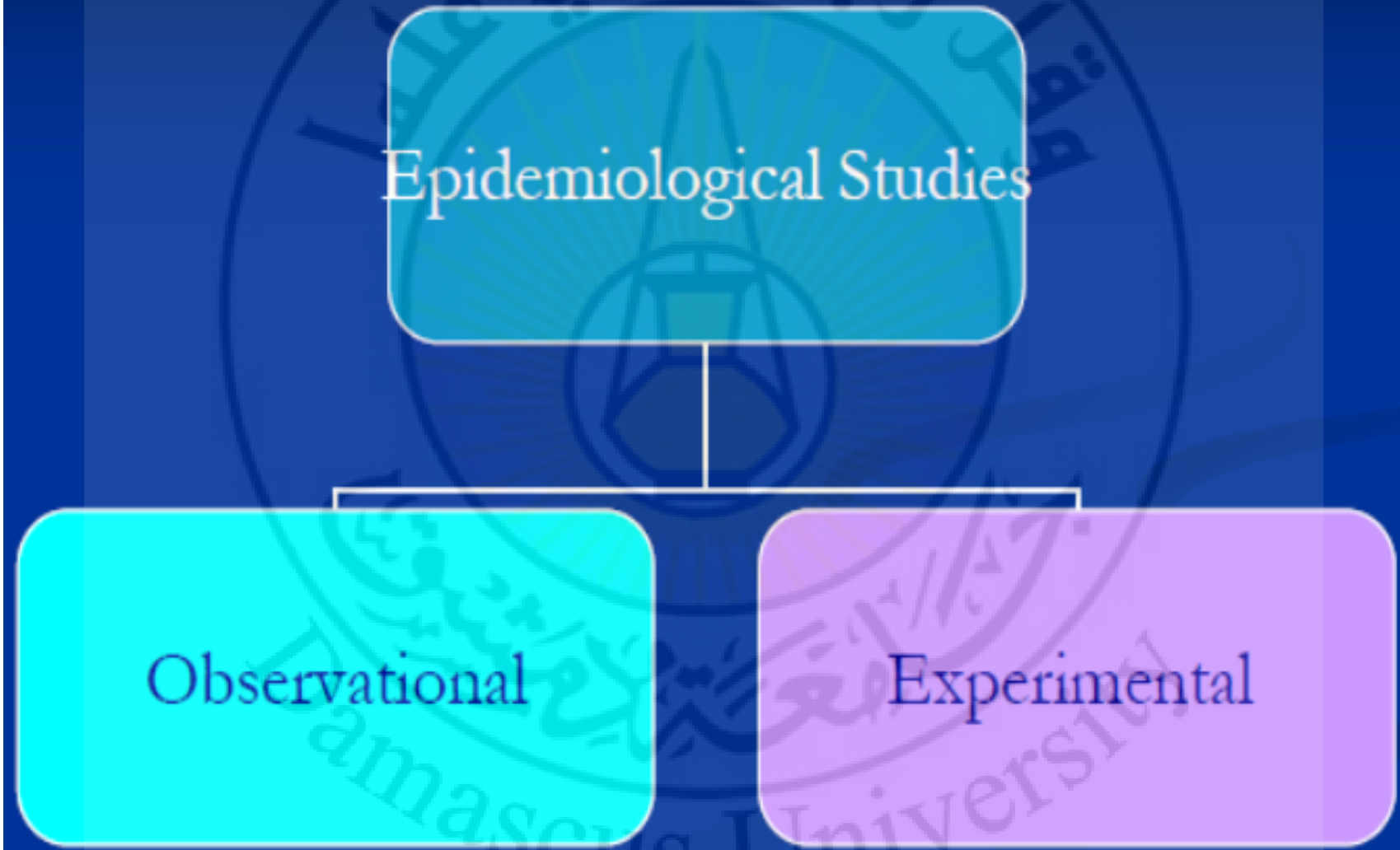
If incident cases are not resolved, but continue over time, then they become existing (prevalent) cases.

Relationship between prevalence & incidence

If incident cases are not resolved, but continue over time, then they become existing (prevalent) cases.

Prevalence = incidence × duration

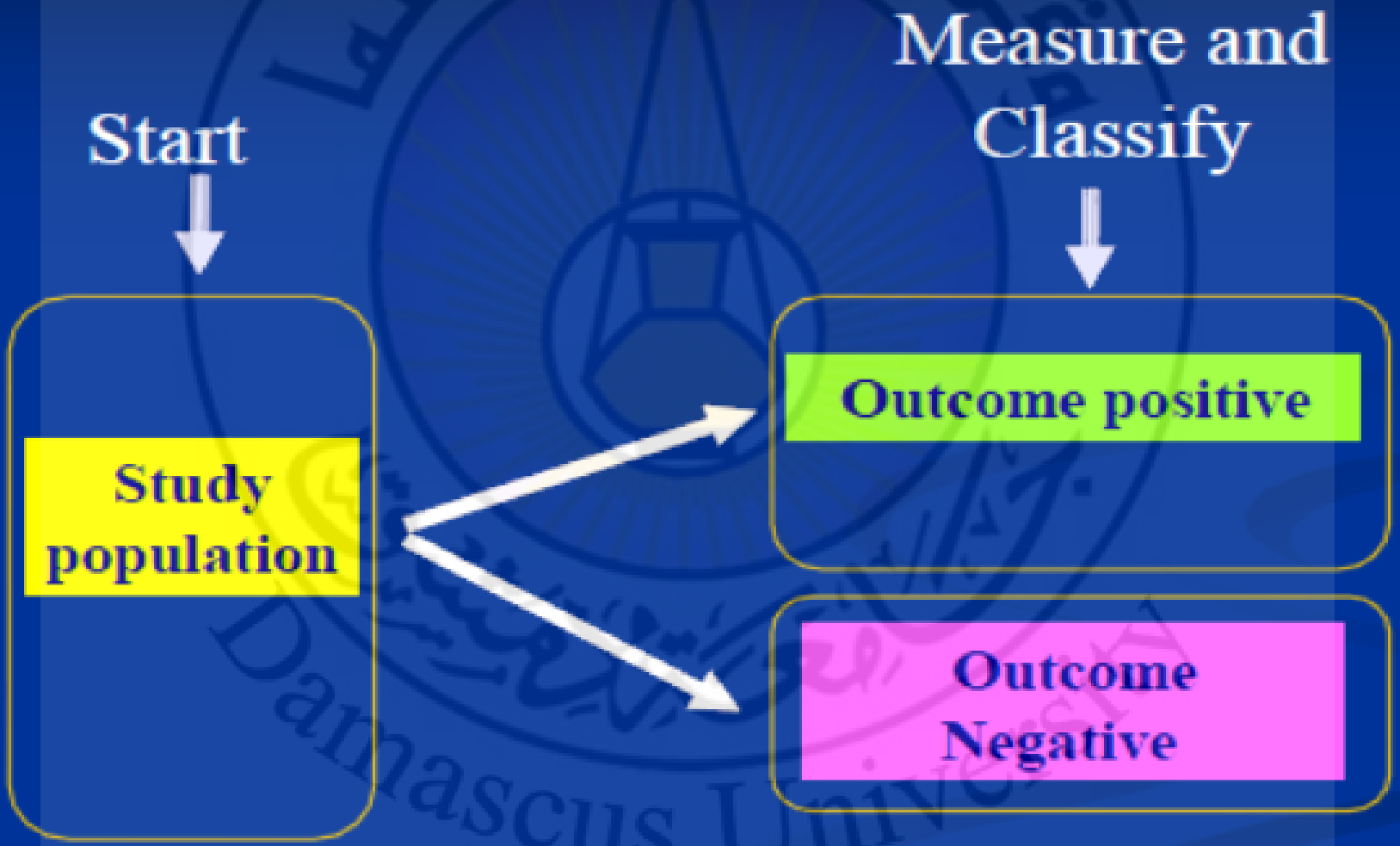
Epidemiological Studies



Cross-sectional studies

- Measure existing disease and current exposure levels at one point in time
- Sample without knowledge of exposure or disease
- Ex. Prevalence studies

Cross Sectional studies



Cross-sectional studies

- Advantages
  - Often early study design in a line of investigation
  - Good for hypothesis generation
  - Relatively easy, quick and inexpensive…*depends on question*
  - Examine multiple exposures or outcomes
  - Estimate prevalence of disease and exposures

Cross-sectional studies

- Disadvantages
  - Cannot infer causality
  - Prevalent vs. incident disease
  - May miss latent disease
  - May be subject to recall bias

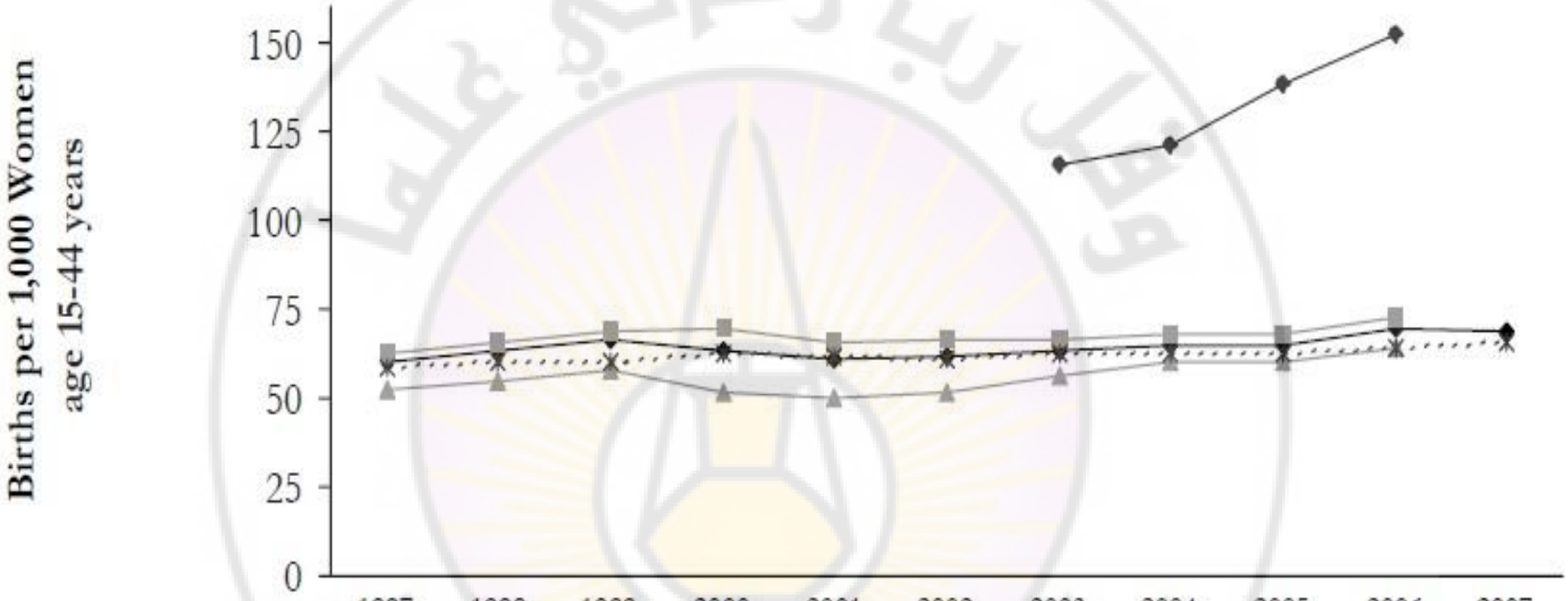
Research Question

- Determine whether there are differences in rates of stroke and myocardial infarction by gender and race among patients.

Hypothesis

- *There will be differences in rates of stroke by gender and race.*
- *There will be differences in rates of myocardial infarction by gender and race.*

General Fertility Rate, Baltimore City by Race and Maryland 1997-2007

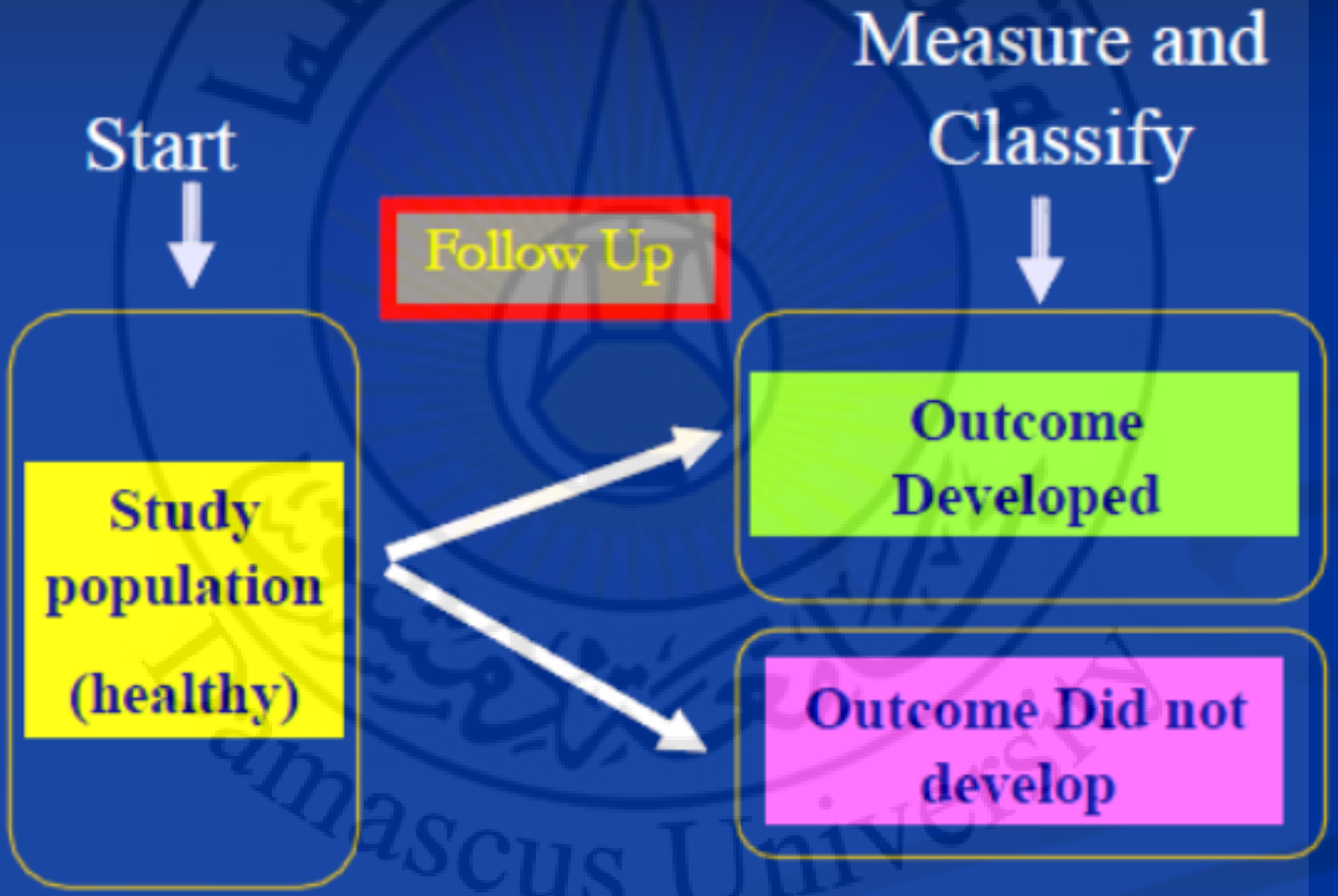


	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007
Baltimore City All Races	60.0	63.0	66.0	63.1	60.7	61.5	63.0	64.6	65.1	69.3	68.5
African American	62.1	65.7	68.4	69.6	65.6	66.3	66.0	67.7	68.2	72.7	
White	52.6	54.9	57.7	51.2	50.3	51.3	55.9	59.9	60.1	64.0	
Hispanic*							115.9	121.0	138.5	152.4	
Maryland All Races	58.8	60.0	60.3	62.4	61.5	60.6	62.3	62.2	62.7	64.1	65.5

Source: Maryland Department of Health and Mental Hygiene, Vital Statistics Annual Report (2007 data are preliminary and not yet available by race/ethnicity)

**Includes all births to mothers of Hispanic origin of any race, data not available prior to 2003*

Longitudinal studies



Longitudinal Study

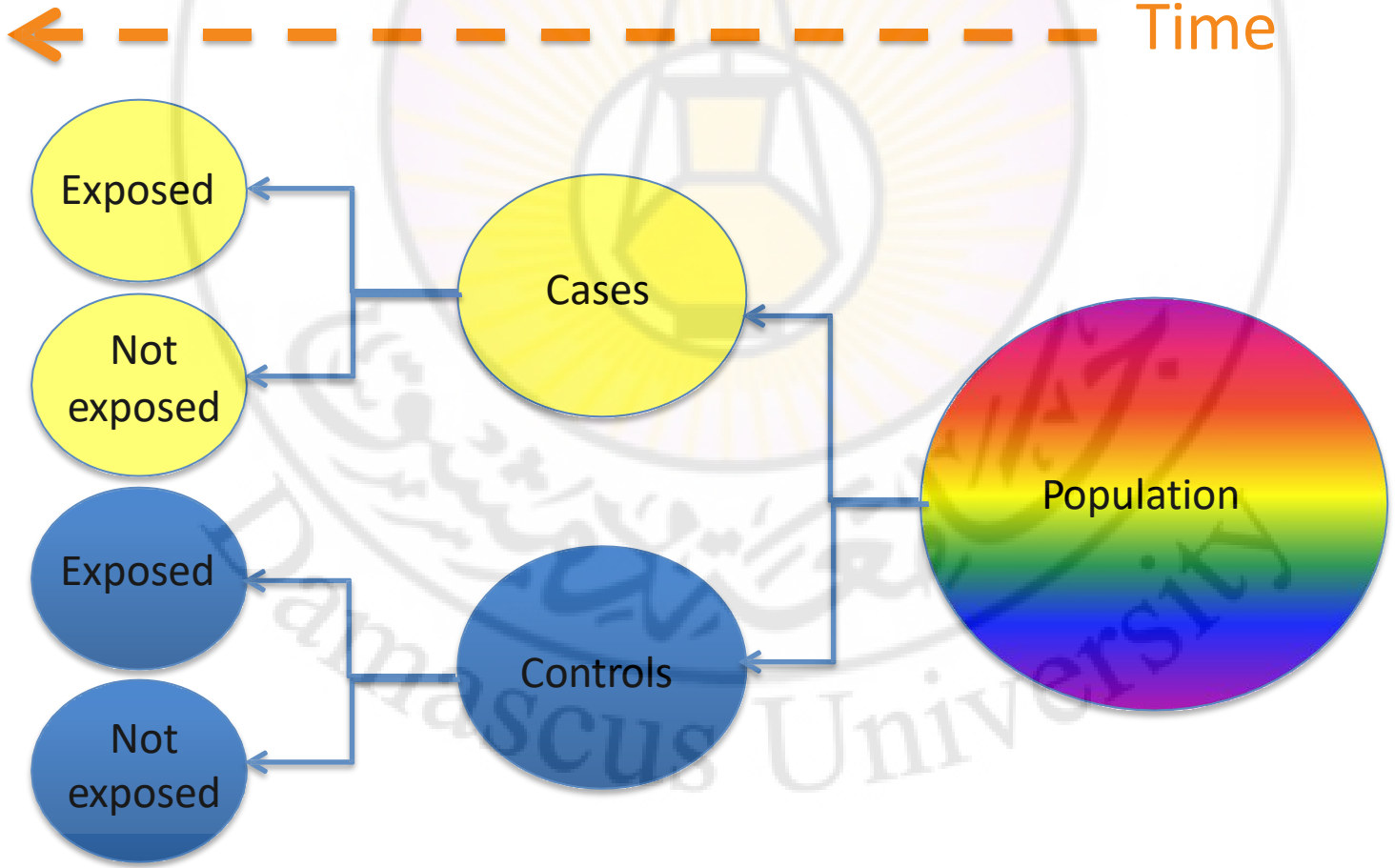
- **Also Known as Incidence Study**

Longitudinal Studies

- Long duration
- Expensive
- Good for changing variable
- Good for frequent diseases
- Follow Up is difficult (Long time!)

Case-Control studies

- Identify individuals with existing disease/s and retrospectively measure exposure



Case-Control studies

- Advantages
  - Good design for rare, chronic and long latency diseases
  - Relatively inexpensive (population size and time)
  - Allows for the examination of multiple exposures
  - Estimate odds ratios
  - Hospital-based studies and outbreaks

Case-Control studies

- Disadvantages
  - Multiple outcomes cannot be studied
  - Recall bias
  - Sampling bias
  - Cannot calculate prevalence, incidence, population relative risk or attributable risk
  - Beware of reverse causation

Neonatal Abstinence Syndrome (NAS) and Drug Exposure

Research question

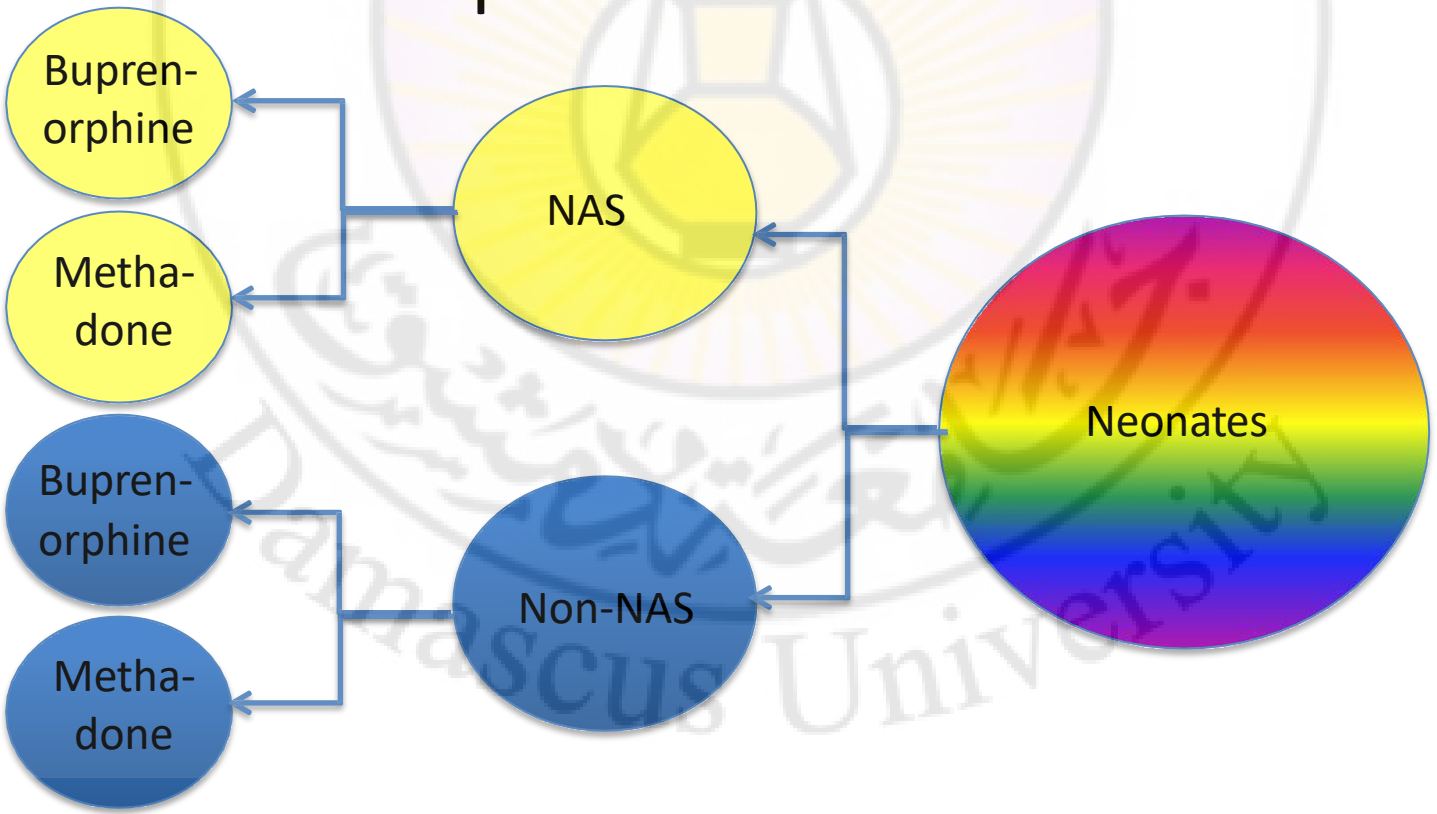
?

Hypothesis 1

Buprenorphine-exposed neonates will exhibit less NAS than methadone-exposed neonates.

Case-Control Study Example

- Hypothesis 1: Buprenorphine-exposed neonates will exhibit less NAS than methadone-exposed neonates.

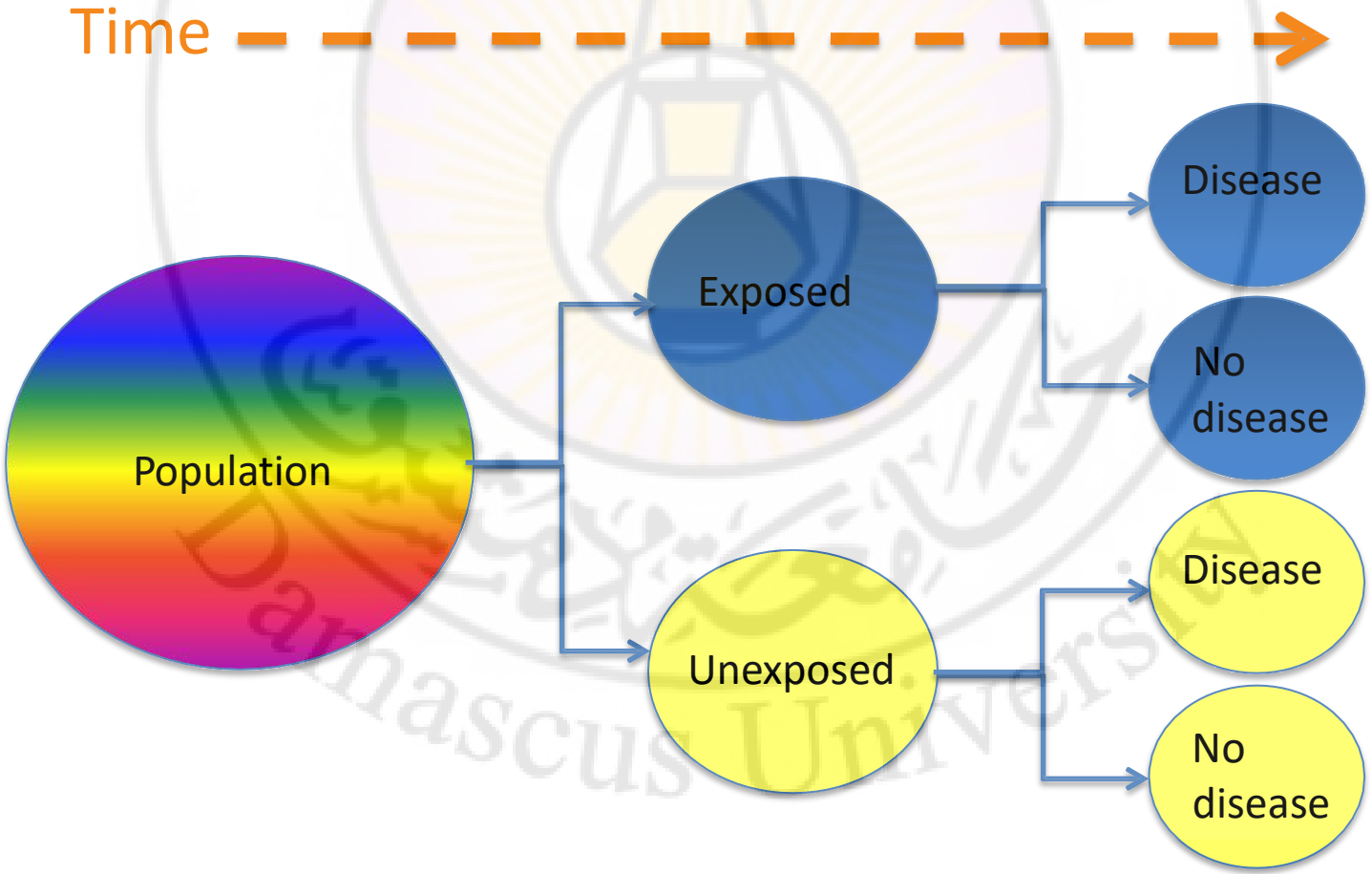


Challenges in Case-Control Studies

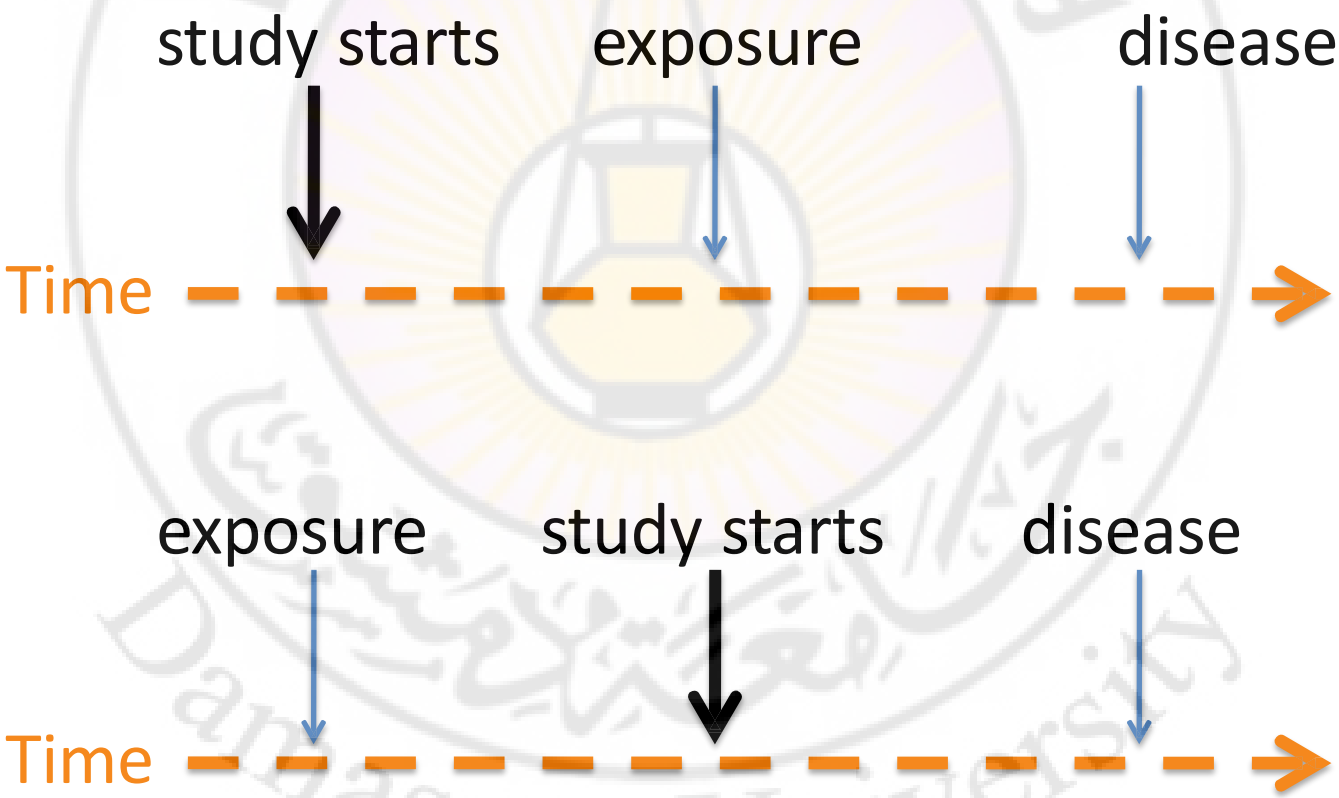
- Selection of Controls
  - Sample size
  - Matching (group or individual)
- Selection of Cases
  - Incident or prevalent disease
- Nested case-control study

Cohort Studies

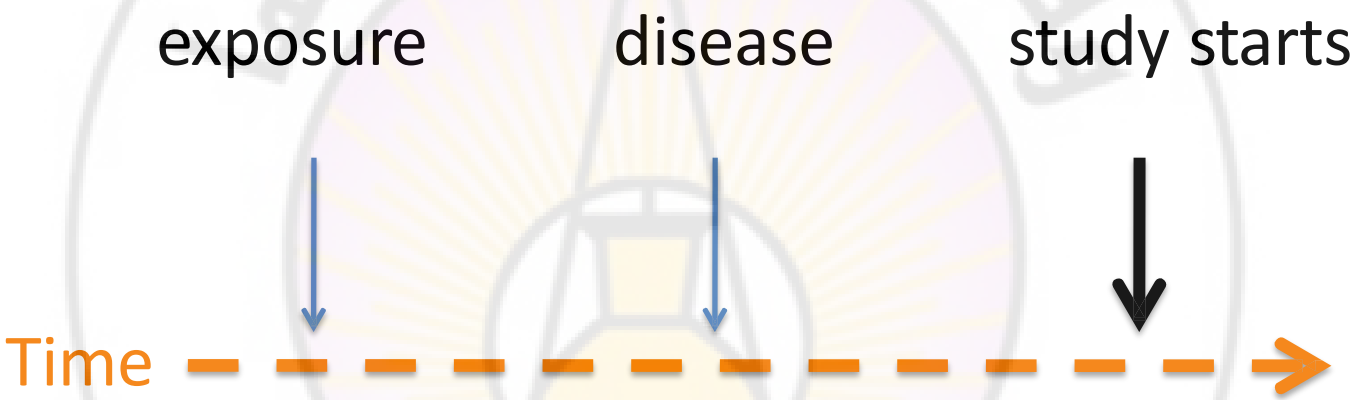
- Identify exposed and unexposed individuals and follow them over time measuring outcome/s (Prospective)



Prospective Cohort Study



Retrospective Cohort Study



Cohort Studies

- Advantages
  - Measure population-based incidence
  - Relative risk and risk ratio estimations
  - Rare exposures
  - Temporality
  - Less likely to be subject to biases (recall and selection as compared to Case-control)
  - Possible to assess multiple exposures and/or outcomes

Cohort Studies

- Disadvantages
  - Impractical for rare diseases and diseases with a long latency
  - Expensive
    - Often large study populations
    - Time of follow-up
  - Biases
    - Design - sampling, ascertainment and observer
    - Study population – non-response, migration and loss-to-follow-up

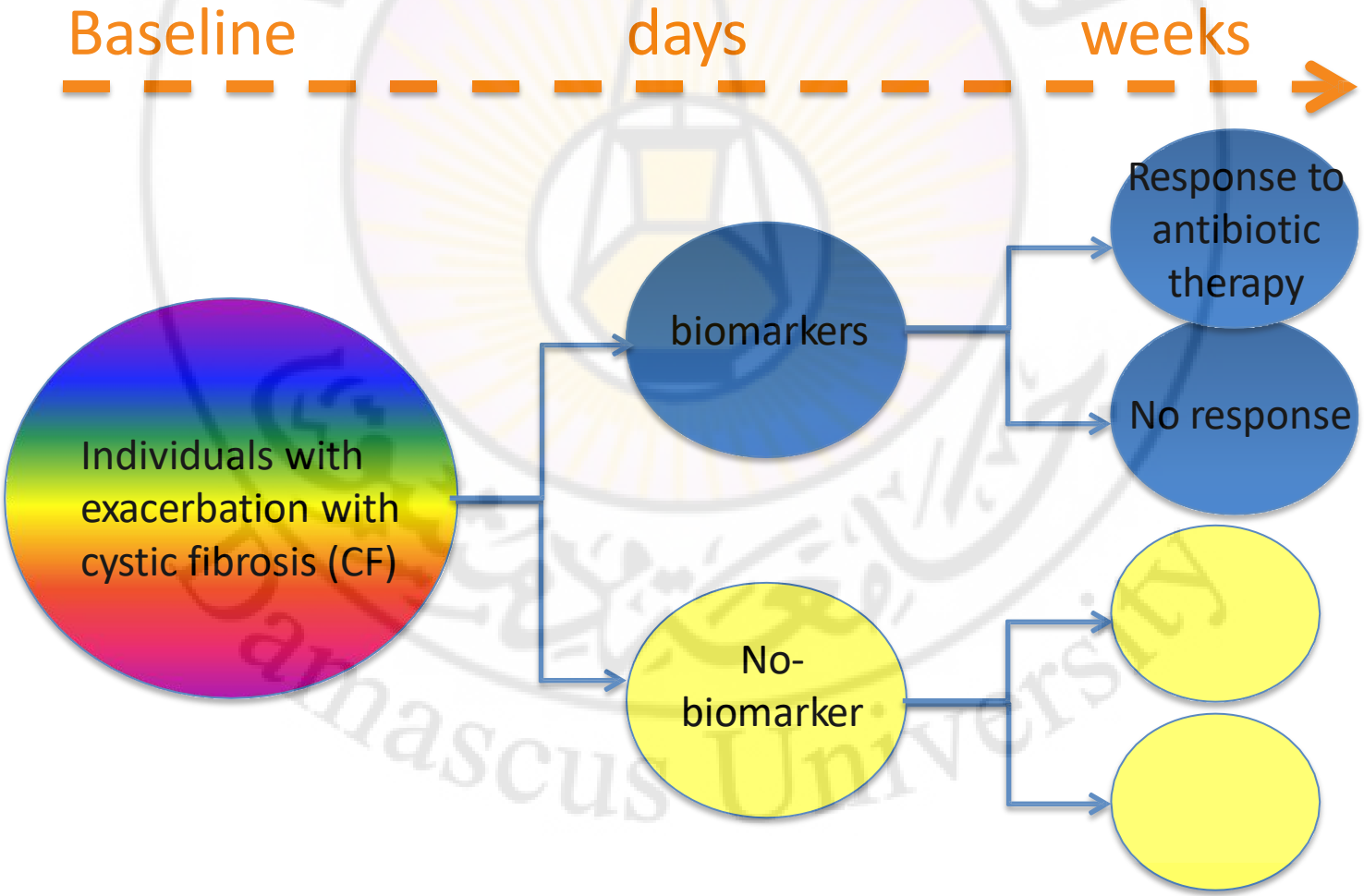
Research Question

Determine whether circulating biomarkers (i.e. C-reactive protein; exhaled breath condensate - pH, hydrogen peroxide, 8-isoprostene, nitrite, nitrate levels; sputum - TNF-α, IL-6, IL-8, IL-1β, neutrophil elastase; and fractional exhaled nitric oxide) predict individuals who will benefit from initiation of antibiotic therapy for the treatment of a mild decrease in FEV_1.

Hypothesis

Biomarkers at the time of presentation with a mild increase in pulmonary symptoms or small decline in FEV_1 can be used to identify which patients require antibiotics to recover.

Cohort Study

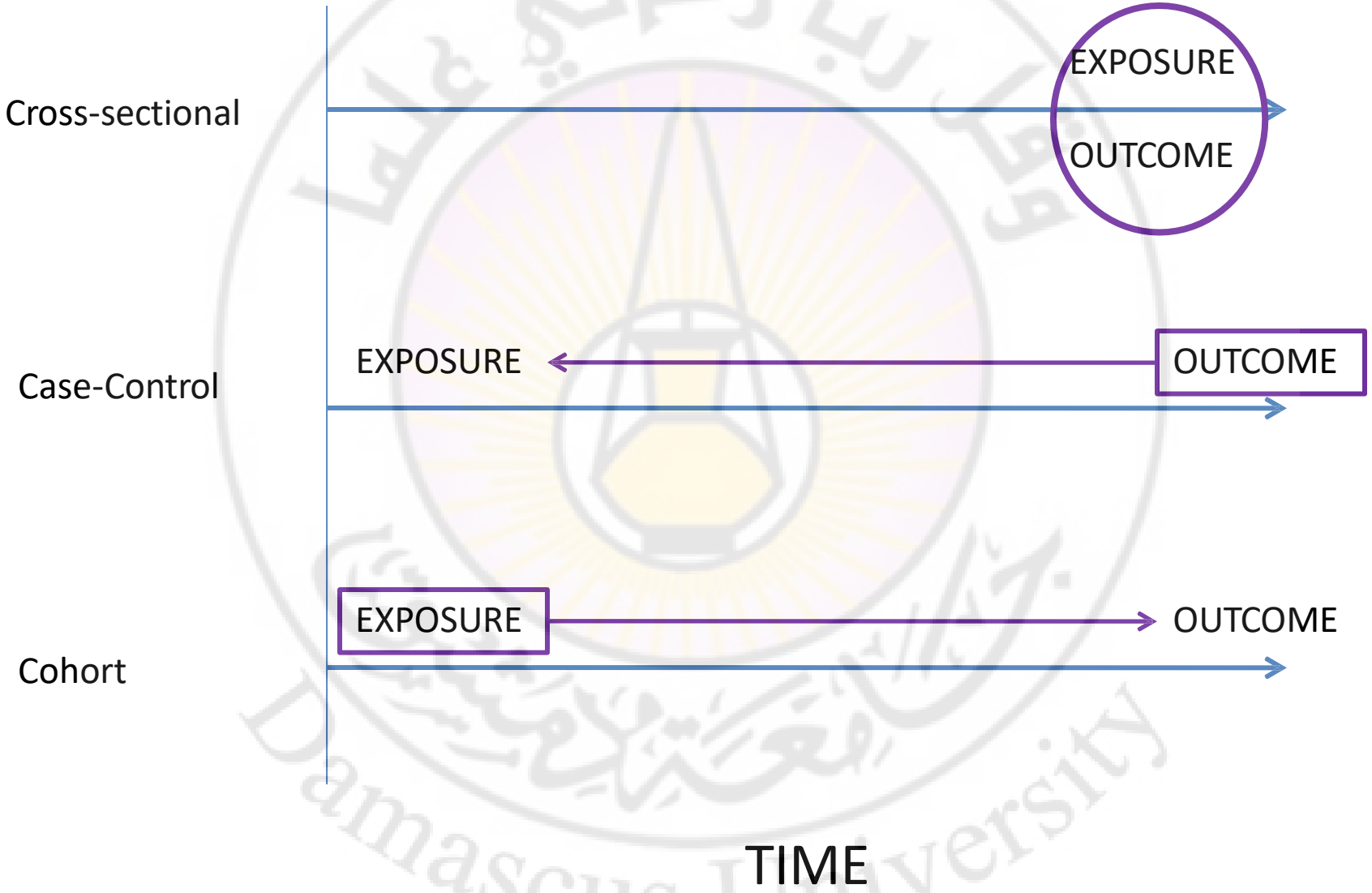


Important features

- How much selection bias was present?
  - Were only people at risk of the outcome included?
  - Was the exposure clear, specific and measureable?
  - Were the exposed and unexposed similar in all important respects except for the exposure?
- Were steps taken to minimize information bias?
  - Was the outcome clear, specific and measureable?
  - Was the outcome identified in the same way for both groups?
  - Was the determination of the outcome made by an observer blinded to treatment?

Important features

- How complete were the follow-up of both groups?
  - What efforts were made to limit loss to follow-up?
  - Was loss to follow-up similar in both groups?
- Were potential confounding factors sought and controlled for in the study design or analysis?
  - Did the investigators anticipate and gather information on potential confounding factors?
  - What methods were used to assess and control for confounding?


Cross-sectional
EXPOSURE
OUTCOME
Case-Control
EXPOSURE
OUTCOME
Cohort
EXPOSURE
OUTCOME
TIME

Applications Of Observational Studies

	Ecological	Cross-Sectional	Case-Control	Cohort
Investigation of rare disease	++++	-	+++++	-
Investigation of rare cause	++	-	-	+++++
Testing multiple effects of cause	+	++	-	+++++
Study of multiple exposures and determinants	++	++	++++	+++
Measurements of time relationship	++	-	+	+++++
Direct measurement of incidence	-	-	+	+++++
Investigation of long latent periods	-	-	+++	-

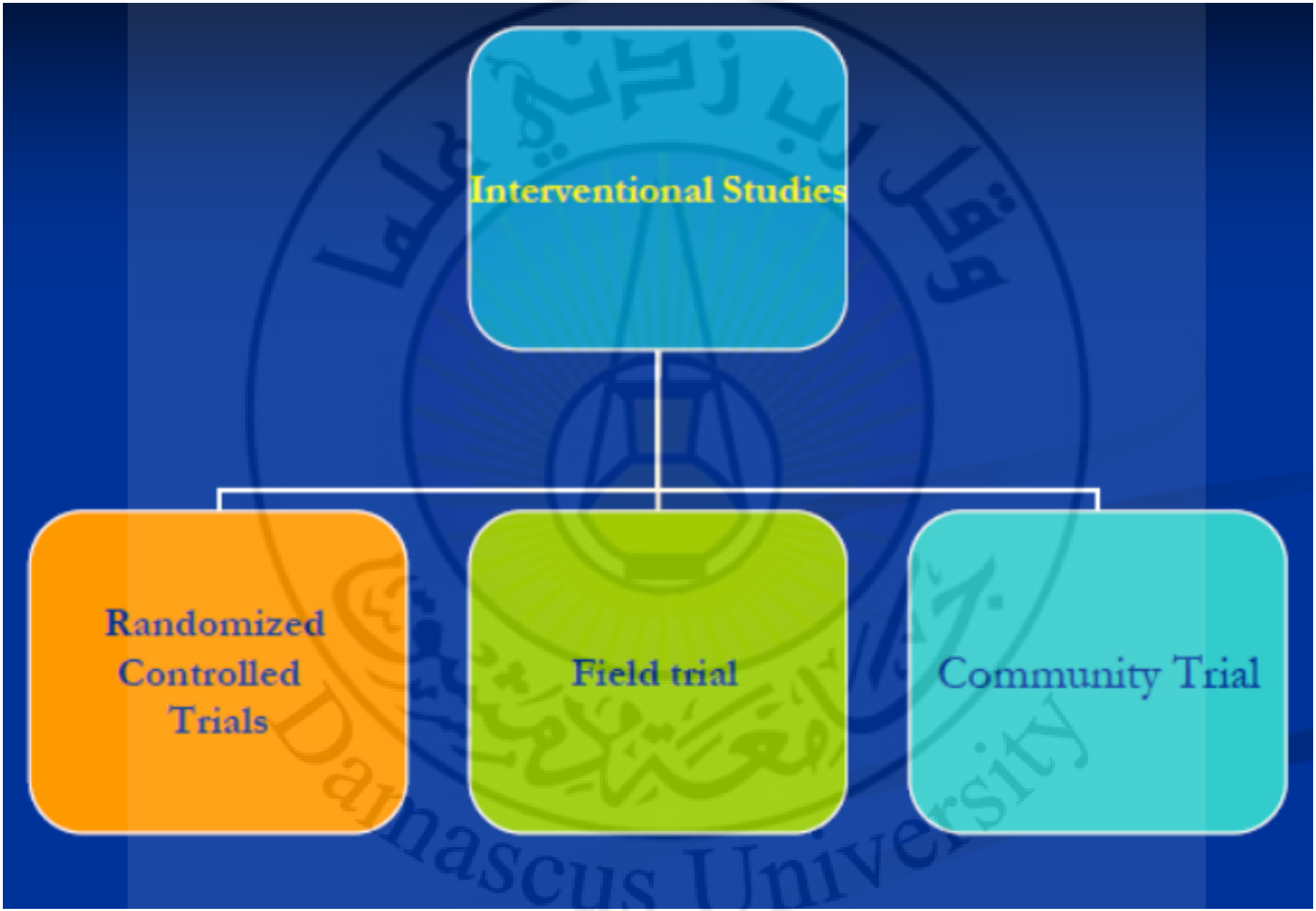
Experimental studies (1)

- The effects of an intervention are measured by comparing the outcome in the experimental group with that in a control group

Experimental studies (2)

- involves attempting to change a variable in one or more groups of people.
  - This could mean the elimination of a dietary factor thought to cause allergy,
  - or testing a new treatment on a selected group of patients.

INTERVENTION


Interventional Studies
Randomized Controlled Trials
Field trial
Community Trial

Experimental Studies

	RCT	Field Trial	Community Trial
Objective	Efficacy of a drug	Efficacy of a vaccine	Efficacy of intervention
Unit of study	Patient	Healthy Individual	Community
Setting	Hospital	Community/Field	Community
Sample Size	100-200	1000-10000	Few
Limitations	++	+	+++

Randomized Controlled Trials (RCTs)

- Experimental: exposure is assigned
- Randomization assignment
  - Random allocation of exposure or treatment
  - Results (or should result!) in two equivalent groups on all measured and unmeasured confounders
- Gold Standard for causal inference

Randomized Controlled Trials

- Advantages
  - Least subject to biases of all study designs (**IF** designed and implemented well…!)

Randomized Controlled Trials

- Disadvantages
  - Intent-to-treat
  - Loss-to-follow-up
  - Randomization issues
  - Not all exposures can be “treatments”, i.e. are assignable
  - Note: for reporting of RCTs see
    Altman DG, et al. CONSORT GROUP (Consolidated Standards of Reporting Trials). Ann Intern Med. 2001 Apr 17;134(8):663-94.

Research Question

- *To determine whether resident's attitudes and skills in diabetes management and counseling change after a curricular intervention.*
- *To determine whether patient outcomes related to diabetes (i.e. weight, smoking status) change after a curricular intervention among residents.*

Hypothesis

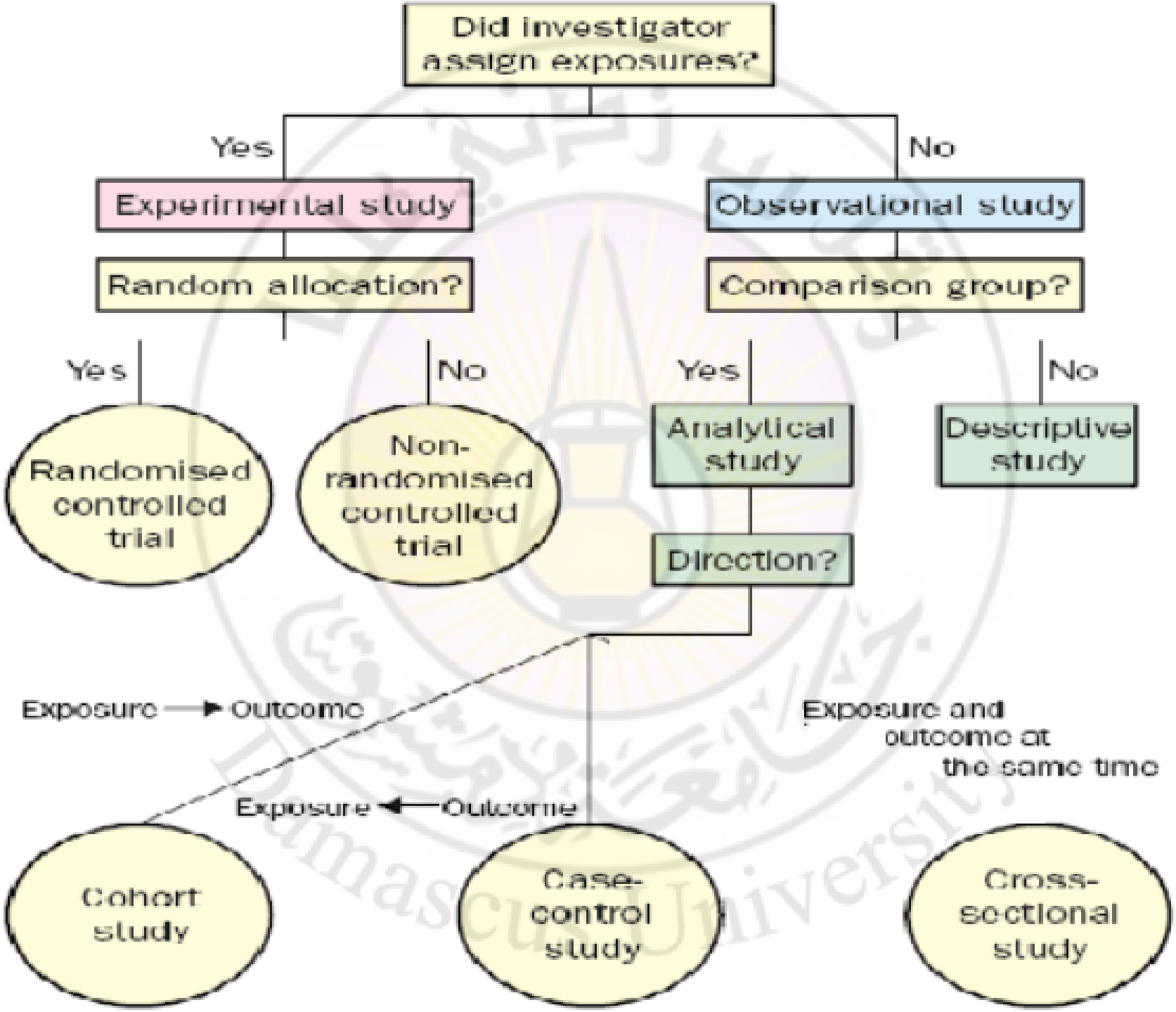
- *Attitudes and skills related to diabetes management and counseling will improve among residents after a curricular intervention.*
- *Fewer patients with diabetes will smoke over time after a curricular intervention among residents.*

Randomization Strategies

- Randomly assigned
- Quasi-randomization
- Block randomization – method of randomization that ensures that at any point in the trial, roughly equal numbers of participants have been allocated to the comparison groups

Ethical Consideration

- ❑ Ethical Approval
- ❑ Consent Form
- ❑ Information Sheet


Did investigator assign exposures?
Yes
No
Experimental study
Observational study
Random allocation?
Comparison group?
Yes
No
Yes
No
Randomised controlled trial
Non-randomised controlled trial
Analytical study
Descriptive study
Direction?
Exposure → Outcome
Exposure ← Outcome
Exposure and outcome at the same time
Cohort study
Case-control study
Cross-sectional study

Study Design

- **Must be defensible**
- **Drives conclusions:**
  What do you want to be able to say at the end of the study?