Principles of Epidemiology

Samuel N. Forjuoh, MB, ChB, DrPH

Associate Professor

Department of Family & Community Medicine

Texas A&M University System Health Science Center

College of Medicine, Scott & White Hospital

Temple, Texas

Outline of Presentation

- Epidemiology: definition, objectives & elements
- Epidemiological measurements
- Screening procedures
- Epidemiological design strategies

Epidemiology:
Definition & Objectives

Definition of Epidemiology

- ***Classic***: Study of the *distribution* & *determinants* of DISEASE frequency in human populations
- ***Last's***: Study of the *distribution* & *determinants* of health-related states, OUTCOMES, or events in specified populations, and the application of this study to control the health problems

Objectives of Epidemiology

- Study etiology & risk factors
- Determine the extent of outcome
- Study natural history & prognosis
- Evaluate new preventive / therapeutic measures / new modes of H / C delivery
- Provide foundations for developing public policy & regulatory decisions

Elements of Epidemiology

- ***Distribution*** of problem
- ***Frequency*** of problem
- ***Determinants*** of problem
- ***Control measures*** for problem

Distribution of Problem

- ***Who*** is affected by problem?
- ***Where*** does problem occur?
- ***What*** is the nature of the problem?
- ***Why*** does problem occur?
- ***When*** does problem occur?
- ***How*** can we prevent the problem?

Frequency of Problem

- Measures of Morbidity
  - Risk => Pr (having the problem)
  - Prevalence => # of existing cases
  - Incidence => # of new cases
- Measures of Mortality
  - Mortality
  - Survival
  - Case Fatality

Determinants of Problem

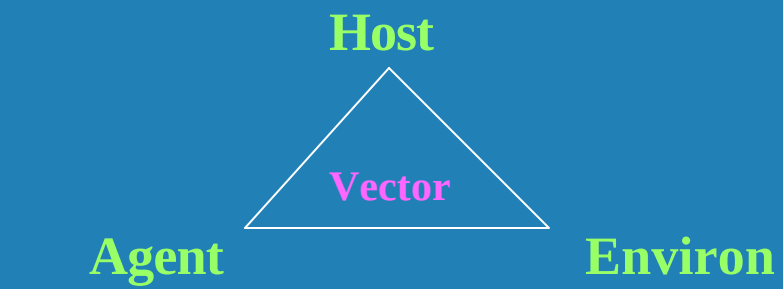
- ***Risk factors*** of problem
- ***Contributory factors*** of problem
- ***Course*** of problem
- ***Prognosis*** of problem

Control of the Problem

- ***Prevention***
  - Primary prevention => pre-event phase
  - Secondary prevention => event phase
  - Tertiary prevention => post-event phase
- ***Promotion (1^0 prevention)***
- ***Acute care (2^0 prevention)***
- ***Rehabilitation (3^0 prevention)***

Dynamics of Transmission

- The epidemiologic triad
  - Host
  - Agent/Vector
  - Environment



- Modes of transmission
  - Direct *vs.* indirect transmission
  - Horizontal *vs.* vertical transmission
- Endemic, Epidemic & Pandemic

Risk & Protective Factors

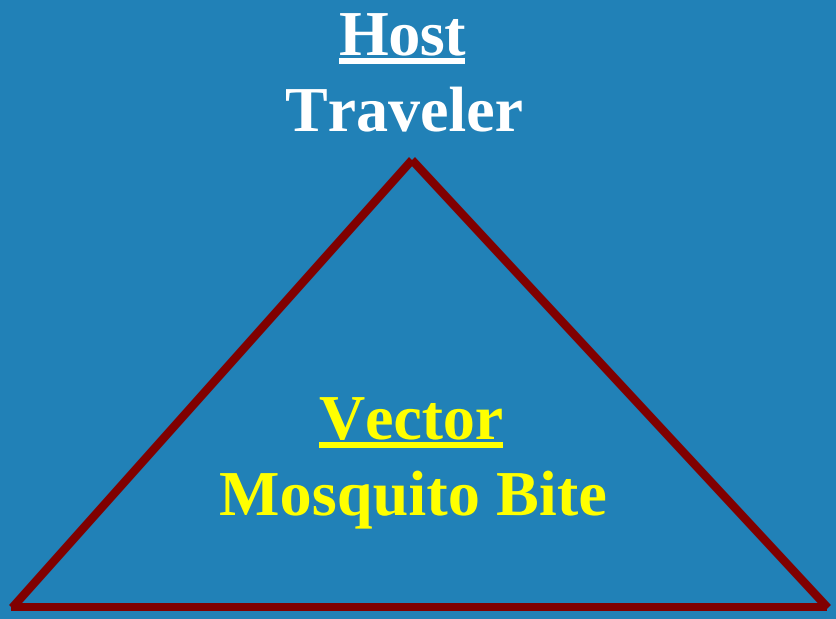
HOST	AGENT	ENVIRONMENT
Age, sex, race	Biologic (bacteria)	Temperature
Religion	Chemical (poison)	Overcrowding
Occupation	Physical (trauma)	Neighborhood
Education	Nutritional (lack)	Housing
Marital status	Energy (thermal)	Radiation

Modes of Transmission

- Direct transmission
  - person-to-person *via* contact, e.g., cold, TB
- Indirect transmission
  - *via* vector e.g., water-borne diseases
- Horizontal transmission => infectious d
  - from one person to another
- Vertical transmission => genetic d
  - from one generation to another

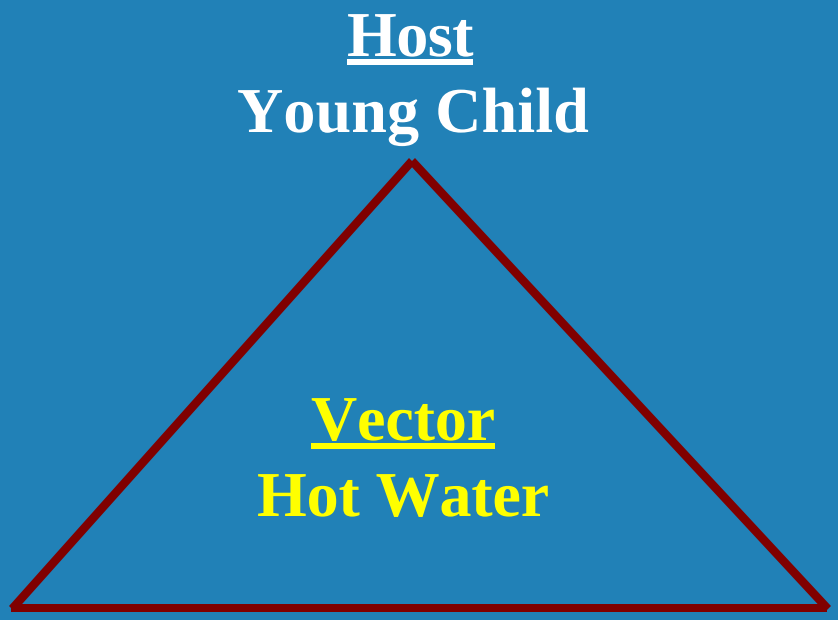
Application to Disease

MALARIA



Application to Injury/Trauma

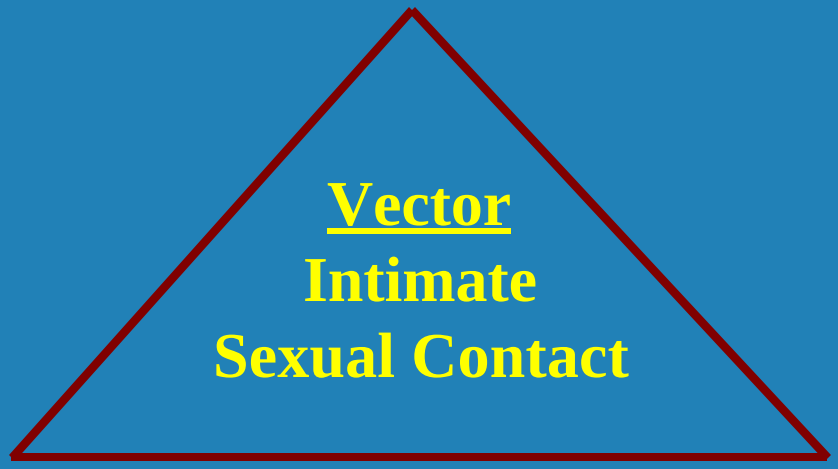
BURN



Application to Social Problem

TEEN PREGNANCY

Host
Teenager



Agent
Sperm

Environ
Lack of Protection

Endemic, Epidemic & Pandemic

- Endemic condition
  - Usual occurrence of a condition within a geographic area
- Epidemic condition
  - Occurrence of a condition clearly in excess of normal expectancy
- Pandemic condition
  - An epidemic occurring over a very wide area usually affecting a large population

Investigation of an Epidemic

- Define the epidemic *re: elements of Epid.*
- Examine distribution of cases
- Look for interactions of relevant variables
- Develop hypothesis based on:
  - existing knowledge, if any, of the condition
  - analogy to conditions of known etiology
- Test hypotheses
- Recommend control & prev^{n} measures

Epidemiological Measurements

Risk $\mathbf{R = \frac{\#\ New\ Cases}{Persons\ At\ Risk}}$

- Measure of the occurrence of new cases in a pop^{n} as a f^{n} of *persons at risk* at a specified time
- *a.k.a.* Cumulative incidence
- Proportion; no units
- 1 > R > 0 *or* can be expressed as a %
- Reduced by 1^{0} prevention

Prevalence

$$P = \frac{\text{\# Existing Cases}}{\text{Total Population}}$$

- Measure of existing cases in a pop^{n} as a f^{n} of *total pop*^{n} at a specified time
- *a.k.a.* Point prevalence
- Proportion; no units
- $1 > P > 0$ *or* can be expressed as a %
- Reduced by 2^{0} prevention

Types of Prevalence

- Point Prevalence
  - A snap shot of time
    - Are you here because of abuse?
- Period Prevalence
  - Over a specified period of time
    - Have you been abused in the past 6 months?
- Life-time Prevalence
  - Over a life time
    - Have you ever been abused?

Incidence

$$\mathbf{IR} = \frac{\text{\# New Cases}}{\text{Person-Time Obs}}$$

- Measure of new cases in a pop^{n} as a f^{n} of *person-time of observation*
- *a.k.a.* Incidence density
- Rate; units of person-time
- **4** > IR > 0
- Reduced by 1^{0} prevention

Computation Example - I

- Consider a town with a pop^{n} of 5,000. If 500 of 4,000 persons without disease A develop it over a 10-year period:
  - What is the risk of acquiring A? $R = \frac{500}{4{,}000}$
  - What is the prevalence of A? $P = \frac{1{,}500}{5{,}000}$
  - What is the incidence of A? $IR = \frac{500}{40{,}000}$

Indices of Mortality

- Crude mortality rate
- Proportionate mortality rate
  - attributable to a specified condition
- Specific mortality rate
  - attributable to a specified segment of pop^{n}
- Standardized mortality ratio
- Case fatality ratio (rate)
- Survival rate

Standardization of Rates

- Why standardize?
  - To account for the possible effects of differences in age distribution *or* other confounding variables
- Direct standardization
  - Uses a suitable standard pop^{n} (combined)
- Indirect standardization
  - Uses a set of standard age-specific rate

Screening & Diagnostic Tests

Screening

- Application of a test *or* procedure to asymptomatic persons for the purpose of classifying them *w.r.t.* their likelihood of having a particular condition
- Aims at early detection before the development of symptoms
- Ex: Partner Violence Screen - 3 Qns

Attributes of a Screening Test

- Validity
  - Ability of the test to do what it is supposed to do
  - Accuracy of the test
- Reliability
  - Consistency of results when repeat tests are performed on the same persons under the same condition
  - Precision *or* repeatability of the test

Measures of Validity

- Sensitivity
  - Pr. of testing +ve if condition truly exists
- Specificity
  - Pr. of testing -ve if condition is truly absent
- Predictive value positive
  - Pr. of having disease given a +ve test
- Predictive value negative
  - Pr. of not having disease given a -ve test

The 2x2 Contingency Table

- Epidemiologist's way of summarizing data for analysis / test of association
- Two variables represented by columns and rows

		Variable 1 +	Variable 1 -
Variable 2	+	a	b
	-	c	d

Sensitivity

$$\text{Sensitivity} = \Pr(T^+ / D^+) = \frac{\text{True +ves}}{\text{All w/ Cond}} = \frac{TP}{TP+FN}$$

TEST RESULTS	WITH CONDITION	WITHOUT CONDITION
POSITIVE	Have Condition Have +ve Test Result = TRUE +VES (TP) a	No Condition Have +ve Test Result = FALSE +VES (FP) b
NEGATIVE	Have Condition Have -ve Test Result = FALSE -VES (FN) c	No Condition Have -ve Test Result = TRUE -VES (TN) d

Specificity

$$\text{Specificity} = \Pr(T^- / D^-) = \frac{\text{True -ves}}{\text{All w/o Cond}} = \frac{TN}{TN+FP}$$

TEST RESULTS	WITH CONDITION	WITHOUT CONDITION
POSITIVE	Have Condition Have +ve Test Result = TRUE +VES (TP)	No Condition Have +ve Test Result = FALSE +VES (FP)
NEGATIVE	Have Condition Have -ve Test Result = FALSE -VES (FN)	No Condition Have -ve Test Result = TRUE -VES (TN)

Predictive Value Positive

$$PV+ = Pr(D^+ / T^+) = \frac{\text{True +ves}}{\text{All Testing +ve}} = \frac{TP}{TP+FP}$$

TEST RESULTS	WITH CONDITION	WITHOUT CONDITION
POSITIVE	Have Condition Have +ve Test Result = TRUE +VES (TP)	No Condition Have -ve Test Result = FALSE +VES (FP)
NEGATIVE	Have Condition Have -ve Test Result = FALSE -VES (FN)	No Condition Have -ve Test Result = TRUE -VES (TN)

Predictive Value Negative

$$PV- = Pr(D^-/T^-) = \frac{\text{True -ves}}{\text{All Testing -ve}} = \frac{TN}{TN+FN}$$

TEST RESULTS	WITH CONDITION	WITHOUT CONDITION
POSITIVE	Have Condition Have +ve Test Result = TRUE +VES (TP) a	No Condition Have -ve Test Result = FALSE +VES (FP) b
NEGATIVE	Have Condition Have -ve Test Result = FALSE -VES (FN) c	No Condition Have -ve Test Result = TRUE -VES (TN) d

Computation Example - II

	C^+	C^-	
+ve	40	20	60
-ve	15	25	40
	55	45	100

$$\text{Sens} = \frac{40}{55} \times 100\% = 73\%$$

$$\text{Spec} = \frac{25}{45} \times 100\% = 56\%$$

$$PV^+ = \frac{40}{60} \times 100\% = 67\%$$

$$PV^- = \frac{25}{40} \times 100\% = 63\%$$

Measures of Reliability

- Percent Agreement
  - Observed proportion of agreement between 2 observers
- Percent Agreement by Chance
  - Proportion of agreement between 2 observers arising by chance
- Cohen's Kappa Statistic
  - Extent to which agreement between 2 observers is above and beyond the amount of agreement expected by chance alone

Percent Agreement

		Physician 2		
		Abnormal	Normal	
Physician 1	Abnormal	16	2	18 (40%)
	Normal	16	11	27 (60%)
		32 (71%)	13 (29%)	45 (100%)

$$\% \text{ Agreement} = \frac{\text{Total \# Agreed}}{\text{Grand Total}} = \frac{16+11}{45} \times 100\% = 60\%$$

% Agreement by Chance

		Physician 2		
		Abnormal	Normal	
Physician 1	Abnormal	12.8	5.2	18 (40%)
	Normal	19.2	7.8	27 (60%)
		32 (71%)	13 (29%)	45 (100%)

Note: Exp=RTxCT/GT

$$\%\ \text{Agr by Ch} = \frac{\text{Tot. Exp. \# Agr}}{\text{Grand Total}} = \frac{12.8+7.8}{45} \times 100\% = 45.8\%$$

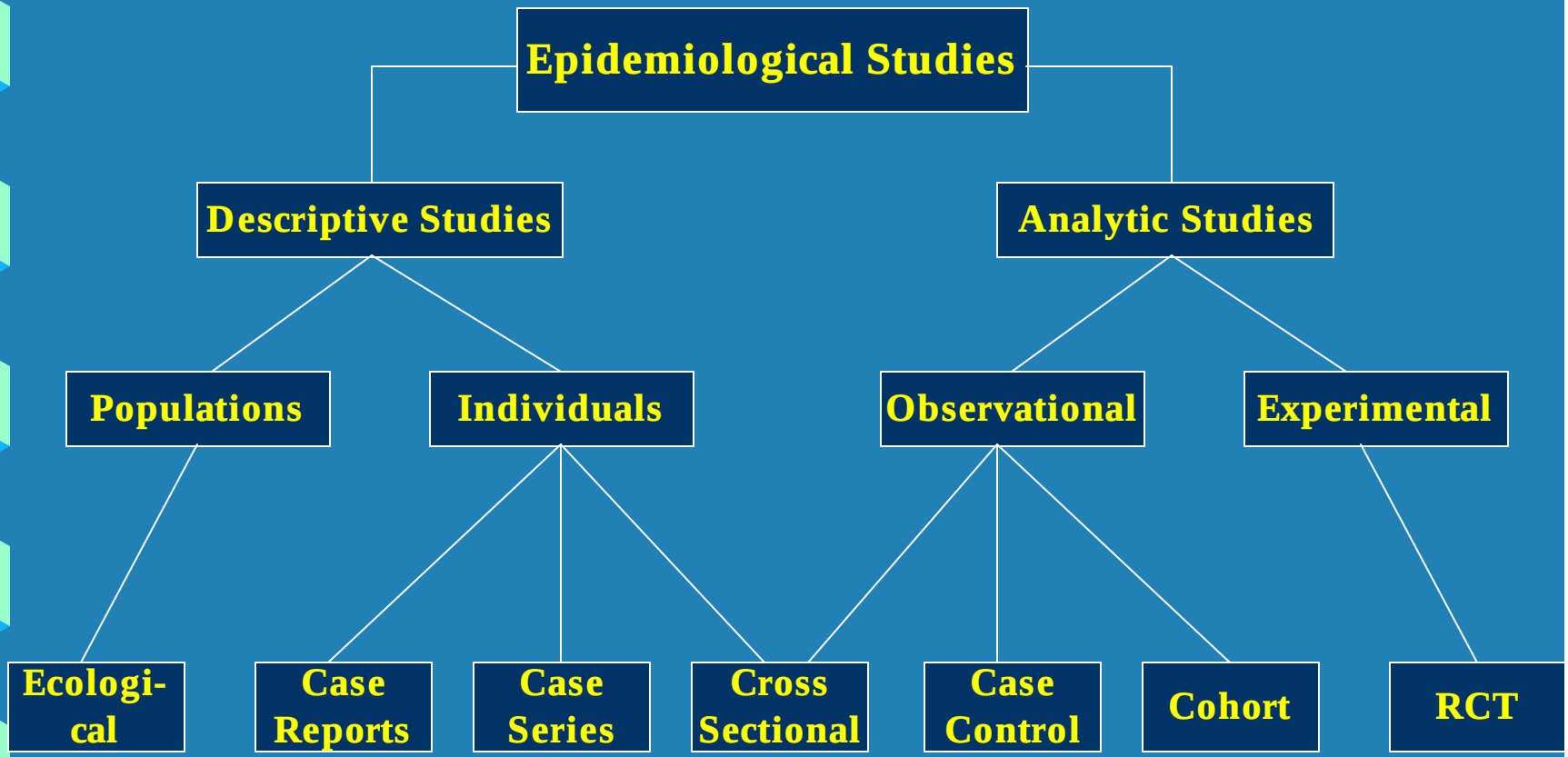
Cohen's Kappa Statistic

$$\text{Kappa} = \frac{\text{\% Agreement - \% Agreement By Chance}}{\text{100\% - \% Agreement By Chance}}$$

$$= \frac{\text{60\% - 45.8\%}}{\text{100\% - 45.8\%}} = \frac{\text{14.2\%}}{\text{54.2\%}} = .26$$

Epidemiological Design Strategies

Epidemiological Design Strategies



Descriptive Studies . . .

- Studies normally concerned with the distribution of diseases or health-related outcomes including consideration of the subgroups affected, and how the frequency of occurrence of the condition varies over time
- No hypothesis testing

A Case Report

- Consists of a careful, detailed report on the profile of a single patient
- Most *basic* type of descriptive study design of individuals
- Ex: Patient, i.v. drug abuser; weight loss, etc.

A Case Series

- Consists of a compilation of case reports describing the characteristics of a number of individual patients with a particular health-related outcome
- Ex: Several patients seen at San Francisco with similar (AIDS) s/s

An Ecological Study

- Study of whole populations or groups of persons rather than individual persons
- Unit of analysis = populations
- Ex: County death rates by per capita cigarette consumption

Analytical Studies . . .

- Studies designed to examine associations which may have putative *or* hypothesized cause-effect relationships
- Involve hypothesis-testing

A Cross-Sectional Study

- Status of individuals are observed and assessed *w.r.t.* presence or absence of both exposures and outcomes at the same time
- *a.k.a.* Prevalence study
- Considered descriptive & analytical
- Most *common* study design used

A Case-Control Study

- Study subjects are selected based on whether or not they have an outcome of interest and then compared *w.r.t.* to their current or past exposure histories
- *a.k.a.* Retrospective study
- Provides strength of association

A Cohort Study

- Study subjects are selected based on whether or not they have an exposure of interest and are then followed over time to assess the occurrence of outcome(s)
- *a.k.a.* Prospective / Follow-up study
- Provides measures of relative risk of incidence

A Randomized Controlled Trial

- A type of prospective cohort study in which the researcher assigns the exposure of interest randomly between the study subjects
- *a.k.a.* Intervention study
- Most *advanced* study design

Criteria for Causal Inference

- Strength of the association
- Dose-response effect
- Lack of temporal ambiguity
- Consistency of findings
- Biological plausibility of hypothesis
- Coherence of the evidence
- Specificity of the association

Descriptive *vs.* Analytic Designs

Descriptive Studies

- **No hypothesis testing**
- No elucidation on the criteria for causation
- **Lack power to generalize results**

e.g. Case Reports

Analytic Studies

- **Can hypothesize on cause-effect relationships**
- Elucidation on the criteria for causation
- **Results may be normally generalizable**

e.g. Case-Control Study

Hierarchy of Epi Design Strategies

Complexity

Confidence

Case Reports

Case Series

Cross-Sectional Surveys

Case-Control Studies

Cohort (Follow-Up) Studies

Randomized Controlled Trials