**Chinsamba Abel Waison**

**A9UM913339**

COURSE NAME:

**(Introduction to Epidemiology)**

ATLANTIC INTERNATIONAL UNIVERSITY

**April/2025**

**Introduction 3**

**Question 1 3**

1.1 Table 1.1 indicates that there were over 40 times more cholera cases in one district than in another. Did this reflect the risk of catching cholera in each district? 3

1.2 How could the role of the water supply in causing deaths from cholera have been tested further? 4

1.3 Why do you suppose the study shown in Figure 1.2 was restricted to doctors? 4

1.4 What conclusions can be drawn from Figure 1.2? 5

1.5 Which factors need to be considered when interpreting geographical distributions of disease? 5

1.6 What changes occurred in the reported occurrence of rheumatic fever in Denmark during the period covered in Figure 1.7? What might explain them? 7

1.7 What does Table 1.2 tell us about the contribution of asbestos exposure and smoking to the risk of lung cancer? 7

**Chapter 2 8**

2.1 What are the three epidemiological measures of disease frequency and how are they related? 8

2.2 Is prevalence rate a useful measure of the frequency of type 2 diabetes in different populations? What are the possible explanations for the variation in diabetes prevalence rates indicated in Table 2.3 8

2.3 What is the population attributable risk or attributable fraction (proportion) for smokers in the example in Table 2.4? 9

2.4 What measures are used to compare the frequency of disease in populations and what information do they provide? 9

2.5 The relative risk of lung cancer associated with passive smoking is low, but the population attributable risk is considerable. What is the explanation for this? 10

2.6 What is the main reason for standardizing rates to a population with a standard age distribution (for example, the WHO world standard population)? 10

2.7 If you want to know where the most cancer deaths per capita occur within a country, which is appropriate: crude death rates or age-standardized rates? 10

2.8 The crude death rate per 100 000 population for all cancers in Côte d’Ivoire is 70.5 and the age-standardized death rate is 160.2 per 100 000 population. What explains the large difference between these two rates? 11

2.9. The crude death rate for all cancers in Japan is 241.7 per 100,000 population and the crude death rate for all cancers in Côte d’Ivoire is 70.5 per 100,000 population. Is the death rate in Japan higher than that in Côte d’Ivoire? 11

**Question 3 11**

3.1 What are the applications and disadvantages of the major epidemiological study designs? 11

3.2 Outline the design of a case-control study and a cohort study to examine the association of a high-fat diet with bowel cancer. 12

3.3 What is random error and how can it be reduced? 13

3.4 What are the main types of systematic error in epidemiological studies and how can their effects be reduced? 14

3.5 Describe in which studies the relative risk (RR) and the odds ratio (OR) are used. Outline the reasons why they would be used in a particular study and not in another. 14

3.6 In the case of a rare disease, the OR and RR can have very similar values. Explain the reasons behind this similarity. 15

3.7 A cross-sectional study of Downs syndrome has found an association with birth order. What could be a cause of confounding and how would you avoid it? 15

**Conclusion 16**

**Bibliography Error! Bookmark not defined.**

# Introduction

The assignment will cover three topics namely, definition of epidemiology, measuring heath and diseases and types of studies

Epidemiology is the study of the distribution and determinants of health-related

states or events in human populations and the application of this study to the control

of health problem. This involves scientific approach of determining the diseases that are both prevalence and incidence in nature. Epidemiology is cross cutting subject as it is becoming relevance in medicine, nursing, environmental management, veterinary sciences, molecular and genetics and even in biotechnology. It is so helpful in theses sectors in policy development and planning purposes. It is therefore the aim of this assignment to elaborate more on this epidemiology concept with public health

Secondly, health and diseases are well measured in epidemiology this involves defining health and disease,Diagnostic criteria, Measuring disease frequency, Population at risk and Mortality among others.

Finally, in epidemiology there are mainly two types of studies namely observational which include descriptive, ecological studies, cohort studies among others. It is very important in local and international health well being of population in addition to environmental health. The following are questions and answers below is body of Assignment

# Question 1

## 1.1 Table 1.1 indicates that there were over 40 times more cholera cases in one district than in another. Did this reflect the risk of catching cholera in each district?

The risk of catching cholera in districts of London by drinking water from Southwark and Lambeth companies is reflected. As per John Snow’s research, he located the home of each person who died from cholera in London this was done comparing risk of affected by cholera in each district thus every population of London was exposed after noting an apparent association between the source of drinking water and the deaths. John Snow noted that people who were drinking water from Lambeth and Southwark were the ones dying after drinking and conclusion can be that 40 times more cholera cases in one district is really reflecting risk of catching cholera in each district.

## 1.2 How could the role of the water supply in causing deaths from cholera have been tested further?

Recently, there are many intervention that have been done so far. Just an example, many technologies have been developed to detect Vibrio Cholerae, which is a bacteria responsible for cholera disease. These technologies include water testing and even water treatment for public consumption like chlorination. Further more, some tests have been done through researches which include case control studies, cohort studies. The tests have also been tested further by application of statistical analysis like calculations of mortality rate, regression, ANOVA, Z test and designing research methods for comparing cases of cholera and later develop a policy and that facilitate human safety against cholera. Lastly there is also Geo spatial applications for mapping by using Geographic Information System. This is an Artificial Intelligent tool which helps to locate and demarcate the prone areas of cholera,

##

## 1.3 Why do you suppose the study shown in Figure 1.2 was restricted to doctors?

The study was restricted to doctors for the following reasons:

Data accuracy, Doctors were likely to give information which is credible for analysis and decision as they are so precise.

It was so easy to make follow up as cohort study. Since doctors are registered professionals, it makes so easy to trace there whereabouts

Homogeneity of data. This approach helped to sort out the problem of socio-economic differences among risk population to lung cancer.

Mortality and cause of death could be precise due to their access to healthcare and accurate medical report

##

## 1.4 What conclusions can be drawn from Figure 1.2?

The following are the conclusions that can be drawn:

Smoking are the primary cause of lung cancer.

Smoking reduces life expectancy

## 1.5 Which factors need to be considered when interpreting geographical distributions of disease?

There are there four main factors to consider when interpreting geographical distributions of diseases and these points are well discussed below.

Firstly, socioeconomic factors. The outbreaks of some diseases like cholera are still

frequent among poor populations, especially in developing countries. For example 2006, Angola reported 40 000 cholera cases and 1600 deaths; Sudan reported 13 852 cases resulting in 516 deaths in the first few months of the year. This shows that those regions which are developed tend to have advanced technologies which are accessible to people. These technologies are used in educational and public awareness which are used to detect, prevent and control diseases.

Secondly,Cultural background. The assertion is that cultural and behavior of people can facilitate certain type of diseases. For example, some cultural practices encourages people to take diet of a lot fats or eating meat or meat product which is not raw in nature like raw milk or drinking blood of animals. Furthermore, some traditional rituals during funerals can facilitate certain type of diseases like cholera, HIV /AIDS, lung cancer just mention a few.

Thirdly, demographic and biological factors. These factors include number of people per particular area thus population density, age structure meaning either it has youthful or aging. Some diseases tend to be common among children and youth than among older people of that particular area.

Lastly, literacy level. The high in come regions mostly high literacy level tend to have few cases of diseases comparing to those regions with low income level which mostly have have people with low literacy level. Some diseases are well preventable

### Case studies

Figure 1 pictures showing different practices in different geographical locations,



Practices that increases waterborne diseases



Practices that reduces waterborne diseases

below shows effects of geographical distribution on cholera case. It can be confirmed that areas with high population density and low income people tend to have high cases of cholera and this should be considered when interpreting cases of diseases like cholera

Figure 2: Geographical distribution of confirmed and suspected cholera cases (n=36 943) and deaths (n=1210) by District in Malawi, 3 March 2022 to 3 February 2023.

Source: Malawi Ministry of Health and Population

## 1.6 What changes occurred in the reported occurrence of rheumatic fever in Denmark during the period covered in Figure 1.7? What might explain them?

The changes were observed and can be categorized into three main changes:

The first change was between 1870 - 1900. During this time in Denmark rheumatic fever was common diseases as a result of poor living conditions with high level of streptococcus bacteria and staying in an overcrowded houses.

The second change was observed in the years between 1900 - 1930. During this period in Denmark rheumatic disease started to decline. As per epidemiological researches it shows that during this time there was some improvements in living standards, government started to take serious measures on campaign awareness on infectious diseases as a rheumatic public health measures with some medical treatments.

Third change was observed between 1930 - 1960. During this time in Denmark experienced and observed sharp decline thus it was decreasing at an increasing rate. So far epidemiological research papers indicates that government introduced policies which helped to expand health care which later helped to deal with rheumatic completely. Moreover drugs like antibiotics were used in health cares to treat streptococcus hence rheumatic diseases declined sharply

## 1.7 What does Table 1.2 tell us about the contribution of asbestos exposure and smoking to the risk of lung cancer?

As it is shown as people exposed to asbestos alone were 11 depicting 1.8 % risk (11+602) , those people exposed to smoke were 123 depicting 16.9% risk (123+302) while combination of asbestos and smoking is 602 depicting 79.6% risk (11+123+602). From this statistical analysis it can be concluded that a person who is smoking and also is exposed to asbestos is at high risk of suffering from lung cancer comparing to those that are not either exposed to asbestos or smoking or the one who is exposed to either of them.

# Chapter 2

## 2.1 What are the three epidemiological measures of disease frequency and how are they related?

These are Incidence, prevalence and case mortality rate.

Incidence is the proportion of number of new cases of a disease over a period of time to number of persons exposed to a risk during that particular period. This incidence can be categorized as cumulative incidence or incidence density. Prevalence refers to all people in a defined population with the disease or condition at a given point in time or over a given period of time. Prevalence can be in the form of point prevalence, period prevalence or lifetime prevalence. Mortality rate is the total number of deaths due to disease in a population over a given specific time. Mortality rate can be crude mortality rate, case fatality rate or cause specific mortality rate. Measuring prevalence, incidence and mortality involves the counting of cases in defined populations at risk. Numerically Prevalence is same as incidence multiply by average duration of disease. However high mortality rate reduces prevalence on the other hand it increases incidence like aggressive cholera cases.

## 2.2 Is prevalence rate a useful measure of the frequency of type 2 diabetes in different populations? What are the possible explanations for the variation in diabetes prevalence rates indicated in Table 2.3

Prevalence rate is indeed a useful measure of the frequency of type 2 diabetes. The following are the possible explanations:

Generally men of Chinese, Singapore and Sri Lanka origin have a high percentage of type 2 diabetes comparing to women of Chinese, Singapore and Sri Lanka origin thus considering gender of the population

Generally people from urban areas have high percentages of type 2 diabetes comparing to people from rural areas in people from Fiji and South India thus considering socio-economic status of the population.

## 2.3 What is the population attributable risk or attributable fraction (proportion) for smokers in the example in Table 2.4?

The attributable risk of stroke = (incidence of stroke disease in population - incidence of stroke among unexposed)/ (incidence of stroke disease in population)

 = (30.2 -17.7)/(30.2)

 = 0.414

 = 41.4%

41.4% means that of all strokes in USA in 1976 are attributable to smoking.

Relative Risk = (Incidence in smokers /Incidence in never -smokers

 = 49.6/17.7

 = 2.8

Attributable Fraction = (Relative Fraction - 1)/Relative Fraction

 = (2.8 - 1)/2.8

 = 0.643

 = 64.3%

This means that 64.3% among smokers are attributable to smoking.

## 2.4 What measures are used to compare the frequency of disease in populations and what information do they provide?

The first on is mortality rate which provide information on the ratio or proportion of number of deaths per total population over a period of time of which numerically its number deaths /total population

Secondly, case fatality rate which gives information on the proportion of people with disease who die from it thus ((number of deaths from disease/number of diagnosed cases) × 100%)

Thirdly, attack rate show cases proportion of a population that develop disease during a short outbreak of which numerically its number of new cases per population at risk

The fourth point is cumulative Incidence that provide direct proportional of number of people who get disease during specified period of time per number of people free of the disease in the population at at the beginning of the period usually per 1,000 population

Furthermore, Incidence rate or density shows number of new cases of a disease occurring in a population during a specified time period, divided by the total person time at risk

Lastly Prevalence shows a population with the disease or condition at a given point in time or over a given period of time

## 2.5 The relative risk of lung cancer associated with passive smoking is low, but the population attributable risk is considerable. What is the explanation for this?

The explanation can be based on the measure on how more likely passive smokers are to develop lung cancer compared to unexposed exposed ones and this is Relative Risk (RR) . the other explanation can be based on the estimation on the proportion of lung cancer cases in the population that could be prevented if passive smokers were eliminated which is Population Attributable Risk

##

## 2.6 What is the main reason for standardizing rates to a population with a standard age distribution (for example, the WHO world standard population)?

The main reason is to allow for valid comparisons across different populations or over time by removing effects of age distributions, trend analysis over time and age confounders.

## 2.7 If you want to know where the most cancer deaths per capita occur within a country, which is appropriate: crude death rates or age-standardized rates?

Age-standardized rates is the best choice over crude death rates. The reason being that age-standardized rates considers variations in age distribution which provides a fair comparisons between regions with different age demography unlike using crude death rates for crude death rates just divide the number of cancer deaths but the total population in an area which has a problem of whereby sometimes areas with aged populations will have high cancer crude death rates comparing to youthful population hence biasness,

##

## 2.8 The crude death rate per 100 000 population for all cancers in Côte d’Ivoire is 70.5 and the age-standardized death rate is 160.2 per 100 000 population. What explains the large difference between these two rates?

Due to factors of higher infectious disease mortality rate and lower life expectancy makes fewer people to live older ages where cancer risk is highest hence age -standardized death rates is up to 160.2 per 100000 as such the true underlying risk of cancer is made known. Moreover in Cote d'Ivoire the population is young hence low crude death rate of 70.5 per 100000.

##

## 2.9. The crude death rate for all cancers in Japan is 241.7 per 100,000 population and the crude death rate for all cancers in Côte d’Ivoire is 70.5 per 100,000 population. Is the death rate in Japan higher than that in Côte d’Ivoire?

The death rate in Japan is higher by 3.4 times than that of the Côte d’Ivoire. This is due to aging population in Japan and also availability of reports and diagnosis, lifestyle like smoking in Japan and also availability of other diseases which are infectious like HIV/AIDS in Côte d’Ivoire due to low income status.

# Question 3

## 3.1 What are the applications and disadvantages of the major epidemiological study designs?

The first major epidemiological study design is observational study. This study design can be used in public health when studying risk factors of certain disease, social science when analyzing certain lifestyle and medical and clinical research when investigating long-term effects treatments. However observations study has the challenges in controlling variables thus establishing causation rather just shows correlation, confounding factors as hidden variables may influence results, biasness from observer’s point of view and observation study has problem of ethical and practical limitations.

The second major study design is experimental study design. This can be used in medicine and clinical trials in comparing treatment outcomes, agriculture and biology in studying biotechnological studies like genetic modification plants and animals, engineering and technology in new materials testing for durability. However experimental study design has ethical limitations as some experiments can expose humans to harmful substances, it is expensive like setting up controlled environments such as clinical trials, it is time consuming as longitudinal experiments in tracking disease progression over years, sometimes it might not be realistic as it may lack real world applicability unlike observational study design which reflects real world behaviour.

## 3.2 Outline the design of a case-control study and a cohort study to examine the association of a high-fat diet with bowel cancer.

### *Case-control study*

The design looks back at the effect of high fat diet to bowel cancer causation.

The following are steps:

*Selection of cases*

This involves identifying individuals who has already been diagnosed with bowel cancer due to high fat diet which are confirmed by medical practitioners

*Selection of controls*

This involves choosing healthy individualists without bowel cancer matched to cases by age, sex, and other relevant factors to deal with confounders

*Exposure Assessment*

This is use of dietary questionnaire or interviews or nutritional records to asses past high fat diet consumption like frequency , types of fats

*Analysis*

Comparison of proportion of cases of high fat and controls with a history of high fat diet and estimating the association thus relative risk by use of odds ratio (OR) , which is the ratio of the odds of exposure among the cases to the odds of exposure among the controls

### *Cohort control study*

It is also called follow up study

Objective: Determine whether long-term high fat consumption increases bowel cancer

Study design:

*Data source*

Pre-existing cohort like electronic health records Exposure. This involves historical dietary records like health screenings

*Outcome*

Bowel cancer cases identified via cancer registries or hospital data Comparison groups

The comparison should be between high fat consumers historical records against low fat consumers

*Follow-up*

Follow up is done retrospectively maybe 8 years of past data

*Analysis*

The statistical analysis is applied to the data. This include hazard ratio for bowel cancer in high fat diet against low fat diet. Dose response analysis should also be analysed as fat intake which is continuous variable adjustments for confounders Age, sex, lifestyle and family history should be considered.

## 3.3 What is random error and how can it be reduced?

Random error is a type of measurement error that occurs due to unpredictable and uncontrollable variations in experimental or observational studies. There are three major sources of random error which are individual biological variation, sampling error and

measurement error. Sampling error is usually caused by the fact that a small sample is

not representative of all the population’s variables. The best way to reduce sampling

error is to increase the size of the study. Individual variation always occurs and no

measurement is perfectly accurate. Measurement error can be reduced by stringent

protocols, and by making individual measurements as precise as possible. Investiga

tors need to understand the measurement methods being used in the study, and the

errors that these methods can cause. Ideally, laboratories should be able to document the accuracy and precision of their measurements by systematic quality control

procedures.

## 3.4 What are the main types of systematic error in epidemiological studies and how can their effects be reduced?

Systematic error (or bias) occurs in epidemiology when results differ in a systematic

manner from the true values and below are types of systematic error

Selection error which occurs when there is a systematic difference between the characteristics of the people selected for a study and the characteristics of those who are not. For example, self-selection bias and loss of follow up. The selection error can be reduced by randomizing the target sample, replication or multiple recruitment sources and statistical application of inverse probability weighting to account for missing data.

Measurement error or information bias which occurs when the individual measurements or classifications of disease or exposure are inaccurate – that is, they do not measure correctly what they are supposed to measure. This may include recall bias, interview bias, and misclassification bias. This can be reduced by use of objective measures, validating measurement tools for precision, and also replicating or use multiple sources of data

## 3.5 Describe in which studies the relative risk (RR) and the odds ratio (OR) are used. Outline the reasons why they would be used in a particular study and not in another.

The relative risk compares the probability of a disease occurring in an exposed group versus an unexposed group. It can be used in cohort studies, randomized controlled trials and cross sectional studies. This is so because relative risk studies are direct risk interpretation , requires known incidence and easy for public health decisions as it is simple to communicate. This shows that it is invalid in case control studies and rare outcome studies.

The odd ratio compares the odds of a disease occurring in one group versus another. It is used in case control studies, logistic regression, cross-sectional studies and when pooling different study types thus meta analyses. This is the case because Odds ratio work with fixed samples, it is convenient mathematically and approximates relative ratio when outcome is rare. The odds ratio cannot be preferred in cohort studies and in public heath communication

## 3.6 In the case of a rare disease, the OR and RR can have very similar values. Explain the reasons behind this similarity.

There are mainly three reasons for the similarity and below are these similarities

Disease prevalence is low. In a rare diseases, the probability of disease in the unexposed group is very small

When disease is very rare, Relative Risk is approximated by Odds Ratio.

There is convergence mathematically. That is when disease is very rare, the denominators in the odds are almost one making the odds almost equal to the risk. That means the ratio of odds becomes nearly identical to the ratio of risks

## 3.7 A cross-sectional study of Downs syndrome has found an association with birth order. What could be a cause of confounding and how would you avoid it?

There can be many other confounding factors that can cause Downs syndrome and these include:

Firstly maternal age. Women who have high birth order children and those women that are at advanced maternal age are likely to give birth to a child with Down’s syndrome comparing to those women who are having their child.

Second point is confounding mechanisms . that is if maternal age is not accounted for , the observed association between birth order and Down’s Syndrome could simply reflect the underlying relationship between maternal age and chromosomal abnormalities

On the part of prevention of Down’s syndrome, stratification can be used. This involves data analysis within strata of maternal age thus if association between birth order and Down syndrome disappears within age groups, maternal age was likely the confounder

The second prevention is by use of multivariate regression. This involves adjustment for maternal age in a logistic regression model to isolate the independence effect of birth order

The third one is by matching. This involves case control design match Down syndrome cases and controls based on maternal age to eliminate its confounding.

Restriction is the fourth prevention guide. This is the study to mothers within narrow age range to reduce variability in maternal age

# Conclusion

The assignment have articulated three main points to me and analytical skill being the first. Epidemiological data from different countries with different cases had to be analysed and concluded at the end. This is a profound skill as of now e very data from any other field being financial, medical,gender based data, veterinary, agriculture and the list is endless that I can have, I can draw conclusions from it.

Observational skill is the second skill. Before this assignment I was not aware whenever am surrounded with people. Surprisingly am now able to observe the behaviour wise of people not the they are sick or stricken by any other disease. People can show different behaviors due to their past experiences like diseases and accidents. The conclusion from this point is that who so ever had any case before being disease or accidents or near death experiences tends to be careful to risk factors comparing to those that had never had them before. This observational skill is from this assignment.

The third being the last is writing skill. This assignment required me to source information from different sources and combine them to have one concrete answer. This is amazing for me that I have done it like that in that way. There are so many books or resources that I had to consult and turn them into writing for easy communication . this is an experience that I have acquired..

Bibliography

1. Snow J. On the mode of communication of cholera. London, Churchill,1855.

(Reprinted in: Snow on cholera: a reprint of two papers. New York, Hafner

Publishing Company, 1965).

1. Last JM. A dictionary of epidemiology, 4th ed. Oxford, Oxford University Press,

2001.

1. Shibuya K, Boerma T. Measuring progress towards reducing health inequalities.

Bull World Health Organn 2005;83:162.

1. Soleman N, Chandramohan D, Shibuya K. Verbal autopsy: current practices and challenges. Bull World Health Organn 2006;84:239-45.

5. World Health Report 2005: Make every mother and child count. Geneva, World

Health Organization, 2005.

1. World Health Report, 2023 Disease outbreak . Malawi, World Health Organization ,February, 2023