**AFRICAN INSTITUTE FOR PROJECT MANAGEMENT STUDIES**

**(AIPMS)**

**COURSE NAME: DIPLOMA IN PUBLIC HEALTH (2019/2020)**

**STUDENT NAME: FRANCIS MORI MOGA.**

**CONTINUOUS ASSESSMENT TEST (CAT-3)**

**MODULE 3 : EPIDEMIOLOGY.**

**Question 1:**

**Distinguish between descriptive epidemiology and analytical epidemiology.**

Epidemiology is the science concerned with the study of the factors determining and influencing the frequency and distribution of disease, injury, and other health-related events and their causes in a defined human population for the purpose of establishing programs to prevent and control their development and spread. Epidemiology is a very common and important activity in public health departments. It clarifies clinical and demographic characteristics of diseases and conditions. It identifies who is at risk and provides clues to the causes of disease. And finally, it guides preventive measures and interventions.

Epidemiologic studies ask five questions: what, who, where, when, and why. These five Ws remind us to organize questions about possible exposures (or risk factors).

What includes the diagnosis or clinical information?

Who, where, and when include person, place, and time information?

Why involves causes of disease, risk factors, and modes of transmission.

Epidemiologists study what, who, where, and when, or in other words, clinical data plus person, place, and time information through descriptive epidemiology, which investigates the distribution of diseases or conditions, while they study why and how, or in other words, causes, risk factors, and modes of transmission, through analytic epidemiology, which investigates the determinants of

diseases or conditions.

Epidemiologic studies /activity is often divided into two types, descriptive and analytic, and each of these types of epidemiology uses specific kinds of studies.

Descriptive epidemiology:

Descriptive studies examine patterns of disease occurrence, with a focus on person, place, and time. It primarily concerned with documenting the magnitude of a disease or injury problem and describing it in terms of the personal characteristics and behaviour of those at risk, and the place and timing of occurrence. It also is a leading source of research hypotheses These studies use relatively accessible data for program planning, to estimate caseloads, to determine the amount of public health resources needed, or to identify high-risk groups. Many public health practitioners carry out descriptive epidemiologic studies within their jurisdictions.

Epidemiologists also use descriptive studies to generate hypotheses that need to be confirmed or ruled out by analytic studies.

Person:

People’s sociodemographic characteristics and behaviour variously increase or decrease their risk for contracting a disease or becoming injured. For example, the poor are much more likely than the to live in sub-standard housing in heavily polluted areas with their attendant health hazard.

* Demographic characteristics e.g. age, sex, race, marital status, number of children.
* Socio-economic characteristics e.g. Social class, employment status, occupation.
* Life style behaviour e.g. drinking alcohol/ smoking marijuana and driving.

Time:

Time is important in characterizing disease or injury occurrence, if incidence rates or case numbers markedly increase or decrease, if there is a seasonal effect, and for determining whether a communicable disease is attributable to a point-source infection, a continuing common source infection, or an intermittent exposure.

* Are diseases (injury) rates or case numbers variable or constant
* Do rates or case numbers vary seasonally?
* Is the disease attributable to a point source of infection or propagated transmission?

Place:

Place information may include where people become sick, such as in the home or a vacation spot.

Place also includes where an exposure occurred, such as in a restaurant serving contaminated food or a cruise ship.

And finally, place also includes the source of contamination,

such as a farm or a poultry-packing plan.

Descriptive epidemiology describes condition or diseases. Descriptive studies identify patterns in person, place, and time, and falls under three basic types of studies: a case report, a case series, and an incidence study.

A case report is a detailed description of the person place, and time information of a specific case of disease or condition. Case reports usually are about unexpected symptoms in disease, an unexpected event while treating a patient, or unique therapeutic approaches.

A case series describes the person, place, and time information about a group of cases. It can be retrospective, looking back in time, or prospective, looking forward in time, and usually involves small number of patients, such as those who were given similar treatment.

Case reports and series permit discovery of new diseases, unexpected effects, and provide data for generating hypotheses. Data from a case series may be used in analytic studies to investigate possible causal factors.

An incidence study describes the incidence, or number of new cases of a disease

or condition, during a specific time in a specific population. One advantage of incidence

studies is that they allow calculation of true rates of disease occurrence for

**Analytic epidemiology:**

Analytic studies, on the other hand, are usually larger and more complex than descriptive studies. It often used to assess determinants of diseases, focus on risk factors and causes, and analyse the distribution of exposures and diseases. A key feature of analytic studies is that it uses comparison groups. In contrast to descriptive studies, which generate hypotheses, analytic studies is used to test hypotheses. it used to look for and measure associations.

Epidemiologists conduct two main types of analytic studies: experimental and observational.

Experimental studies use a randomized selection process. A process based on chance is used to assign study subjects to different exposure groups.

Experimental studies may be either clinical, such as studying a new drug to prevent disease, or community-based, for example, studying the overall effectiveness of a new drug in preventing a specific disease in the community.

**Experimental studies:**

Experimental studies involve assigning subjects to exposures randomly and following them over time to determine if they develop or recover from disease. Experimental studies fall into two categories: clinical trials and community trials.

Clinical trials use data from individual people. The investigator randomly determines the type of exposure, for example, to a new drug to treat cancer. The study participants are then followed to determine the effect of treatment on them. People who received the new drug are

compared with people who received an older drug or no drugs at all.

In community trials, on the other hand, the study group is the entire community, rather than individuals. Researchers might investigate whether a media campaign to reduce smoking was effective. The researchers would select a community to receive the media campaign. They would then compare the smoking rates over time in this community with those in another community that did not to receive the intervention.

**Observational studies:**

In observational studies, the researcher does not determine who receives the exposure. The researcher simply observes or records the study participants and their outcomes.

Observational studies come in four main types: Cohort, case-control, cross-sectional, and ecologic.

In cohort studies, researchers determine the study population’s exposure and observe over time who gets ill.

In case-control studies researchers identify people who are ill and select or identify a comparison or control group of people who aren’t ill. Researchers then compare the prior exposures of the two groups.

In cross-sectional studies, researchers survey both the exposure and the condition or disease among individuals at a single moment in time.

In ecologic studies researchers survey the community level exposures and the condition or disease, but entire populations are compared rather than individual.

**Question 2:**

**Write down and explain the mathematical expression of the following.**

**i. Incidence**

**ii. Prevalence**

**Incidence:**

Incidence is the rate of new (or newly diagnosed) cases of the disease. It is generally reported as the number of new cases occurring within a period of time (e.g., per month, per year). It is more meaningful when the incidence rate is reported as a fraction of the population at risk of developing the disease (e.g., per 100,000 or per million population

* Incidence rate describes how quickly disease occurs in a population.
* Express the risk of acquiring the disease.
* Relate the number of new cases of the disease to the population "at risk" of getting the disease, in a certain period of time

Incidence rate (IR) of disease "X" in year "Y" and place "Z":

Numerator (N): Number of new cases of disease "X" that occurred in place "Z" during year "Y“

Denominator (P): Population in "Z" place at risk of getting the disease "X" during year "Y“

Base (C): constant, usually 100,000

Hence,

**N(new)**

**IR = ------------ x C**

**P(at-risk)**

**Example:**

In one regional hospital, (Wau) teaching hospital in South Sudan in 2018, the number of new cases of Hepatitis B virus was reported to be 8,255 cases. The estimated population of the town was approximately 1,655,722 in that particular year. Calculate the incidence rate of Hepatitis B virus in 2018 in Wau town.

Numerator = new cases of Hepatitis B = 8255

Denominator = 1,655,722 estimated population

10n = 100,000

Incidence rate = (8,255 / 1,655,722) x 100,000

IR = 498.5 new cases of Hepatitis B infection per 100,000 population.

**Prevalence:**

Prevalence is the proportion of persons in a population who have a particular disease or attribute at a specified point in time or over a specified period of time. Prevalence differs from incidence in that prevalence includes all cases, both new and old or pre-existing ones in the population at the specified time.

* The numerator of an incidence proportion or rate consists only of persons whose illness began during the specified interval. The numerator for prevalence includes all persons ill from a specified cause during the specified interval **regardless of when the illness began**. It includes not only new cases, but also pre-existing cases representing persons who remained ill during some portion of the specified interval.
* Prevalence is based on both incidence and duration of illness. High prevalence of a disease within a population might reflect high incidence or prolonged survival without cure or both.
* Conversely, low prevalence might indicate low incidence, a rapidly fatal process, or rapid recovery.
* Prevalence rather than incidence is often measured for chronic diseases such as diabetes or osteoarthritis which have long duration and dates of onset that are difficult to pinpoint

**Method for calculating prevalence of disease**

All new and pre-existing cases during a given time period

Population during the same time period

Prevalence rate (PR) of disease "X" in year "Y" and place "Z":

Numerator (N): Number of new and old cases of disease "X" present in place "Z" during year "Y“

Denominator (P): Population in "Z" place during year "Y“

Base (C): constant, usually 100,000

Hence,

**N(new+old)**

**PR = --------------- x C**

**P(total)**

**Example:**

In a survey carried out in one clinic to determine the number of pregnant mothers who attend antenatal services, it was reported that 250 mothers out of the 720 pregnant mothers do attend antenatal services. Calculate the prevalence of the pregnant mothers visiting the clinic for antenatal services.

Numerator = 250 mothers who visit the clinic

Denominator = 720 total pregnant mothers.

Prevalence = (250 / 720) x 100 = 0.347 x 100 = 34.7%.

**Question 3:**

**Apart from Randomized trials, describe four (4) other epidemiological research designs.**

Research design is defined as a framework of methods and techniques chosen by a researcher to combine various components of research in a reasonably logical manner so that the research problem is efficiently handled. It provides insights about “how” to conduct research using a particular methodology. Every researcher has a list of research questions which need to be assessed – this can be done with research design.

**Crossover Design**

A controlled trial where each study participant has both therapies, e.g., is randomised to treatment A first, at the crossover point they then start treatment B. Only relevant if the outcome is reversible with time, e.g., symptoms.

*Advantages:*

* all subjects serve as own controls and error variance is reduced thus reducing sample size needed;
* all subjects receive treatment (at least some of the time);
* statistical tests assuming randomisation can be used;
* blinding can be maintained.

*Disadvantages:*

* all subjects receive placebo or alternative treatment at some point;
* washout period lengthy or unknown;
* cannot be used for treatments with permanent effects

**Cohort Study.**

“Cohort “study is used to describe/define a group or set of people (subjects, cases, patients etc.), who are investigatively followed for a certain period to observe what changes have occurred to the members of that cohort over that particular time period

Data are obtained from groups who have been exposed, or not exposed, to the new technology or factor of interest (e.g. from databases). No allocation of exposure is made by the researcher. Best for study the effect of predictive risk factors on an outcome.

*Advantages:*

* ethically safe;
* subjects can be matched;
* can establish timing and directionality of events;
* eligibility criteria and outcome assessments can be standardised;
* administratively easier and cheaper than RC

*Disadvantages:*

* controls may be difficult to identify;
* exposure may be linked to a hidden confounder;
* blinding is difficult;
* randomisation not present;
* for rare disease, large sample sizes or long follow-up necessary

**Case-Control Studies**

Case-control studies take a representative group of subjects, and consider known outcomes in that group, and will, by dint of study design, and careful choice of controls, attempt to refer backwards to determine what bearing any number of possible exposures (and the degree to which subjects were exposed to them) had on these outcomes. In a case-control study, the distribution of the disease outcomes (i.e. how many are ill, who is ill, how ill they are etc.) is also known at the beginning of the study period.

Patients with a certain outcome or disease and an appropriate group of controls without the outcome or disease are selected (usually with careful consideration of appropriate choice of controls, matching, etc) and then information is obtained on whether the subjects have been exposed to the factor under investigation.

*Advantages:*

* quick and cheap;
* only feasible method for very rare disorders or those with long lag between exposure and outcome;
* fewer subjects needed than cross-sectional studies

*Disadvantages:*

* reliance on recall or records to determine exposure status;
* confounders;
* selection of control groups is difficult;
* potential bias: recall, selection.

**Cross-Sectional Survey**

A study that examines the relationship between diseases (or other health-related characteristics) and other variables of interest as they exist in a defined population at one particular time (i.e. exposure and outcomes are both measured at the same time). Best for quantifying the prevalence of a disease or risk factor, and for quantifying the accuracy of a diagnostic test.

*Advantages:*

* cheap and simple;
* ethically safe.

*Disadvantages:*

* establishes association at most, not causality;
* recall bias susceptibility;
* confounders may be unequally distributed;
* Neyman bias;
* group sizes may be unequal.

**Question 4:**

**Data from hospital records are one of the most important sources of information in**

**epidemiologic studies.**

**a) Outline the limitations of using hospital data.**

Data in Health refers to routinely or automatically collected datasets, which are electronically captured and stored. It is reusable in the sense of multipurpose data and comprises the fusion and connection of existing databases for the purpose of improving health and health system performance. The analysis whether it is structured or unstructured, usually requires significant logistic efforts and computing power. Data analysis consists of generating and collecting data, storing and processing, and, finally the distribution and analysis of the relevant data. Applications or use of hospital data contribute to the following achievements:

* Increasing the effectiveness and quality of treatmentsby e.g.: earlier disease intervention,
* reduced probability of adverse reactions to medicines,
* Less medical errors,
* determination of causalities, understanding of co-morbidity,
* cross-linkage of health care providers and professionals,
* intensification of research networks, and
* fusion of different networks such as social networks, disease networks or medicine networks,
* widening possibilities for the prevention of diseasesby identification of risk factors for disease at population, subpopulation, and individual levels, and by improving the effectiveness of interventions to help people achieve healthier behaviours in healthier environments,
* the improvement of pharmacovigilance and patient safety through the ability to make more informed medical decisions based on directly delivered information to the patients,
* prediction of outcomes, e.g. containment and improvement of chronic diseases, global infectious disease surveillance through evolving risk maps and better **un-**derstanding of demographic challenges and trends as well disease trans-mission pathways,
* knowledge dissemination, e.g. help physicians to stay current with the latest evidence guiding clinical practice, and
* reduction in inefficiency and waste, improvement in cost containment

**Potential Limitations of hospital data:**

1. Hospital detected cases are not inclusive and are selected according to:

(a) Personal characteristics, e.g., age, race, sex, socioeconomic status

(b) Severity of disease with a tendency to advanced cases

(c) Associated conditions

(d) Administrative admission policies.

2. Difficulty of finding adequate control groups.

3. Hospital records are not primarily designed for research, because of:

(a) Incomplete and unstandardized information

(b) Diagnostic variability among hospitals.

4. The community population at risk cannot be precisely defined.

5. Duplicate admissions raise problems in determining incidence and prevalence rates.

**6. Privacy and security** are of importance for health care businesses. Successful attacks on health care data can be extremely lucrative for criminals and extremely damaging for organizations. And the financial costs of data breaches may be just the beginning – reputational costs are harder to measure but may linger for long periods. And individuals whose data is stolen may suffer most of all, since health records contain personal data ranging from credit card numbers to details about diagnoses and lab tests, raising threats of identity theft and even blackmail.

**7. Data retention:** Health data must stay accessible for a period (possibly for more than 5 years). That means businesses need to take a long-term approach to data stewardship and keep track of when the data gets accessed, by whom, and for what purpose. Medical data management software allows users to establish access privileges and processes, such as those that give temporary data viewing capabilities to representatives in different departments in a hospital. These products can index data and notes and track when data entered the system. Organizations must put processes in place to periodically sort through the data to delete it when appropriate or modify and anonymize it to use it in new ways, such as to gauge trends across several years.

8. **Data management:** Health organizations face big-data-related challenges that can impact patient safety. All data that health organizations collect needs to be described, formatted, deduped and checked for accuracy, and made accessible for various uses — medical, billing, administrative — and the volume and velocity of big data makes this task more difficult.

9. **Data accessibility:** All data management strategies fall short if they don't result in content that's accessible and in the correct format for reporting. Data analysts must be empowered to access the data they need and share what the data reveals.

**Q4 b**

**Describe the possible sources of error in interview surveys**

**Questionnaire as a source of error.**

Ideally a question will convey to the respondent the meaning of interest to the researcher. However, several linguistic, structural, and environmental factors affect the interpretation of the question by the respondent. These factors include the specific question wording, the structure of each question (open versus closed), and the order in which the questions are presented. Question wording is often seen as one of the major problems in survey research; although one can standardize the language read by the respondent or the interviewer, standardizing the language does not imply standardization of the meaning. In addition, a respondent's perception of the intent or meaning of a question can be shaped by the sponsorship of the survey, the overall topic of the questionnaire, or the environment more immediate to the question of interest, such as the context of the previous question or set of questions or the specific response options associated with the question

**Respondent as a source of error.**

Once the respondent comprehends the question, he or she must retrieve the relevant information from memory, make a judgment as to whether the retrieved information matches the requested information, and communicate a response. The retrieval process is potentially fraught with error, including errors of omission and commission. As part of the communication of the response, the respondent must determine whether he or she wishes to reveal the information. Survey instruments often ask questions about socially and personally sensitive topics. It is widely believed, and well documented, that such questions elicit patterns of underreporting (for socially undesirable behaviours and attitudes) as well as overreporting (for socially desirable behaviours and attitudes).

**Interviewer as a source of error**.

For interviewer-administered questionnaires, interviewers may affect the measurement processes in one of several ways, including:

* Failure to read the question as written;
* Variation in interviewers' ability to perform the other tasks associated with interviewing, for example, probing insufficient responses, selecting appropriate respondents, or recording information provided by the respondent; and
* Demographic and socioeconomic characteristics as well as voice characteristics that influence the behaviour and responses provided by the respondent.

The first two factors contribute to measurement error from a cognitive or psycholinguistic perspective in that different respondents are exposed to different stimuli; thus variation in responses is, in part, a function of the variation in stimuli. All three factors suggest that interviewer effects contribute via an increase in variable error across interviewers. If all interviewers errored in the same direction (or their characteristics resulted in errors of the same direction and magnitude), interviewer bias would result. For the most part, the literature indicates that among well-trained interviewing staff, interviewer error contributes to the overall variance of estimates as opposed to resulting in biased estimates.

**Other essential survey conditions as source of error.**

Any data collection effort involves decisions concerning the features that define the overall design of the survey, here referred to as the essential survey conditions. In addition to the sample design and the wording of individual questions and response options, these decisions include:

* Whether to use interviewers or to collect information via some form of self-administered questionnaire;
* The means for selecting and training interviewers (if applicable);
* The mode of data collection for interviewer administration (telephone versus face to face);
* The choice of respondent rule, including the extent to which the design permits the reporting of information by proxy respondents;
* The method of data collection (paper and pencil, computer assisted);
* The extent to which respondents are encouraged to reference records to respond to factual questions;
* Whether to contact respondents for a single interview (cross-sectional design) or follow respondents over time (longitudinal or panel design);
* For longitudinal designs, the frequency and periodicity of measurement;
* The identification of the organization for whom the data are collected; and
* The identification of the data collection organization.

No one design or set of design features is clearly superior with respect to overall data quality. For example, as noted, interviewer variance is one source of variability that obviously can be eliminated through the use of a self-administered questionnaire. However, the use of an interviewer may aid in the measurement process by providing the respondent with clarifying information or by probing insufficient response

**Question 5;**

**Explain the main determinants of health**.

Health determinants are a range of factors that influence the health status of individuals or populations. At every stage of life, health is determined by complex interactions between social and economic factors, the physical environment and individual behaviour. There are different perspectives in expressing the determinants of health of an individual or a community. There are mainly five major health determinants, and these are:

**Biological and genetics.**

Biological or genetics are Affects or characteristics of the body that are genetic in origin or have a genetic component that may directly and measurably determine health of an individual. Every Human being is made of genes. In addition, there are factors, which are genetically transmitted from parents to offspring. As a result, there is a chance of transferring defective trait and/or disability (e.g inheritance of the BrCa1 gene increases individual susceptibility to breast and ovarian cancer; genetic causes of birth defects/ congenital abnormalities,). Some biological and genetic factors affect specific populations more than others. Sickle cell is another example of a genetic determinant of health. Sickle cell is a disease condition that people inherit when both parents carry the gene for the sickle cell.

**Environmental Factors:**

Environmental factors are any external agents that contributes to poor health being either (biological, chemical, or physical) .

Environmental factors that could influence health include:

a. Life support, food, water, air etc

b. Physical factors include plants, weather, climate, Rain fall etc.

c. Biological factors: microorganisms, toxins, Biological waste,

d. Psycho-social and economic e.g. Crowding, income level, access to health care

e. Chemical factors: industrial wastes, agricultural wastes, air pollution, etc

Poor health outcomes are often made worse by the interaction between individuals and their social and physical environment. For example, millions of people in the United States live in places that have unhealthy levels of ozone or other air pollutants. In counties where ozone pollution is high, there is often a higher prevalence of asthma in both adults and children compared with state and national averages. Poor air quality can worsen asthma symptoms, especially in children.

**Behavioural factors:**

Lifestyle (Behavior): is an action that has a specific frequency, duration, and purpose, whether conscious or unconscious. It is associated with practice. It is what we do and how we act. Recently lifestyle by itself received an increased amount of attention as a major determinant of health. Lifestyle of individuals affects their health directly or indirectly. Behaviours , activities, actions, or patterns of actions undertaken by individuals that have the potential to influence health, including behaviours undertaken to promote, protect or maintain health, whether or not such behaviours are objectively effective towards that end (adapted from WHO 1998 ‘Health behaviour’). Examples include eating contaminated food, tobacco smoking. Unsafe sexual practice, alcohol use, drugs use etc.

**Socio-economic factors:**

**Socio-economic** factors – are social determinants that influence health, and are associated with an individual income (including levels and distribution) and associated levels of poverty and inequality, education, the existence of social support networks, etc. The effects of each of these factors on health are, however, highly dependent on other socioeconomic variables as well as the policy context, including accessibility and effectiveness of health and welfare systems’. These socio-economic factors are further subdivided into social and economic determinants:

Social determinants of health reflect the social factors and physical condition of the environment in which people are born, live ,play, work and age, including social, cultural, and gendered roles, religious belief or spirituality, health cognition; and other societal contributors such as social attitudes, class/ caste systems, community involvement and social and support systems.

Economic determinants that influence health, including education, literacy and health literacy; employment status, occupation and industry; living standard, financial resources; and the services, systems and policies that are provided by or govern societies and states including the health system, services and policies.

.

**Heath service provision;**

Both access to health services and the quality of health services can impact health.  Barriers to accessing health services include lack of availability. high cost, lack of insurance coverage, limited language access.  These barriers to accessing health services lead to unmet health needs, delays in receiving appropriate care, inability to get preventive services, hospitalizations that could have been prevented.

Effective Health service in the community should be or able to address the following.

a. Availability of health service to People living in areas where there is no access to health service.

b. Health services provided in the community should be affordable.

c. Acceptability of the service by the community

d. Accessibility: in terms of physical distance, finance etc

e. Quality of care that mainly focuses on the comprehensiveness, continuity and integration of the health care.

(NB: *Scarcity of Health Services leads to inefficient health service and resulting in poor quality of health status of people)*

References:

* Miller, K., Willis, G.B., Eason, C., Moses, L., and Canfield, B. (2005). Interpreting Results of Cross-Cultural Interviews: A Mixed Methods Approach. Methodological Aspects in Cross National Research, J.H.P. Hoffmeyer-Zlotnik and J.A. Harkness (eds). ZUMA, Mannheim, 79–92. Available
* Friis RH, Sellers TA. Epidemiology for public health practice, 4th Edi. Sudbury, MA: Jones and Bartlett publishers;2009
* Bortz WM 2005. Biological basis of determinants of health, American Journal of Public Health, 95(3):389-392.
* Andersen, R., Kasper, J., Frankel, M., and Associates. 1979. *Total Survey Error—Applications to Improve Health Surveys.* San Francisco, CA: Jossey-Bass
* Determinants of Health | Working in Health Promoting Ways [Internet]. Dhhs.tas.gov.au. 2018 [cited 13 April 2018]. Available from: http://www.dhhs.tas.gov.au/wihpw/principles/determinants\_of\_health