Dissecting the large field of studying health related issues, known as epidemiolo...

Science, Epidemiology

The World Health Organization defines epidemiology as “ the study of the distribution and determinants of health-related states or events (including disease), and the application of this study to the control of diseases and other health problems.

Epidemiology is a huge field containing several sub-fields such as disease etiology, outbreak investigation, disease surveillance, and comparisons of treatment effects. It is utilized often to not only discern the causes of disease, but as a key player in preventative healthcare. John Snow is commonly referred to as the father of modern epidemiology after researching the causes of cholera outbreaks in the 1800s and tracing the cause of a severe outbreak in London and eliminating the cause of water pollution. In the early 1900s, many more contributions were made including the use of statistical methods to test for correlation. Presently, epidemiology is a multi-faceted field, with several components.

Epidemiology is generally split into “ descriptive” and “ analytical” parts. The descriptive side deals with the “ who, what, where, and when” of disease occurrence and are the factors that are determined through careful observation. The people affected may share certain underlying factors such as race, gender, or age as well as occupation, socioeconomic status, and daily habits. The outbreak at hand also needs to be diligently monitored and described to see the effects and timeline of symptoms so that it can be tracked in others. Identifying the area in which the outbreak occurs is also vital, because the environment or an indigenous species may be the cause of transmission. Finally, knowing the time of the outbreak intertwines with the

previous factor, because certain areas are affected by seasonal changes differently than others. For example, monsoon season in tropical areas would be likely to bring flies, and if the flies carry malaria, more infections could occur during that time period in that area. A hypothesis would then be created and tested to see if the weather was indeed affecting the rate of malaria infections and preventative measures could be instituted before the next monsoon season.

Analytical epidemiology, meanwhile, is generally used to test the hypotheses generated by descriptive epidemiology. Analytical epidemiology is the experimental side and also features comparison groups, while its descriptive counterpart does not. All clinical and community trials, as well as randomized control trials, are part of analytical epidemiology as they all help to solidify and support or reject a hypothesis formed by the former. It is important to note that although a hypothesis may be strongly supported with data from several trials, correlation is not the same as causation. It is much more difficult to scientifically claim that one thing causes another, especially when it comes to linking behaviors to illnesses that are not pathogen-borne like diabetes, hypertension, or cancer. However, in order to potentially claim certain evidence as causal rather than merely correlating, a set of criteria known as the Bradford Hill criteria must be met and further testing and analysis is generally required. The Bradford Hill criteria consist of nine aspects, such as strength or consistency of the evidence in question (Hill, 1965).

Analytical epidemiologists utilize a number of different studies in order to test their hypotheses. The studies may vary on what they are testing for, whether it's incidence of heart attack before 50, side effects from a new pharmaceutical, or the rate of multiple drug resistant tuberculosis cases per year in a specific area. Other factors depend on time and funding, as some studies are significantly more expensive and time intensive than others.

Case-control studies are one of the more popular and widely used analytical methods. They utilize people who already have the affliction to be studied and are used to determine if an exposure is linked to the disease. The affected group is compared to a control group, which should be from the same population and hopefully well-distributed and large in size. Within these groups, individuals are identified as exposed or unexposed to a specific factor. An odds ratio is used to determine if the exposure can be deemed as having an effect on the acquisition of the disease, as well as the magnitude of the effect it may have. Using the table below, a statistic is calculated.

Odds Ratio =(exposed cases x unexposed cases)

If the odds ratio is less than one, the exposure may aid in protecting from the disease. If the odds ratio equals or is close to 1, exposure has no effect, and if it is greater than one, the chance of acquiring the disease is increased.

To be considered accurate, studies must have a certain amount of cases fitting the 95% confidence interval relating to the odds ratio. The confidence interval works hand-in-hand with the odds ratio, as it defines the precision of the odds ratio. The smaller the confidence interval, the more precise the

odds ratio is considered to be. Although not exactly the same as calculating the p-value at a 95% confident interval, the interval itself is also necessary. The p-value at the 95% confidence interval is used to determine statistical significance of the evidence and whether it supports or rejects the original hypothesis.

In case-control studies, the necessary number of cases at the confidence interval are related to the odds ratio, which can make it difficult because the data does not always fit nicely into that pattern, even if there is an obvious effect from the exposure. Case-control studies are comparatively quicker, less expensive, and simpler than other types of studies. They are particularly useful for investigating outbreaks and studying rare diseases or outcomes. However, they are also subject to more bias, such as recall bias, because each individual may not accurately remember if they have been exposed and how many times they were potentially exposed.

Cohort studies are another tool widely used in analytical epidemiology. The subjects selected for cohort studies are unaffected by the disease in question at the beginning of the study, however they should have been exposed and might develop the disease later in the study. These individuals are compared to controls, such as in the case-control studies. However, the controls are not to be exposed at all. Relative risk is calculated rather than the odds ratio, which is interpreted the same way statistically, but may have a greater power than the odds ratio as it can be determined that the actual risk of disease with the exposure is being calculated rather than the risk that may have come from being potentially exposed. Cohort studies are often more

useful because the exposure is already identified as being questionable, and the subjects are chosen because they have been exposed, rather than because they have a disease and might have been exposed, like in case-control studies. Another advantage is that multiple outcomes can be observed, rather than being limited to simply affected or unaffected. A few disadvantages to cohort studies include the need for a much larger sample size, the timeframe involved, as it can take many years for the disease to develop, and the cost that is increased by those factors.

Sources of bias are common in any broad study, but they may increase depending on the type of study. Validity is vital in statistics, and knowing the various kinds of bias that can occur may prevent their frequency. Two major sources of bias are selection bias and recall bias. The response difference and retention rate in the control and exposed groups may differ, and if not accounted for, the results may be skewed. Recall bias affects case-control studies more than cohort studies, but is still a major cause of misinformation. It occurs when subjects are asked if they have been exposed, because they may not recall it, even if they have, which alters which group they are placed into. This occurs because many individuals may not think it’s relevant at the time of exposure to remember the incidence and duration of exposure. A final way results could be skewed is seen when incidence of diseases seem to increase over time, when the explanation could very well be that the screening methods just improved and physician awareness increased. This occurred when investigating the incidence of gastrointestinal stromal tumors in Taiwan from 1998-2008, and better screening technologies were to blame.

Once accounted for, the data was normalized once again (Chiang et al. 2014).

Along with descriptive and analytical fields of epidemiology, new fields are emerging with today's scientific advances and technology. A form of epidemiology not often heard of is social epidemiology. Social epidemiology focuses on social issues like civil wars and famines in context of illnesses and death toll, using the social aspect in tandem with environmental and seasonal factors, because they often play a huge role in acquiring proper nutrition, shelter, and hygiene habits (Muntaner 2013). For example, a family who has been relocated due to civil strife will possibly only have the clothes on their back, and in times of desperation while escaping may drink contaminated water because there is no other choice, while someone of their demographics who isn't in a hostile situation would not be drinking dirty water and would not acquire the same illnesses. A rather new, innovative field is molecular pathological epidemiology, which focuses on integrating epidemiology with molecular and cellular biology to help understand the basis of several illnesses, including cancer, which are caused by changes to the genome and other cellular alterations (Ogino et al. 2012).

The epidemiology of cancer is essentially applying all the tools of epidemiology covered previously in reference to how cancer is spread and addressing possible preventative measures. Cancer is a very complex group of diseases, spawning from both exogenous and endogenous mutations, as well as the body itself hijacking its own healing systems to create masses proficient in gaining size and a power supply. Many people will be affected by

cancer either by being diagnosed at some point in their lifetime or by watching family and friends fight valiantly against it. If it wasn't enough to study the epidemiology of cancer based on how devastating it can be, it should be recognized that nearly every person will have contact with cancer personally or through a loved one and by volume alone, it deserves its own place in the epidemiologic field.

The two main forms of bias in this instance are lead time bias and length time bias. Lead time bias is essentially the idea that early diagnosis will falsely give the impression of survival statistics inflation, without really decreasing the mortality rate. Length time bias is the impression that slow growing tumors will be easier to screen and diagnose. However, often the diagnosis is unnecessary if the tumor is benign, and may cause unnecessary procedures and thus, side effects to the patient. A fear related to length time bias is over diagnosis. Many patients may have benign growths that are treated harshly as though they were malignant, for fear that they might become so one day. For example, an enlarged prostate, which can be caused by several harmless factors, will increase PSA, which is also a warning sign for prostate cancer. However, PSA alone should not diagnose prostate cancer, yet it might lead the patient to have an unnecessary procedure for further testing, or cause them mental anguish if they think it indicates cancer.

Some hormones do play a part in cancer formation, however. High levels of estrogen plus progestin are associated with increased breast cancer risk, but many women still need E+P treatment for menopause relief. To determine

whether the risk was increased or decreased with a pretreatment of E+P, a nested case-control study was conducted within The Women's Health Initiative randomized clinical trial of E+P therapy. The 1000+ women in the trial were followed for an average of 5. 6 years to see if they developed breast cancer. Then case and control subjects were compared for baseline hormone levels, and tested again a year later. It was found through this testing as well as further testing that women with lower pretreatment levels of estrogen were at an increased chance to develop breast cancer as opposed to their counterparts (Farhat et al. 2103).

Several observational epidemiological studies such as case-control studies and cohort studies, and the study listed above are often used to show associations between certain risk factors or behaviors and cancer. Some risk factors are uncontrollable, such as age or genetic mutations, but others such as smoking, eating an unhealthy diet, and acquiring certain STDs are preventable, which is why the epidemiology of cancer is so vital. As more preventable factors that are identified, people can change their lifestyles for the better or get screening for the uncontrollable aspects, but informing the public is the bridge between analysis and application.