The objective of this module is to understand the basic principles of ophthalmic epidemiology and learn how to apply them in practice. Application includes the ability to read and understand epidemiological studies and in some situations apply epidemiological knowledge to designing studies in your setting.
What Is Epidemiology?

- Epidemiology is the study of factors that affect the health and illness of populations.
- It is frequently used as an integral part of interventions conducted in the interest of public health and preventive medicine.
- It is especially useful in regard to evidence-based medicine in identifying risk factors for disease and determining optimal treatment approaches to clinical practice.
- Typically epidemiologists create study designs, collect data on communicable and non-communicable diseases in the field, develop disease models, and conduct statistical analysis.

Epidemiology is the science of factors that affect disease and health of populations and is applied to all kinds of life in addition to humans. Most commonly it is involved in public health interventions but is frequently associated with evidence-based medicine to identify risk factors for disease and help determine the best approaches to treatment of disease. The job description of an epidemiologist might be to help design studies, collect data on communicable and non-communicable diseases or conditions in the field, analyze the data, and construct disease models for other health care professionals to understand so that individuals can be better treated.
The prevalence of a condition or disease tell us how common or rare the condition or disease is. It is normally expressed as a percentage or numerator and denominator. The prevalence has to be defined in the context of a specific population and a time. For example, a study might find that the prevalence of cataract in Brazil during 2008 was 2% (200 per 10,000 individuals).
In order to determine the prevalence of a condition in a population, one needs a reliable sampling strategy; that is a method using a small enough sample that is representative of the larger population. In ophthalmology, the cluster method, which involves sampling the general population in many specific locations (the “clusters”) is common. Studies that determine prevalence are often called cross-sectional studies because they determine the prevalence of a condition at a single point in time. Another important factor is that the technique used to identify the condition of interest be reliable.
Very often we want to know whether an individual has a particular disease or condition. For example, if we screen individuals for glaucoma, we want to know if an individual really has glaucoma because if he or she does, we want to start treating them immediately. However, if our screening test indicates that some individuals have glaucoma, when in fact they do not, this is a problem because we would then start treating them for glaucoma when they do not have the disease. These issues form the basis of sensitivity and specificity for diagnostic tests.

**Sensitivity and Specificity**

- Sensitivity is the chance that if an individual has a given condition that the test will show he or she has the condition.
- Specificity is the chance that if an individual does not have a given condition that the test will show he or she does not have the condition.
Let us look at a table to see how to calculate sensitivity and specificity. In our glaucoma population of 100 individuals we have 67 who have glaucoma but our screening test data shows that we only pick up 78% of these. This is the sensitivity of the test (i.e. 52 out of 87). On the other hand, of the 33 individuals who do not have glaucoma, our screening test only rules out 25 of them. Therefore, our test has a specificity of 76% (i.e., 25 out of 33).
Two other parameters are commonly calculated in diagnostic tests that are influenced heavily by the prevalence of the disease or condition in the population being studied. These are the positive and negative predictive values. Thus the PPV is the proportion of individuals who have the disease or condition and are confirmed by the test as having the disease or condition. Likewise, the NPV is the proportion of individuals who do not have the disease or condition and are verified as not having it by the test. Note that the PPV decreases considerably when the prevalence of the condition is low but the NPV decreases when the prevalence is high.
Let us look at the same table to see how to calculate PPV and NPV. The PPV is calculated by taking the proportion of individuals who have glaucoma and test positive out of all the individuals who test positive for glaucoma in our screening test. Likewise, the NPV is calculated by taking the proportion of individuals who do not have glaucoma and test negative out of all the individuals who test negative for glaucoma.

<table>
<thead>
<tr>
<th></th>
<th>Disease Present</th>
<th>Disease Absent</th>
<th>Total</th>
<th>PPV</th>
<th>NPV</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Test Positive</strong></td>
<td>True Positive (TP) 52</td>
<td>False Positive (FP) 8</td>
<td>Total=60</td>
<td>PPV=52/60 (87%)</td>
<td>NPV=25/40 (63%)</td>
</tr>
<tr>
<td><strong>Test Negative</strong></td>
<td>False Negative (FN) 15</td>
<td>True Negative (TN) 25</td>
<td>Total=40</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Incidence

- Incidence is a way of looking at the change in prevalence of a condition or disease over time.
- It is often expressed as the number of new cases over a period of time.
- For example, in Argentina, there might have been 9,000 new cases of macular degeneration diagnosed in population over 60 years old in 2007.

Incidence is a method of expressing the change in prevalence over a period of time. So, if there were 9,000 new cases of macular degeneration reported in Argentina in 2007 in the over-60 years age group, and we assumed a population count of 5 million, then the change in prevalence of the disease from the end of 2006 to the beginning of 2008 would be 9,000/5 million, or 0.18%.
A complementary study to the cross-sectional design is called a cohort or longitudinal study. Typically a group of individuals, (called the “cohort”) is followed for a specific period of time to learn how a particular condition or disease changes in respect of time. For example, we might want to assess how the outcomes of cataract surgery change over a period of 10 years due to other ocular comorbidities that develop. Longitudinal studies are expensive and difficult to carry out the longer the study runs. A 1-year study might be relatively easy, but a 10-year study would likely suffer from high follow-up losses, and would need staff to keep track of individuals for 10 years and administer tests after 10 years.
Case-control Studies

- Case control studies are a common design in epidemiology
- They are typically retrospective studies of groups of individuals, one of which had the outcome of interest and one which did not.
- What factors differentiate the cohorts?
- Individuals can be matched between cohorts to try and eliminate common possible factors (e.g., age, gender).

The case-control study is an inexpensive retrospective investigation to look at an outcome and try to determine what factors caused the outcome. The study begins with definition of the cases (e.g., those with newly acquired diabetic retinopathy) and investigators identify suitable people for this group and another group of people who do not have diabetic retinopathy. Based on personal recall of individuals and medical and other records, researchers look for factors that might be involved in the development of diabetic retinopathy.
Epidemiology is frequently used to identify risk factors for a given condition or disease. Case-control studies can provide odds ratios (ORs) for given factors. Cohort studies can then follow groups of individuals and look at incidence of diseases or conditions of interest and correlate outcomes with risk factors of interest.

Epidemiology is frequently used to identify risk factors that predispose an individual to a given disease or condition. For example, a case-control study of age-related macular degeneration might try to uncover which nutritional factors could be involved in the development of macular degeneration as an individual ages. If prominent risk factors are identified, a more expensive cohort study can then be instituted to follow groups of individuals as they age and see if these nutritional factors play a part in the development of macular degeneration. Finally, if the evidence is strong enough, controlled clinical trials can be conducted by actively giving specific nutritional factors to find out if these factors actually slow down the incidence of macular degeneration.
Odds Ratios

• In epidemiology we often use the odds ratio to quantify the risk of an event that might occur in one group compared to another.
• Suppose that in a sample of 1,000 Hispanic people, 98 have been diagnosed with uncorrected refractive errors in the last year but in same numerical sample of Asian people, 175 have been diagnosed with uncorrected refractive errors.
• The odds of having this condition in each group is 98/902 (0.109) and 175/825 (0.212), respectively. The odds ratio (OR) is 0.212/0.109 or 1.94.

Odds ratios are often used to determine the risk of an event occurring compared to a control group. Let us suppose we are studying 1,000 Hispanic individuals of whom 98 have been diagnosed with uncorrected refractive errors in the last year. We wish to compare these people to 1,000 Asian individuals of whom 175 have been diagnosed with uncorrected refractive errors. We first calculate the odds of having the condition in each group and then using the lowest odds for the denominator, calculate the odds ratio, which is this case is 1.94.
Odds ratios can be difficult to understand, which is why relative risk (RR) is often used. In calculating the risk of having uncorrected refractive errors in each group we use the number of individuals with the condition divided by the total number in each group. The relative risk is then calculated like the odds ratio, using the lowest risk as the denominator. Note that the odds ratio is higher than the relative risk in our example. Part of the reason for the popularity of odds ratios is that the regression options used in most statistical programs provide odds ratios only. The relative risk and associated confidence intervals are much harder to estimate from regression.
Differences between the calculated odds ratios and relative risks vary according to the prevalence of the disease or condition and the exposure to it. In this simulation the prevalence of exposure has been fixed at 0.5. This might be equivalent to half the general population being at risk for glaucoma. The prevalence of the disease has been varied from 5% to 45%. Note first that when the relative risk is 1, so is the odds ratio of having the disease. Second, when the relative risk is less than 1, the odds ratio varies, with the largest differences at the highest prevalence of the disease. Likewise, when the relative risk is greater than 1, the largest differences in odds ratios are observed when the prevalence of disease is the highest. Thus, when the relative risk is very high or low, the odds ratio appears to be exaggerated.

Based on Zocchetti et al, Int J Epidemiol 1997;26:220-3