# Genetic epidemiology

1- *Definition*

Genetic Epidemiology is the study of how genetic factors contribute to health and disease in families and populations, and how genes interplay with environmental factors.

Traditionally, the study of the role of genetics in disease progresses through the following study designs, each answering a slightly different question:

Familial aggregation studies: Is there a genetic component to the disease, and what are the relative contributions of genes and environment?

Segregation studies: What is the pattern of inheritance of the disease (e.g. dominant or recessive)?

Linkage studies: On which part of which chromosome is the disease gene located?

Association studies: Which allele of which gene is associated with the disease?

This traditional approach has proved highly successful in identifying monogenic disorders and locating the genes responsible.

More recently, the scope of genetic epidemiology has expanded to include common diseases for which many genes each make a smaller contribution (polygenic, multifactorial or multigenic disorders). This has developed rapidly in the first decade of the 21st century following completion of the Human Genome Project, as advances in genotyping technology and associated reductions in cost has made it feasible to conduct large-scale genome-wide association studies that genotype many thousands of single nucleotide polymorphisms in thousands of individuals. These have led to the discovery of many genetic polymorphisms that influence the risk of developing many common diseases. The genetic epidemiology can also be skewed by the presence of evolutionary pressures that induce negative selection during molecular evolution. This negative selection can be determined by tracking the skewness of the distribution of mutations with putatively severe effects as compared to the distribution of mutations with putatively mild or absent effect.

2- *Approaches*

Genetic epidemiological research follows 3 discreet steps, as outlined by M.Tevfik Dorak:

a- Establishing that there is a genetic component to the disorder.

b- Establishing the relative size of that genetic effect in relation to other sources of variation in disease risk (environmental effects such as intrauterine environment, physical and chemical effects as well as behavioral and social aspects).

c- Identifying the gene(s) responsible for the genetic component.

These research methodologies can be assessed through either family or population studies.

3- *Genetic and cancer (adapted from [https://www.cancer.gov/about-cancer/causes-prevention/genetics](https://www.cancer.gov/about-cancer/causes-prevention/genetics))*

Cancer is a genetic disease. It is caused by changes in genes that control the way cells grow and multiply. Scientists have found hundreds of DNA and genetic changes (also called variants, mutations, or alterations) that help cancer form, grow, and spread.

Cancer-related genetic changes can occur because:

a- they are random mistakes in our DNA happen as our cells multiply our DNA is altered by carcinogens in our environment, such as chemicals in tobacco smoke, UV rays from the sun, and the human papillomavirus (HPV)

b- they were inherited from one of our parents

DNA changes, whether caused by a random mistake or by a carcinogen, can happen throughout our lives and even in the womb. While most genetic changes aren't harmful on their own, an accumulation of genetic changes over many years can turn healthy cells into cancerous cells. The vast majority of cancers occur by chance as a result of this process over time.

3.1- Is cancer hereditary?

Cancer itself can't be passed down from parents to children. And genetic changes in tumor cells can't be passed down. But a genetic change that increases the risk of cancer can be passed down (inherited) if it is present in a parent's egg or sperm cells.

For example, if a parent passes a mutated BRCA1 or BRCA2 gene to their child, the child will have a much higher risk of developing breast and several other cancers.

That's why cancer sometimes appears to run in families. Up to 10% of all cancers may be caused by inherited genetic changes.

Inheriting a cancer-related genetic change doesn't mean you will definitely get cancer. It means that your risk of getting cancer is increased.

3.2- What is a family cancer syndrome?

A family cancer syndrome, also called a hereditary cancer syndrome, is a rare disorder in which family members have a higher-than-average risk of developing a certain type or types of cancer. Family cancer syndromes are caused by inherited genetic variants in certain cancer-related genes.

With some family cancer syndromes, people tend to develop cancer at an early age or have other noncancer health conditions.

For example, familial adenomatous polyposis (FAP) is a family cancer syndrome caused by certain inherited changes in the APC gene. People with FAP have a very high chance of developing colorectal cancer at an early age and are also at risk of developing other kinds of cancer.

But not all cancers that appear to “run in families” are caused by family cancer syndromes. A shared environment or habits, such as exposure to air pollution or tobacco use, may cause the same kind of cancer to develop among family members.

Also, multiple family members may develop common cancers, such as prostate cancer, just by chance. Cancer can also run in a family if family members have a combination of many genetic variants that each have a very small cancer risk.

3.3- Should I get genetic testing for cancer risk?

Certain genetic tests can show if you’ve inherited a genetic change that increases your risk of cancer. This testing is usually done with a small sample of blood, but it can sometimes be done with saliva, cells from inside the cheek, or skin cells.
Genetic tests can help families with a history of breast and ovarian cancer make screening and treatment decisions.

Not everyone needs to get genetic testing for cancer risk. Your doctor or health care provider can help you decide if you should get tested for genetic changes that increase cancer risk. They will likely ask if you have certain patterns in your personal or family medical history, such as cancer at an unusually young age or several relatives with the same kind of cancer.

3.4- How do genetic changes cause cancer?

Genetic changes can lead to cancer if they alter the way your cells grow and spread. For example, some DNA changes raise the levels of proteins that tell cells to keep growing. Other DNA changes lower the levels of proteins that tell cells when to stop growing. And some DNA changes stop proteins that tell cells to self-destruct when they are damaged.

For a healthy cell to turn cancerous, scientists think that more than one DNA change has to occur. People who have inherited a cancer-related genetic change need fewer additional changes to develop cancer. However, they may never develop these changes or get cancer.

As cancer cells divide, they acquire more DNA changes over time. Two cancer cells in the same tumor can have different DNA changes. In addition, every person with cancer has a unique combination of DNA changes in their cancer.

3.5- What kinds of genetic changes cause cancer?
Fusion proteins, which can occur when parts of different chromosomal regions are joined, may drive the development of many cancers in children.

Mutation in the DNA code. Some variants affect just one nucleotide which could be missing or replaced by another nucleotide. These are called point mutations.

For example, around 5% of people with cancer have a point mutation in the KRAS gene that replaces G with A. This single point mutation creates an abnormal KRAS protein that constantly tells cells to grow.

Cancer-causing genetic changes can also occur when segments of DNA—sometimes very large ones—are rearranged, deleted, or copied. These are called chromosomal rearrangements.

For example, most chronic myelogenous leukemias (a type of blood cancer) are caused by a chromosomal rearrangement that places part of the BCR gene next to the ABL gene. This rearrangement creates an abnormal protein, called BCR-ABL, that makes leukemia cells grow out of control.

Some cancer-causing DNA changes occur outside genes, in sections of DNA that act like “on” or “off” switches for nearby genes. For example, some brain cancer cells have multiple copies of “on” switches next to genes that drive cell growth.

Other DNA changes, known as epigenetic changes, can also cause cancer. Unlike genetic variants, epigenetic changes (sometimes called epimutations) may be reversible and they don’t affect the DNA code. Instead, epigenetic changes affect how DNA is packed into the nucleus. By changing how DNA is packaged, epigenetic changes can alter how much protein a gene makes.

Some substances and chemicals in the environment that cause genetic changes can also cause epigenetic changes, such as tobacco smoke, heavy metals like cadmium, and viruses like Epstein-Barr virus.

*4- Statistical analyses*

Analysis of genetic epidemiology needs statistical methods. Here are some definitions to help you navigate the interpretation of the data.

**Linkage**, as related to genetics and genomics, refers to the closeness of genes or other DNA sequences to one another on the same chromosome. The closer two genes or sequences are to each other on a chromosome, the greater the probability that they will be inherited together.

A **confidence interval** is the mean of your estimate plus and minus the variation in that estimate. This is the range of values you expect your estimate to fall between if you redo your test, within a certain level of confidence. Confidence, in statistics, is another way to describe probability.

**LOD score** is a statistical estimate of whether two genetic loci are physically near enough to each other (or "linked") on a particular chromosome that they are likely to be inherited together. A LOD score of 3 or higher is generally understood to mean that two genes are located close to each other on the chromosome. In terms of significance, a LOD score of 3 means the odds are 1,000:1 that the two genes are linked and therefore inherited together. Also called logarithm of the odds score.

**Penetrance** is the extent to which a particular gene or set of genes is expressed in the phenotypes of individuals carrying it, measured by the proportion of carriers showing the characteristic phenotype.

**Multipoint linkage analysis** is commonly used to evaluate linkage of a disease to multiple markers in a small region. Multipoint analysis is particularly powerful when the IBD relations of family members at the trait locus are ambiguous.