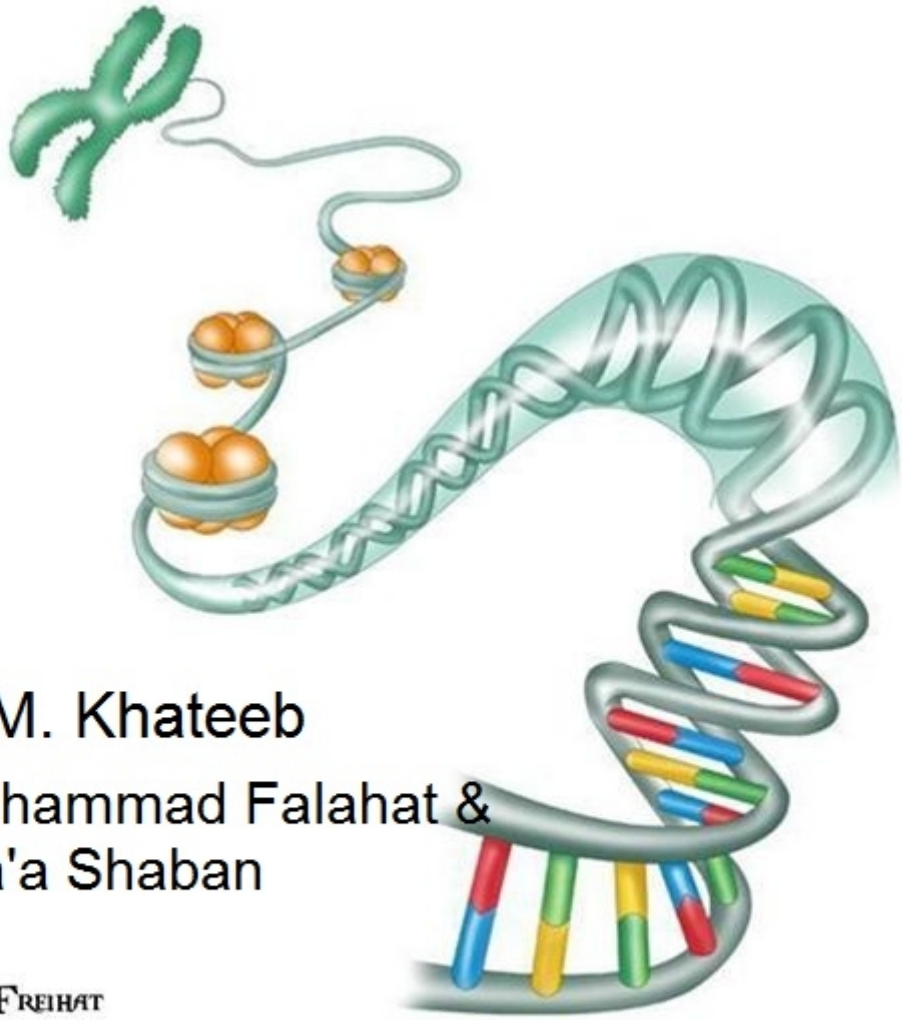
University of Jordan
Faculty of Medicine
Batch 2013-2019

Medical Committee
The University of Jordan

Genetics & Molecular Biology

○ Slides ● Sheet ○ Handout ○ other.....



Sheet#: 8

Dr. Name: M. Khateeb

Done By: Mohammad Falahat &
Ala'a Shaban

Designed by Nadeen Al-Freihat

Multifactorial diseases

Genetic diseases are classified into: Unifactorial (e.g.: in single gene inheritance), Chromosomal, and Multifactorial.

We talked about unifactorial diseases before. Now we'll talk about multifactorial diseases.

A comparison between Unifactorial and Multifactorial diseases:

Unifactorial (Single gene)	Multifactorial
Low prevalence in the population	High prevalence in the population
High penetrance	Low penetrance
An example: Thalassemia - It's incidence in Jordan is 1:2500 (Low prevalence) - The clinical picture: If a person has Thalassemia, phenotypic changes appear immediately. (High penetrance)	**An example: Diabetes** - It's prevalence in the population aged 25 years and older is about 17% (High prevalence) - The clinical picture: If a person has diabetes, phenotypic changes and complications start after 10-15 years. (Low penetrance)
Due to a mutation	Due to a polymorphism The Dr. said: "You remember what is the meaning of polymorphism, it is NOT a single area where we can have these types of characteristics."

Multifactorial genetic diseases:

Combinations of a single gene, multiple genes, and the environmental conditions that interfere with those genes' expression. In this type of diseases, a gene rarely acts completely alone; environmental factors affect and modify the expression of that gene.

Examples on the contribution of genes or environment or both:

- Mendelian: due to a single gene
- Polygenic: due to two or more genes
- Multifactorial: in addition to the polygenes, environmental factors also interact.
  An example on multifactorial genes: **height of the person**
  It depends on many genes (two, three, or more). The expression of these genes is affected by nutrition, environment in utero, or outside utero. Abnormalities in these factors can affect the function of these genes.
- Complex: where the relative contributions of genes and environment are not yet established.
- Polygenic inheritance:

  More than one gene inheritance, the effect of these genes can cause one clinical picture or one phenotypic trait. Generally, there is an **additive** effect in the function of these genes. **Additive** implies that the effects of the genes are cumulative, i.e. no one gene is dominant or recessive to another. (for example: gene 1 will act 5%, gene 2 will act 10%, and so on for the other genes)

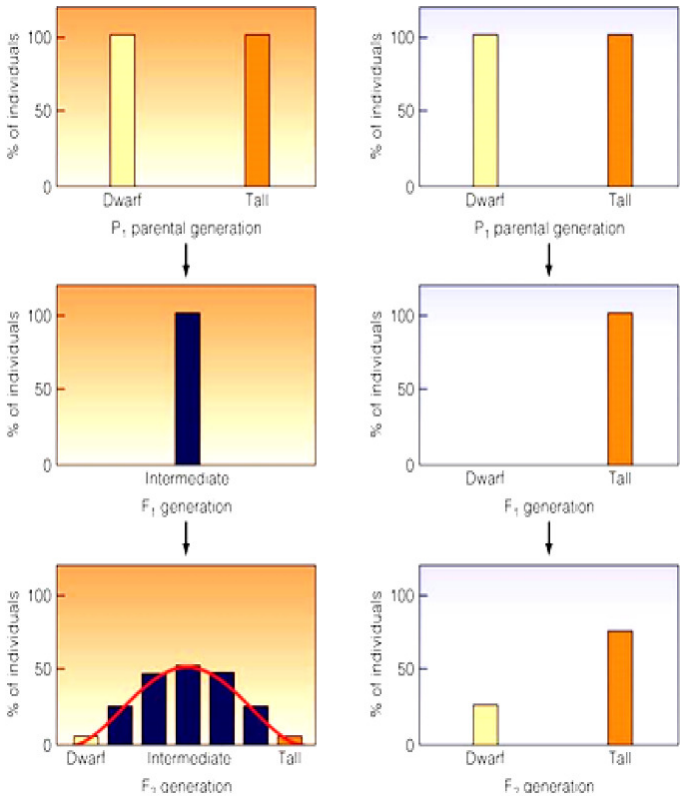
  This additive effect can't be seen in single gene dominance.

Polygenes and single genes:

When we talk about single genes, a single gene is either present or absent (and hence the phenotype), this is called **discontinuous** type of inheritance. If the gene exists, its clinical picture exists too (All or none).

In polygenes, multiple genes control the trait, an extra gene adds an effect to the phenotype. Hence, the phenotype differences are more variable and distributed in the population. This is called **continuous** type of inheritance, as we can see an array of phenotypes (scale).

As seen in the diagrams, in single gene inheritance, the gene either exists or not, and hence the phenotypic trait, in peas for example, the height is controlled by a single gene, in mendelian dominance, thus it's either tall or dwarf (right diagrams). However, in polygenes, as more genes are added, the phenotypic trait is more prominent, like the height of human, it's controlled by multiple genes, thus more (tall) genes means a taller person, producing a scale of heights ranging from dwarf to tall (left diagram).

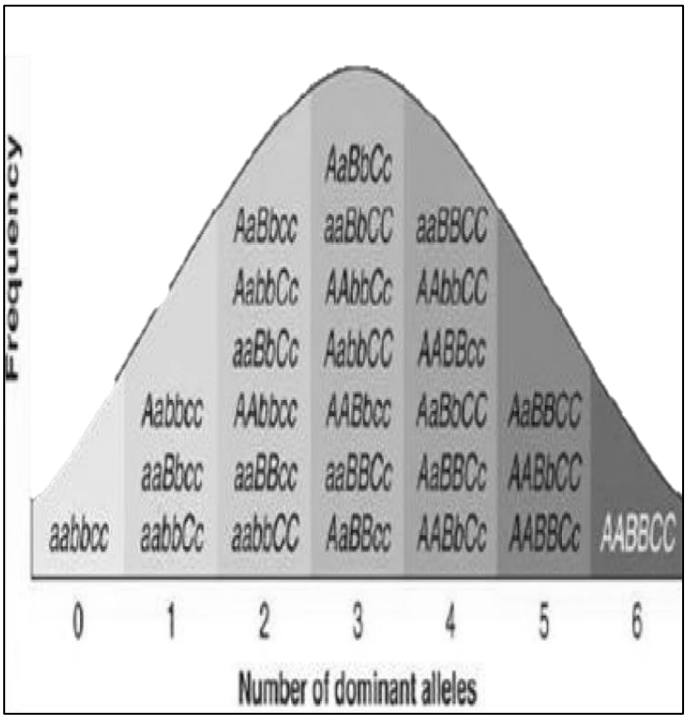


Other examples on **polygenes** are:

Hair color, Eye color, Skin color, Body weight, Intelligence (the following mechanism applies to any of them). These traits can be measured and we can see the prevalence of each of their phenotypes.

Skin color:

Three pairs of genes at least control the skin color, (A, B, C – dark / a, b, c – light), the lightest skin color must have the genotype (aabbcc), while the darkest skin color must have (AABBCC), an intermediate skin color will have any of these genotypes (AABbcc, AAbbCc, AaBbCc, AaBBcc, AabbCC, aaBBCc, aaBbCC), and any other degree of skin color on the scale between light and dark can exist depending on the number of dark and light genes. These genetic differences are distributed in the population. As seen in the diagram.



When can we suspect that a disease is multifactorial or not?

- The first characteristic is that the condition is **relatively common** in the population (not pale), like hypertension, cholesterol abnormalities, and diabetes.
  It's incidence in relatives is lower than single gene diseases.
  (These conditions show a definite familial tendency, but the incidence in close relatives of affected individuals is much lower than would be seen if these conditions were caused by mutations in single genes.).

> Don't mix:
>
> In the population: Multifactorial diseases are more prevalent.
>
> Among relatives: Single gene diseases are more prevalent (if a mutation exists).

- The risk that siblings may have the disease is similar to the risk that children may have it. But it's **less** in the other second and third degree relatives. The incidence in siblings and children is relatively high, but it's lower in cousins (the incidence drops as the relation is further).
- The risk related to index case (the first person identified having the disease). If there is a very severe case in a family, the possibility to have a second case is **more**.
- The number of affected people in the family. If two children have the disease, the possibility that a third child will have it is **more** than if only one child had it.

> For your information:
>
> First degree: parents, siblings, children
>
> Second degree: uncles and aunts, nephews and nieces, Grandparents & grandchildren
>
> Third degree: First cousins, Great grandparents & great grandchildren

Examples of multifactorial diseases:

- Congenital malformations:
  Heart problems, neural tube defects, cleft lip/palate, pyloric stenosis, and congenital hip dysplasia.

- Common non-communicable acquired diseases:
  Asthma, schizophrenia, diabetes mellitus, hypertension, coronary heart diseases, and familial hypercholesterolemia.

Frequency of different types of genetic diseases:

The prevalence of multifactorial diseases is **very much higher** in the population than any other type of genetic diseases.

Type	Incidence at Birth (per 1,000)	Prevalence at Age 25 Years (per 1,000)	Population Prevalence (per 1,000)
Diseases due to genome/chromosome mutations	6	1.8	3.8
Disease due to single gene mutations	10	3.6	20
Disease with multifactorial inheritance	~50	~50	~600

(In "diseases due to genome/chromosome mutations", the prevalence drops in adults because it's mainly a problem in infancy)

How can we study multifactorial abnormalities?

Using: Threshold model, Liability, Recurrence risk.

- **Threshold in a population**: a measure of how many people in the population have the disease (for example diabetes, hypercholesterolemia...)
- **Liability**: how many people from that population have risk factors for the disease (genetic, environmental, or both) (measures the possibility of getting the disease, the liable people phenotypically don't have the disease).
- **Recurrence risk**: estimates the risk that the disease will recur in more individuals.

A definition for Liability from the book:

(Emery's Elements of Medical Genetics 14 ed)

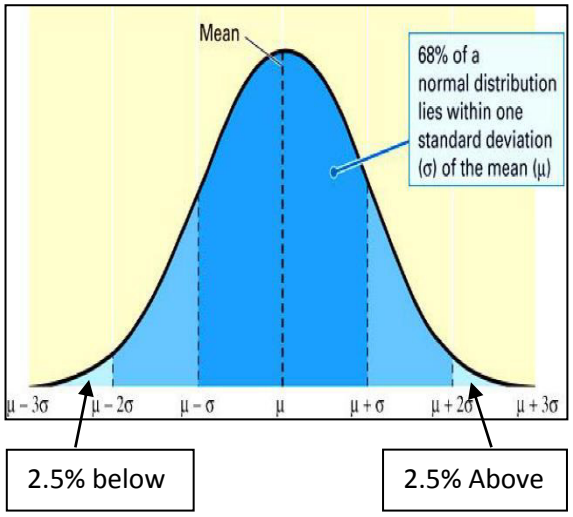
The liability includes all factors that contribute to the cause of the condition.

Looked at very simply, a deleterious liability can be viewed as consisting of a combination of several 'bad' genes and adverse environmental factors. Liability cannot be measured but the mean liability of a group can be determined from the incidence of the disease in that group using statistics of the normal distribution.

Multifactorial abnormalities have population distribution curves, how can we analyze them and why do we use them?

(Definition of Threshold)

Generally, any normal distribution curve has a standard deviation, for example if cholesterol levels are measured for a group of people, and the mean is 200 mg/dL, the standard deviation expresses how the values are distributed below and above the mean.

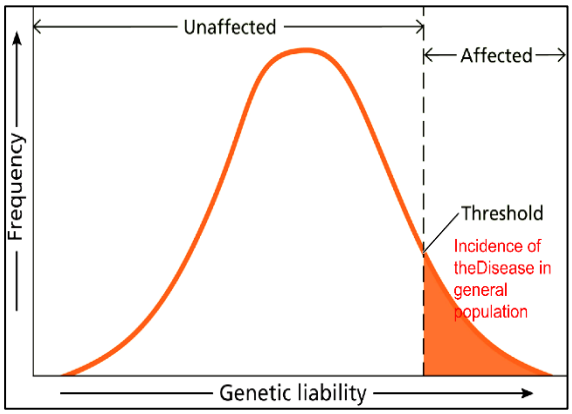


The proportion of the population in a selected area under that curve depends on how many standard deviations are taken above and below the mean in that area, one, two, or three standard deviations.

This depends on the sensitivity of the test that is done. Generally, if one standard deviation above and below the mean is taken, only around 68% of the population is included in that area, if two standard deviations are taken above and below the mean, around 95% of the population is included, if three standard deviations are taken above and below the mean, almost everyone (99.7%) of the population is included.

Generally, in all the studies we take 95% of the population which means two standard deviations above and below the mean, there is a reason after that; if a person does a test and the test is repeated by another person, we expect around 5% difference, which is called **personal difference**, so always there is 2.5% difference above and 2.5% difference below the mean.

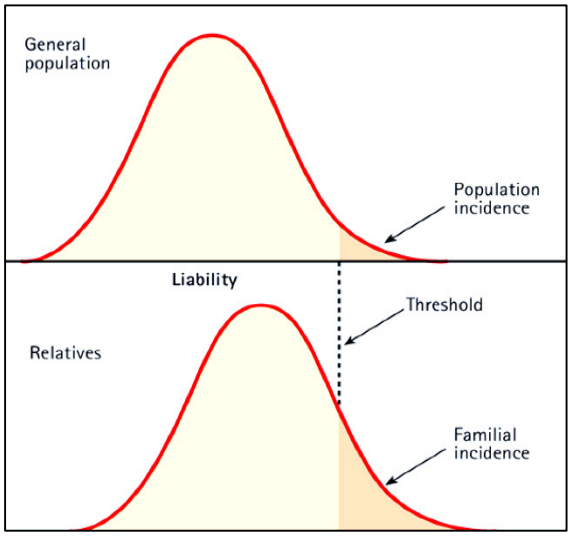
The threshold of a multifactorial disease represents the amount of people that have the disease. If the cholesterol levels are measured for a group of people and represented in the first diagram, considering people with cholesterol levels higher than 200 mg/dL as having hypercholesterolemia, then people after the threshold have cholesterol levels higher than 200 mg/dL.



To find if there is any familial tendency for any of these conditions, we take those people who are found in the threshold group (index case; who already have hypercholesterolemia) and examine their first degree relatives (parents, siblings, children), and see how many of them have the disease, and put the results in a new diagram. If the threshold shifts to the left, this means that the disease has a familial tendency, and we get a diagram similar to the second one in the second figure. If there is no movement to the left, this means that there is no familial tendency and mainly this is an environmental condition.

To prove that a disease has a familial tendency, we increase the relation degree to more relatives to include second and third degree relatives and study them again, if the threshold goes back to the right and the curve returns to the normal state (of population), this proves that the disease has a familial tendency. (See the text box)

Also, we can study the gender, among the people above the threshold (who have the disease), if there are more females, there is more tendency in females, and if there are more males, there is more tendency in males.



Note from the writer:

I think that the curve is shifted to the right and the threshold is fixed, look at the next definition (from the book) to understand why.

Threshold is the margin above which the abnormal phenotype is expressed (I think that threshold on the diagram is the liability (amount of risk factors) above which the disease appears)

So that, when doing the test on first degree relatives, if the disease has a familial tendency (relatives have higher incidence for the disease), more people will have the disease, and the curve is shifted to the right so that more people will lie in the area above the threshold.

Liability:

Quantitative trait that is prevalent in a group of people at a genetic risk to the disease, and the individuals that have liabilities above the threshold will develop the disease.

Heritability

The proportion of the etiology of the disease that can be ascribed (attributed) to genetic factors as opposed to environmental factors. The proportion of the total phenotypic variance of a condition that is caused by the additive genetic variance. So that, the greater the value for the heritability the greater the role of genetic factors. (From the book)

How much is the heritability value of a disease? It's the proportion of the disease that is due to genetic causes (for example: if 50% of the condition is due to genetic effect, 30% is due to genetic effect, or 60% is due to genetic effect, it's called heritability of that condition).

From this we can look for genetic risk factors of diabetes, Alzheimer disease, alcoholism, obesity, and all other diseases.

- How heritability can be measured:

Since we are talking about multifactorial, that genetic background is the summation of the least genes affecting the phenotypic characteristics. If they are three genes, we take those three genes and see their activity.

Empiric risk:

Measures the likelihood that a condition will occur based on an incidence, (in the family, this incidence increases if more previous children are affected). Meaning that if the incidence is 2%, how much that could be recurrent, or if it is 5%, how much that could be recurrent in this population.

Don't forget: there is a difference between prevalence and incidence. Prevalence is the amount of people who have the disease at a certain point of time. Incidence is the difference in the amount of people who have the disease between two points of time (the amount of new cases that get the disease in unit time).

- There are four conditions of the empiric risk:

1. The incidence of the condition in general is greatest among relatives of the most severely affected patients. If a case in a family is diagnosed with a very severe condition, there is an empiric risk that one, two, or three members of that family will get the disease.
2. The recurrence risk increases with increasing the number of previously affected children. More affected children means a higher risk that a new child will have the disease (higher recurrence risk).
3. The risk is greatest among close relatives of the index case and decreases rapidly in more distant relatives.
4. If the condition is more common in individuals of one particular sex (more females or more males are affected), recurrence risk varies according to the sex of the index case.

- An example on the empiric risk: for a certain disease:

  In identical twins, the empiric risk is 40%

  In siblings, the empiric risk is almost 10% (in two brothers or two sisters)

  In the child it's less, as the relation is more distant the empiric risk is even less.

  In the general population it's around 0.5% (All the relatives have higher empiric risks than the general population).

The frequency of a disease in relatives vary depending on the sex. (Male relatives of a male patient) have different risk from (female relatives of a male patient) and different risk from that of (male relatives of a female patient) and from (female relatives of a female patient). (The sex is important in making the risk higher or lower).

- Some multifactorial conditions have an unequal sex ratio, for example:

Condition	Sex ratio (males to females)
Pyloric stenosis	5 to 1
Congenital dislocation	1 to 6
Rheumatoid arthritis	1 to 3
Peptic ulcer	2 to 1

(The prevalence of these conditions is different in males and females)

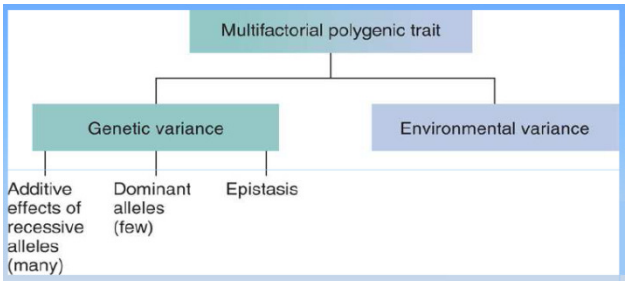
We can calculate the risk of a certain disease in the population using a (2 X 2) table (we took it before, the Dr. said).

How can we analyze a multifactorial disease?

Analysis of multifactorial diseases is not easy; it is a kind difficult to be done. *Human genome sequencing*, *population studies* and *family studies* can give some pieces of information for analyzing the disease; e.g. the frequency of the disease in the population give us the empiric risk, calculating the *heritability* of a trait can estimate the degree of phenotypic variation due to genetic variances in a specific population.

Heriditary is estimated from the proportion of people sharing a trait to the proportion predicted genetically to share that trait.

Among individuals there are genetic variances and environmental variances affecting the presence of a certain multifactorial trait, one way of analyzing a multifactorial disease is by **stabilizing** one of these factors ***(Comparisons between and within families)***, e.g. we can take twins; *monozygotic twins* have the same genetic background in addition to their shared environment in the uterus, *dizygotic twins* have the same uterine environment but they are genetically different (they share 50% of their genes). On the other hand, if we have identical twins (monozygotic) but they have been *raised apart* from each other; their genetics will be the same but the environmental factor is changed here. And the last example to be mentioned here in separating the genetic and environmental factors is the *adopted individuals* who have been put in the same fixed environment but they are genetically different.



Another way for analyzing a multifactorial disease is by doing ***association studies;*** comparing single nucleotide polymorphism patterns between affected and unaffected groups, identify important DNA regions.

** Note: not all the diseases have the same heritability, e.g.: In Schizophrenia 85% cases belong to heritability conditions, while 15%

Disorder	Frequency (%)	Heritability
Schizophrenia	1	85
Asthma	4	80
Cleft Lip = Cleft palate	0.1	76
pylonc stenosis	0.3	75
Ankylosingspondylitis	0,2	70
Club foot .	0.1	68
Coronaryartery disease	3	65
Hypertension (essential)	5	62
Congenital dislocction of the hip	0.1	60
Anencephaly and spina pifida	0.1	60
Peptic Ulcer	4	37
Congenital Heart Disease	0.5	35

only belong to environmental factors. You can look at the percentages in the first figure and notice that diabetes is the least one that belongs to genetic factors. Prevalence and heritability are independent from each other; having a high heritability does not necessarily mean having a high prevalence, e.g. Schizophrenia has a frequency of 1% and a heritability of 85%, while peptic ulcer has a prevalence of 4% and a heritability of 37%. *The doctor in the lecture mentioned the numbers related to Schizophrenia, Spina bifida, Ischemic heart disease, diabetes mellitus, ankylosing spondylitis and peptic ulcer.

Concordance: the percentage of pairs in which both twins express the trait; i.e. if both twins share the same disease like cleft palate, we say it is a concordant disease, but if they do not share the same trait, we say there is no concordance. Concordance is used to determine the heritability (genetic factors).

Monozygotic (MZ) twins have more concordance than do *dizygotic (DZ)* twins have.

	Concordance Values (%)	
Trait	**MZ Twins**	**DZ Twins**
Blood types	100	66
Eye color	99	28
Mental retardation	97	37
Hair color	89	22
Down syndrome	89	7
Handedness (left or right)	79	77
Epilepsy	72	15
Diabetes	65	18
Tuberculosis	56	22
Cleft lip	42	5

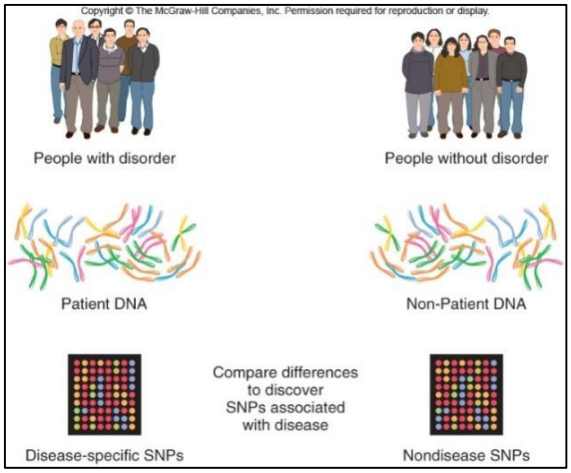
** Even monozygotic twins could have variability in concordance due to different environments (if they have the same amniotic sac or not, same placenta or not, mitochondria are not distributed evenly, or if they have been raised apart or not…).

The second way of analyzing multifactorial diseases as we mentioned is *association studies.* Association studies are undertaken by comparing the frequency of a particular variant in affected patients with its frequency in a carefully matched control group. This approach is often described as a **case-control** study. If the frequencies in the two groups differ significantly, this provides evidence for an association.

The polymorphic HLA histocompatibility complex on chromosome 6 has been frequently studied.

E.g. association studies about diabetes type 1 (you will be asked about it in the exam), we find that diabetes type 1 is associated with HLA DR4 (17% with diabetes type 1, 7% controls). This is an association not heritability. We take groups and control groups then compare between them by applying two by two table.

Another example is ankylosing spondilitis, (FYI: a chronic inflammatory disease of the axial skeleton, with variable involvement of peripheral joints and nonarticular structures.) one of the strongest HLA association known is that between ankylosing spondylitis and the B27 allele. People with this allele have a really high rate of getting the disease compared to those who don't have this allele.



Genome-wide association studies seek **SNP's** (single nucleotide polymorphism) that are shared with much greater frequency among individuals with the same trait than among others, we can for example take one hundred patients of type 2 diabetes and one hundred healthy persons then we can do gene sequencing for both groups looking for polymorphisms, if we find in diabetic patients for example fifty, one hundred, two hundred… different types of polymorphisms, we can say those are associated with diabetes, the other pleomorphisms are not. We did the same thing with prostate cancer starting with 1000 SNP, for example, ending up with thirty to fifty SNP's, so people who do have these polymorphisms are more susceptible to get the disease.

Conclusions:

- Multifactorial disorders are more common than single gene and chromosomal disorders.
- They are caused by the interaction of many genes with environmental factors.
- Optimum preventive measures rely on avoidance of the bad environmental factors since avoidance of inheriting the bad genes is at present not possible.
- These measures can be explained through counseling such as preconception and chronic non-communicable diseases counseling.

How multifactorial diseases can be reflected on the population

Firstly, population is an interbreeding group of the same species within a given geographical area. For example: taking Amman as a population, Jabal Al-Hussein would be a subpopulation, Street of Jordanian University would be a local population. **Gene pool** means the whole genes in a certain population. **Gene flow** means the passage of genes from one generation to another. Gene pool can be affected by:

1. Selection; it can increase or decrease the genetic diversity (positive or negative effect).
2. Mutation; the disease will be prevalent in that population increasing the genetic diversity (Positive effect).
3. Migration; it could increase or decrease the genetic diversity (Positive or negative effect).
4. Genetic drift; it can decrease the genetic diversity (negative effect).
5. Non-random mating; e.g. mating between relatives only, you're concentrating certain genes and deleting others (negative effect).

Secondly, discussing population genetics; we discuss genotype frequency and allele frequency, they are different; gene frequency talks about phenotypic characteristics (clinical picture we see); so talking about genotype frequency means we are talking about the two alleles, while allele frequency identifies the frequency of a certain allele (dominant allele, recessive allele …) in the population. For certain diseases allele frequency and gene frequency are not the same in all population, e.g. taking PKU frequency in different populations, we find it is very rare in Chinese population and even rarer in Japanese population, but it is more prevalent in Turks.

Population	Frequency of PKU
Chinese	1/16,000
Irish, Scottish, Yemenite Jews	1/5,000
Japanese	1/119,000
Swedes	1/30,000
Turks	1/2,600
United States Caucasians	1/10,000

Hardy-Weinberg Theory states that allele and genotype frequencies in a population will remain constant from generation to generation in the absence of other evolutionary influences (these mentioned below)

Hardy-Weinberg principle depends upon the following assumptions:

1. There is **no** selection.
2. There is **no** mutation.
3. There is **no** migration.
4. There are **no** chance events.
5. Individuals choose their mates at random.

**By these assumptions we stabilize the gene pool so we can perform our studies.

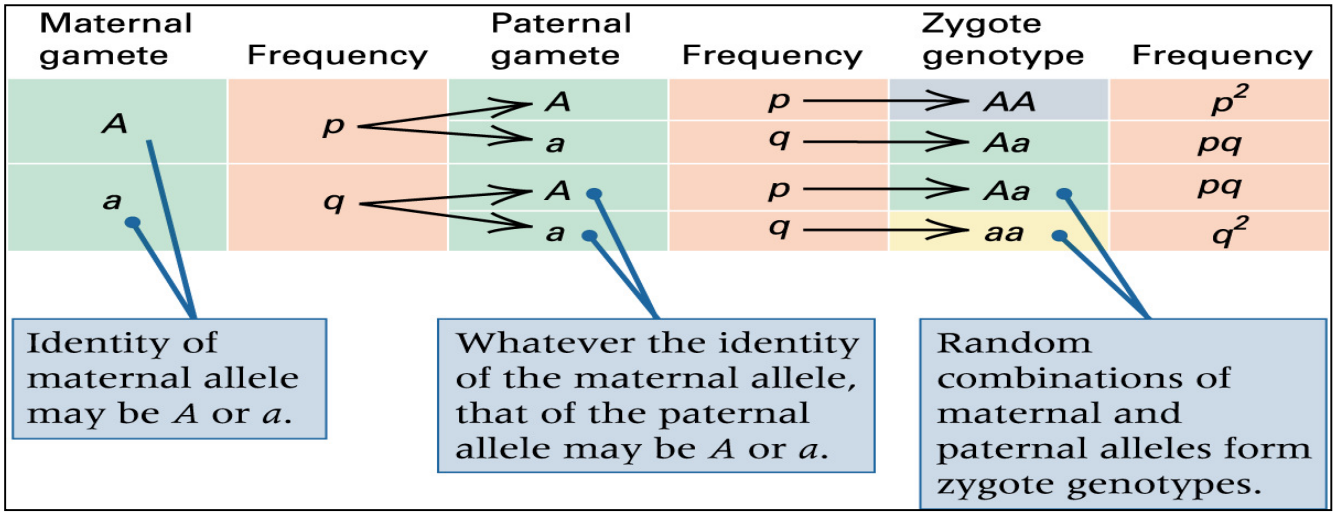
Taking an example to apply the equation:

- Assumptions:

1) Diploid, autosomal locus with 2 alleles: **A** and **a.**

2) Simple life cycle:
Parents (Diploid)⟶ Gametes (Haploid) ⟶ Zygotes (Diploid).

Now, we assume that the frequency of (A) allele to be distributed into the maternal gametes which are going to be fertilized is p for example and that of (a) is 1-p=q. Repeating the same thing for the paternal alleles then looking for fertilization of the maternal and paternal gametes, we get the frequencies of the zygote genotypes. Look at this figure:



- An example with numbers about MN blood groups:

Given:

In a population of 747 individuals (each gene has two alleles so the number of alleles would be 1494 alleles).

Problem:

1) Find the allele frequencies for M^m and M^n.
2) Find the genotypic frequencies of M^m M^m, M^m M^n, and M^n M^n.

Solution:

We performed manual tests in the laboratory and got these results:

Sample Population	Phenotypes	Type M	Type MN		Type N
	Genotypes	M^mM^m	M^mM^n		M^nM^n
747	Numbers	233	385		129
	Contribution to gene pool	2 M^m alleles per person	1 M^m allele per person	1 M^n allele per person	2 M^n alleles per person

For problem 1:

To calculate the frequency of M^m, notice that every person with the phenotype (Type M) has two M^m alleles, thus we multiply the number of people with this phenotype by 2 to know how many alleles exist in this category (2 alleles * 233 people = 466 alleles), also, every person with the phenotype (Type MN) has one M^m allele, thus the number of M^m alleles in this category is equal to the number of people it has (1 allele * 385 people = 385 alleles), the total number of M^m alleles is (466+385= **851 alleles of M^m**). Repeating the same steps for M^n on the phenotypes (Type MN and Type N), we get (**643 alleles of M^n**).

Adding the numbers of the two types of alleles we get the total number of alleles (851+643=**1494 total alleles**), by dividing the number of each allele by the total number of alleles and multiplying the result by 100% we get the frequency of each allele, for M^m ((851 / 1494) * 100% = **57%**), for M^n ((643 / 1494)*100% = **43%**). **The frequency of M^m is 57%, and the frequency of M^n is 43%.**

For problem 2:

To calculate the expected zygotic genotypic frequencies we assume that the individuals mate randomly:

(M^m M^m): 0.57*0.57= **0.32**

(M^n M^n): 0.43*0.43= **0.18**

M^m(paternal) M^n(maternal): 0.57*0.43= 0.25

M^m(maternal) M^n(paternal): 0.57*0.43= 0.25

So that for (M^mM^n)= 0.25+0.25= **0.5**

		SPERMS	
		M^m 0.57	M^n 0.43
EGGS	M^m 0.57	M^mM^m 0.32	M^mM^n 0.25
	M^n 0.43	M^mM^n 0.25	M^nM^n 0.18

The End :D

Best of luck ☺

Made by:

Mohammad Falahat & Ala'a Shaban