Pergamon

European Journal of Cancer, Vol. 33, No. 13, pp. 2160–2166, 1997

Printed in Great Britain
0959-8049/97 $17.00+0.00

PII: S0959-8049(97)00307-9

Review

Basic Population and Cancer Genetics and Their Use in the Assessment of Cancer Risk

G.H. Lyman and N.M. Kuderer

H. Lee Moffitt Cancer Center and Research Institute at the University of South Florida, 12902 Magnolia Drive, Tampa, Florida 33612, U.S.A.; and Albert Ludwigs Universität, Freiburg, Germany

INTRODUCTION

THERE IS considerable evidence that alteration in the content or expression of a cell's genetic information is responsible for the unregulated growth and abnormal differentiation of the malignant state [1]. A number of inherited preneoplastic syndromes are characterised by an increased risk of subsequent malignancy (Table 1). In addition, several neoplasms are inherited in an autosomal dominant manner with incomplete penetrance. These are characterised by early age of onset and the frequent appearance of multiple primary tumours [2, 3]. The genetics of more common cancers such as those of the breast and colon are clearly multifactorial with contributions from several genes as well as environmental factors [4]. Clear familial patterns are often present and risk relates to the degree and number of family members affected. The role of genetic factors in the majority of cancer patients remains poorly defined.

POPULATION GENETICS

Basic genetics

The fundamental laws of genetics are based on the work of the Austrian Monk Gregor Mendel [5]. Each individual possesses two genes for a particular characteristic, only one of which is transmitted to the next generation. A gene pair may be identical (homozygous) or different (heterozygous) in which case the two alleles of the gene code for contrasting characteristics. When two individuals, each homozygous for different alleles are crossed, the offspring of the first generation will be identical to each other, although different from the parents. Different gene pairs segregate independently to the offspring in proportion to their distance from one another on the chromosome.

More than 6000 traits in humans exhibit Mendelian inheritance with a unifactorial pattern of inheritance related to a single gene [6]. Such inheritance is most commonly autosomal but may also be sex-linked. If a trait is expressed in its heterozygous state, it is said to be dominant. Such traits can often be traced through many generations of a family [7]. The offspring of a couple where one has a homozygous normal genotype and the other is heterozygous for a dominant trait will have a 50% chance of being similarly affected. Autosomal dominant diseases include many of the most common and important inherited disorders such as hypercholesterolaemia and familial adenomatous polyposis. Other genes may modify the expression of a mutant allele, reducing the penetrance or expression of the trait [8]. Recessive traits are observed only in the homozygote, so both parents must possess the mutant allele and pass it on to their offspring. It is often difficult to trace an autosomal recessive characteristic or disease through a family unless the mutant gene is common in the population such as in cystic fibrosis or sickle cell anaemia [7]. An X-linked recessive trait is usually manifested only in males (hemizygous). Traits inherited in this manner, such as haemophilia A, are conveyed by heterozygous female carriers to affected male offspring. Affected males transmit the trait to their obligate carrier female offspring. Rare X-linked dominant traits, such as vitamin D-resistant rickets, are manifested in the heterozygous female as well as the male. The affected male transmits the trait to all of his daughters and none of his sons, distinguishing this from autosomal inheritance [9].

Some genes have mutated to form multiple alleles (polymorphic loci) some of which may be dominant and some recessive in behaviour [10]. More than one-third of human genetic loci are polymorphic with two or more alternative phenotypes occurring at a frequency greater than could be accounted for by recurrent mutations [11]. Examples include the ABO, MN and Rh blood groups and the major histocompatability antigens. Such polymorphisms allow for multiple combinations of alleles and thus enormous genetic diversity. Such stable polymorphisms may have originated as mutations, but require favourable selective pressures to sustain their presence in the population as in the case of the sickle cell trait resistance to *plasmodium falciparum* malaria [12, 15]. Since alleles are carried on homologous chromosomes, only one allele for a specific trait can be transmitted to an offspring.

Hardy–Weinburg equilibrium

G.H. Hardy and W. Weinburg proposed independently, at the beginning of the century, that the relative proportions of

Correspondence to G.H. Lyman.
Received 18 Feb. 1997; revised 27 May 1997; accepted 16 Jun. 1997.

Table 1. Hereditary preneoplastic syndromes

Inheritance pattern	Disorder	Malignancy
Autosomal dominant	Down syndrome	Acute leukaemia
	Neurofibromatosis	Sarcomas, gliomas, meningiomas
	Peutz Jegher syndrome	Sarcomas
	Von Hippel–Landau	Phaeochromocytoma, renal cell carcinoma
Autosomal recessive	Bloom syndrome	Acute leukaemia, lymphoma
	Fanconi's anaemia	Acute leukaemia
	Xeroderma pigmentosum	Skin cancer
	Albinism	Skin cancer
	Ataxia telangiectasia	Leukaemias, lymphomas
	Werner's syndrome	Sarcoma
Sex-linked recessive	Bruton's agammeglobulinaemia	Acute leukaemia, lymphoma
	Wiskott–Aldrich syndrome	Acute leukaemia, lymphoma
	Lymphoproliferative syndrome	Lymphoma, plasmacytoma

different genotypes remain relatively constant between generations in a randomly mating population not influenced by outside forces [14]. Consider a gene locus with only two alleles **A** and **a** with frequencies in the population of p and q, respectively. The binomial distribution describes the distribution of two alternative events occurring with probability p and $q = 1-p$ as:

$$f(x; n, p) = \frac{n!}{x!(n-x)!} p^x q^{n-x} \qquad \text{for } x = 0, 1, 2, \ldots n$$

where x is the number of events in n independent trials and p is the probability of the event on each trial. The binomial distribution gets its name from the fact that the values of the distribution are the successive terms in the binomial expansion of:

$$(p+q)^n = \sum_{x=0}^{n} \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

For instance, the expansion of $(p+q)^2$ yields $p^2 + 2pq + q^2$. Table 2 displays the resulting genotype frequencies for the mating of two individuals with allele frequencies of p and q, respectively. If the resulting genotypes mate, the gene frequencies in the second generation are those shown. As can be seen in Table 3, the frequency of each genotype is stable over successive generations (Hardy–Weinburg equilibrium) [15]. The Hardy–Weinburg relationship is of great utility in determining allele frequency and heterozygote carrier frequencies in the population.

Table 2. Allele frequencies

Randomly mating populations				
Generation 1			Male	
			A	**a**
			(p)	(q)
Female		**A** (p)	**AA** (p^2)	**Aa** (pq)
		a (q)	**Aa** (pq)	**aa** (q^2)
Generation 2			Male	
			AA	**Aa**	**aa**
			(p^2)	($2pq$)	(q^2)
Female	**AA** (p^2)	p^4	$2p^3q$	p^2q^2
	Aa ($2pq$)	$2p^3q$	$4p^2q^2$	$2pq^3$
	aa (q^2)	p^2q^2	$2pq^3$	q^4

For autosomal recessive conditions, the disease will be expressed only in the homozygous state with frequency q^2. The frequency of the mutant allele in the population is then: $q = \sqrt{q^2}$ and the frequency of the normal allele is: $p = 1 - q$. The frequency of the heterozygous carrier state is $2pq$. For example, the prevalence of Tay Sachs Disease among Ashkenazi Jews is 1 in 3900 (0.00026). The gene frequency in this population is, therefore, $q = \sqrt{1/3900} = 0.016$. The normal gene frequency is then: $p = 1 - q = 0.984$. The heterozygote carrier frequency is: $2pq = 2(0.984)(0.016) = 0.0315$ or approximately 1 in 30. The frequency of Tay Sachs in non-Ashkenazi Caucasians is approximately 1 in 112 000. With such rare recessive conditions, the normal allele frequency is almost 1. The heterozygote frequency ($2pq$) is, therefore, essentially $2q$ or twice the frequency of the mutant allele. As the frequency of a disorder among homozygotes becomes less, the ratio of heterozygotes to homozygotes in the population becomes larger.

For autosomal dominant disorders, the prevalence of the disorder is represented by $2pq + q^2$. Generally, homozygotes for such conditions are very rare (q^2) and are not considered. Since p is almost 1, the frequency of the mutant allele is approximately one-half of the estimated disease prevalence ($2q$). For instance, the prevalence of autosomal dominant disorder neurofibromatosis is approximately 1 in 3000 (0.00033) and the frequency of the allele is approximately 0.00017 ($0.00033 \div 2$).

For X-linked conditions, the frequencies of normal mutant alleles can be determined directly from the prevalence of the

Table 3. Hardy–Weinburg equilibrium

Randomly mating populations				
		Offspring		
Mating type		**AA** (p^2)	**Aa** ($2pq$)	**aa**(q^2)
AA × **AA**	p^4	p^4		
AA × **Aa**	$4p^3q$	$2p^3q$	$2p^3q$	
AA × **aa**	$2p^2q^2$		$2p^2q^2$	
Aa × **Aa**	$4p^2q^2$	p^2q^2	$2p^2q^2$	p^2q^2
Aa × **aa**	$4pq^3$		$2pq^3$	$2pq^3$
aa × **aa**	q^4			q^4

Offspring

AA: $p^4 + 2p^3q + p^2q^2 = p^2(p^2 + 2pq + q^2) = p^2(p+q)^2 = p^2$

Aa: $2p^3q + 4p^2q^2 + 2pq^3 = 2pq(p^2 + 2pq + q^2) = 2pq(p+q)^2 = 2pq$

aa: $p^2q^2 + 2pq^3 + q^4 = q^2(p^2 + 2pq + q^2) = q^2(p+q)^2 = q^2$

disorder in males (hemizygous). Since females have two X chromosomes, their genotype frequencies are distributed binomially like autosomal genotypes. The prevalence of the X-linked recessive condition haemophilia is approximately 1 in 10 000. Since the frequency of the normal allele is approximately $p = 1$, the frequency of the heterozygous carrier is only $2q$ or 1 in 5000 and the frequency of disease in males equals the frequency of the mutant allele ($q = 0.0001$). In X-linked dominant disorders, the prevalence of affected males is approximately half the number of affected females since the possibility of males receiving the mutant allele is halved.

Factors which alter gene frequencies

The relationships between disease prevalence and gene frequency discussed above hold true only under certain specific conditions. The Hardy–Weinburg equilibrium assumes random mating, constant mutation rate, no selection, no random fluctuations and no migration. Deviations from any of these conditions can alter gene allele frequency in the population over time [16].

In human populations, random mating is seldom observed. Most commonly this is displayed in assortative mating based on some physical trait which may increase the expression of recessive disorders [11]. In its least common form, non-random mating results from consanguinity (inbreeding) which increases the proportion of homozygotes for recessive conditions among the population. Non-random mating may also result from geographical or cultural isolation of a portion of the population. Genetic drift is said to occur when one allele is passed to the next generation with higher frequency. Migration of a population into a new region may result in changes in allele frequencies due to the mixing of populations. Environmental factors may also have a selective influence on a particular phenotype and genotype.

Another important disturbance from the Hardy–Weinburg equilibrium is that caused by a change in the rate of spontaneous mutations. A mutation represents an actual change in the structure of the genome. If such a change occurs in the gamete (germ line), this change may be passed on to subsequent generations. The mutation rate is the frequency of such change expressed as the number of mutations per locus per gamete for each generation [17]. If n is the number of affected individuals without affected parents, and N is the total number of births, the mutation rate for autosomal dominant diseases is $\mu = n/2N$, since alteration in either allele will result in the disorder. Mutation rates can also be estimated as the product of the mutant allele frequency and a measure of selection disadvantage such that $s = 1 - f$, where f is a measure of genetic fitness defined as the ratio of the number of offspring among affected individuals to the number of offspring among controls expressed as a proportion. For most human genes, the mutation rate appears to range from minimum of 0.000001 (10^{-6}) to a maximum of 0.00001 (10^{-5}) per locus per gamete per generation.

It is believed that most gene mutations arise as a result of environmental factors such as ionising radiation and chemical mutagens. In most populations, the introduction of gene mutations is balanced by the loss of mutant alleles due to reduced fitness for survival or propagation. The mutation rate for autosomal dominant conditions in Hardy–Weinberg equilibrium is $\mu = I(1 - f)/2$, where I is the incidence rate. For autosomal recessive conditions, two genes are lost for each homozygote that does not reproduce. Therefore, the mutation rate at equilibrium is $\mu = I(1 - f)$. For sex-linked recessive conditions, three X chromosomes are transmitted per mating per generation and the mutation rate is $\mu = [I_m(1 - f)]/3$, where I_m is the incidence rate in males. Measures of the mutation rate can be useful in estimating gene size and the determination of mutagenic potential for certain diseases or following specific exposures. Selection against recessive mutations is less effective than selection against dominant mutations since only homozygotes are subject to selective influences in the former situation.

Multifactorial inheritance

Many human diseases, including several malignancies, reflect a polygenic pattern of inheritance where the risk of disease relates to the cumulative effect of alterations at several different gene loci, each having a small impact on phenotype. The participation of multiple genes in the phenotypic expression of certain traits such as height, weight and blood pressure often yields measured traits which follow a normal distribution. This can be shown to be consistent with the frequency distribution of two or more gene loci each contributing to the phenotype. As the number of trials in a binomial expansion (n) increases, the binomial distribution representing the participation of multiple gene loci increasingly resembles a normal distribution (Table 4). The greater the number of genetic loci involved (polygenic) and the more alleles involved at each locus (polymorphic), the more closely the phenotype distribution will resemble a normal distribution.

Polygenic disorders such as cancer often run in families in proportion to the degree of closeness of the relationship [19]. First-degree relatives consisting of parents, children, siblings and dizygotic twins share half their genes. Second-degree relatives including grandparents, grandchildren and aunts and uncles, share a quarter of their genes. Third-degree relatives including great-grandchildren and first cousins share one-eighth of their genes. The risk is greatest among first-degree relatives where the frequency is approximately the square root of the population frequency.

Environmental factors are most likely to have an effect on those who have an inherited susceptibility. Phenotypic traits inherited polygenically often represent predispositions to disease or susceptibility to environmental factors increasing the occurrence of disease. Disease susceptibility controlled in such a manner by several genes often interact with environmental risk factors representing multifactorial inheritance. Genetic susceptibility from several gene loci may reach a threshold of susceptibility to environmental factors beyond

Table 4. Genotype and phenotype distributions

Two gene loci each contributing one-half to the phenotype

		Gene locus 1		
		AA (1/4)	**Aa** (1/2)	**aa** (1/4)
Gene locus 2	**BB** (1/4)	1/16	2/16	1/16
	Bb (1/2)	2/16	4/16	2/16
	bb (1/4)	1/16	2/16	1/16

Genotype distribution

	AaBB	**AAbb** **AaBb**	**aaBb**	
AABB	**AABb**	**aaBB**	**Aabb**	**aabb**
1/16	4/16	6/16	4/16	1/16

Phenotype distribution

which disease occurs. Those that develop disease may be at one end of the susceptibility distribution which is, in part, shared by family members that share those genes. The approximately normally distributed susceptibility (liability) to disease is shifted to the right in the relatives of an individual affected by a multifactorial trait or disorder. In such families, the proportion of the individuals lying beyond the threshold expressing the abnormal phenotype is increased. The mean susceptibility can be estimated from the disease's incidence and the underlying normal distribution. The occurrence of a multifactorial disorder in subsequent generations is greatest when more than one family member is affected and when more severe disorders have been observed. Heritability is the proportion of the total phenotypic variance that is due to genetic variance and therefore represents the proportion of the aetiology attributable to genetic rather than environmental factors. A variety of methods exist which may facilitate the identification and isolation of genes which contribute to multifactorial traits and disorders including disease associations, sib-pair analyses and linkage analyses.

THE GENETICS OF CANCER

There are more than 50 Mendelian disorders in which the risk of cancer is high, approaching 100% in some diseases [1–4]. While the occurrence of multiple primary tumours due to sporadic mutations appears to be rare, it is quite common among those with heritable cancers [20, 21]. Table 5 lists some of the commonly recognised heritable cancers and cancer susceptibility syndromes and their affected chromosomes. Although a Mendelian pattern of inheritance is observed in no more than 5% of all malignancies, an additional 15–20% of cancers appear to demonstrate a multifactorial pattern of inheritance with a significant familial component. Virtually all cancers appear to arise as a clonal pattern of growth related to altered genetic control in somatic cells. The development and progression of many cancers may involve the abnormal expression of an entire series of genes. The multiple somatic mutations associated with the development of most human cancers are not independent events since early mutations appear to increase the likelihood of subsequent mutations. Even when environmental factors such as viruses, chemicals or radiation are involved, the common thread that links these carcinogenic agents is their ability to cause mutations which alter the expression of genes that control cellular growth and proliferation. The rapidly developing field of molecular genetics has begun to elucidate more clearly some of the specific genes and processes involved in carcinogenesis [27]. Three distinct types of cancer-related genes are known to be involved in the molecular mechanisms of carcinogenesis: oncogenes, tumour suppressor genes and mismatch repair genes. Somatic mutations may alter such genes by either structural alteration or transcriptional deregulation.

Table 5. Inherited cancer-susceptibility syndromes

Disorder (cancer)	Region affected Chromosome	Gene
Retinoblastoma	13q	*RB1*
Wilms' tumour	11p	*WT1*
Familial adenomatous polyposis (colon, thyroid)	5q	*APC*
Neurofibromatosis type 1 (Von Recklinghausen)	17q	*NF1*
Li–Fraumeni syndrome (sarcoma, breast, CNS, leukaemia, adrenal)	17p	*p53*
Dysplastic nevus syndrome (melanoma)	1p	*CMM1*
Multiple endocrine neoplasia (MEN) 2a (Medullary ca thyroid, Phaeochromocytoma)	10q	*MEN2a*
Von Hippel–Lindau (haemangioblastoma, Phaeochromocytoma)	3p	*VHL*
Hereditary non-polyposis colorectal cancer (HNPCC)		
Lynch I (colon)	18q	*MSH2*
Lynch II (GI, GU)	2p	*LCF2*
Breast/ovary	17p	*BRCA1*

Oncogenes

Proto-oncogenes are genes involved in normal cell growth and proliferation. Such genes produce products involved in signal transduction, DNA-binding as well as secreted growth factors and growth factor receptors [23]. Proto-oncogenes appear to be very highly conserved in evolution, revealing their critical role in cellular growth and replication. An alteration of a proto-oncogene gives rise to an abnormal growth control gene termed a cellular oncogene [24]. Overexpression of such an altered gene may promote the uncontrolled growth and proliferation of cells [25]. More than 50 human oncogenes have been identified. Oncogenes may arise by integration of DNA transcribed from RNA retroviruses via the enzyme reverse transcriptase. It is likely that such viral oncogenes originated from ancestral host proto-oncogenes. DNA sequences from retroviruses may also be integrated into mammalian genome in proximity to a proto-oncogene such as *myc* altering its level of expression. However, most tumour-associated somatic mutations in human cancers are not virally-induced with certain notable exceptions, e.g., HTLV-1, HIV and hepatitis C.

A point mutation giving rise to a single mutant allele in a proto-oncogene such as the *ras* gene can give rise to an oncogene producing an abnormal gene product capable of stimulating cell growth. Overexpression of such gene products (oncoproteins) may result from gene amplification arising from the production of multiple copies of a particular DNA segment containing specific proto-oncogenes. However, it is apparent that mutation of a single proto-oncogene is usually not sufficient for complete malignant transformation. The participation of several activated oncogenes is often necessary for actual malignant transformation.

Numerous non-random chromosomal translocations have also been shown to result in proto-oncogene activation in human cancer [26]. Several chromosomal translocations associated with known human malignancies are shown in Table 6 along with the proto-oncogenes affected. An important illustration of the role of translocations in the expression of human oncogenes is found in chronic myelogenous leukaemia (CML). The majority of patients with CML demonstrate a translocation between the long arm of chromosome 9 and the long arm of chromosome 22 resulting in translocation of the *abl* proto-oncogene on the former to the breakpoint cluster region (*bcr*) on the latter. The *abl* oncogene then comes under the control of the promotor sequence on chromosome 22. The abnormal chimeric protein product of the *abl* and *bcr* regions has increased tyrosine kinase activity

Table 6. Common non-random chromosome translocations

Translocation	Selected malignancies Malignancy	Proto-oncogene affected
Haematopoietic		
t(8;21) (q22;q22)	Acute myeloid leukaemia	AML1-ETO
t(9;22) (q34;q11)	Chronic myelogenous leukaemia	ABL-BCR
t(15;17) (q23;q11)	Acute promyelocytic leukaemia	RAR-PML
t(1;14) (q34;q11)	Acute lymphocytic leukaemia	LCK-TCRD
t(12;21) (p12;q22)	Acute lymphocytic leukaemia	TEL-AML1
t(5;14) (q31;q32)	Acute lymphocytic leukaemia	IL3-IGH
t(8;14) (q24;q32)	Burkitt's lymphoma	MYC-IGH
t(14;18) (q32;q21)	Follicular lymphoma	IGH-BCL2
Solid tumours		
t(11;22) (q24;q12)	Ewings' sarcoma	FLI1-EWS
t(2;13) (q35;q14)	Rhabdomyosarcoma	PAX3-FKHR
t(X;18) (p11;q11)	Synovial sarcoma	SYT-SSX
t(12;16) (q13;p11)	Liposarcoma	CHOP-FUS

which is thought to play a role in the malignant phenotype. Another example of chromosomal translocations altering oncogene expression has been observed in Burkitt's lymphoma, a B cell lymphoma predominantly of children in equatorial Africa. The majority of cases have been found to be associated with a translocation of the *myc* oncogene from the long arm of chromosome 8 to the portion of the long arm of chromosome 14 that contains the immunoglobulin heavy chain gene resulting in overexpression of the *myc* product.

Tumour suppressor genes

It is now known that cellular proliferation is normally regulated by the products of a class of genes known as tumour suppressor genes [28]. The products of such genes are involved in cell cycle regulation, DNA binding and cell–surface interactions. Mutation of such genes can inactivate them and lead to a loss of control of cellular proliferation [29]. Since such mutations are recessive, the function of both alleles must be lost for malignancy to occur. Mutations of tumour suppressor genes appear to play a role in several autosomal dominant Mendelian disorders associated with cancer, including retinoblastoma (RBL), Wilms' tumour (WTI), neurofibromatosis, type I (NFI) and the Li Fraumeni syndrome (p53). Knudson theorised many years ago that such individuals inherit one abnormal gene through a germ line mutation. The other allele, he postulated, must then be inactivated somatically during development for malignancy to occur [30]. This theory suggests that while the susceptibility is inherited dominantly, tumour development occurs in a recessive manner (both alleles affected).

The loss of heterozygosity (LOH) necessary for the loss of normal tumour suppressor gene function may result from mutation of the normal allele, chromosome deletion, somatic recombination or mitotic non-disjunction. Mutations of tumour suppressor genes may occur in a wide variety of sporadic malignancies. Several mutations in different genes may, in fact, be necessary for the development of most malignancies. The sequential loss of tumour suppressor genes on several chromosomes as well as activation of proto-oncogenes has been demonstrated to occur in sporadically-occurring colon cancer which is thought to evolve from pre-malignant adenomatous polyps to invasive cancer [31, 32]. During this transition, there is a progressive accumulation of genetic alterations which include LOH for the *APC* (adenomatous polyposis coli) and *MCC* (mutated in colon cancer) genes on chromosome 5, mutations of the *ras* oncogene and *p53* tumour suppressor gene on chromosomes 1 and 17, respectively, and deletions of portions of chromosome 18 coding for the *DCC* (deleted in colon cancer) gene. LOH has been observed in other human cancer genes including those associated with breast cancer (*BRCA1*, *BRCA2*). The number of genetic alterations appears to be more significant to carcinogenesis than the specific order of such changes. Mutations in the tumour suppressor gene *p53* have been observed in the majority of several common tumours including those of lung, colon, breast and bladder.

Mismatch repair genes

Several autosomal recessive conditions accompanied by chromosomal instability are associated with an increased risk of malignancy, including leukaemia. These include ataxia telangiectasia, Bloom syndrome, Fanconi's anaemia and xeroderma pigmentosum. DNA in these individuals is susceptible to spontaneous chromosomal breakage enhanced by exposure to radiation and chemicals. The altered DNA repair mechanisms evident are the result of defective repair enzymes coded for by specific genes. Only recently has it been demonstrated that heterozygotes for these disorders also have an increased risk of certain malignancies. For example, heterozygotes for the ataxia telangiectasia gene are at increased risk of breast cancer at a young age.

More than one mutational event is generally necessary for a cell to transform completely from a normal phenotype to a malignant cell, suggesting a multistep progression to cancer. These mutations may occur in somatic elements leading to either activation of proto-oncogenes or inactivation of tumour suppressor genes or both. Mutations may occur less commonly in the germline and be passed on to future generations. Mutations occur at a rate of approximately one per million genes per cellular replication. Given the large number of cells in the body actively replicating, cancer would be a frequent event if it were not necessary for a series of mutations to occur in the carcinogenic pathway. Those who have inherited germ line mutations have a greater risk by virtue of being one step closer to the necessary number of changes.

RISK ASSESSMENT AND SCREENING

A clear inherited susceptibility has been established for only a small proportion of human malignancies. Those individuals, however, are often at a greatly increased risk of cancer at an early age and may be affected by multiple malignancies.

Family risk assessment

Risk assessment is relatively straightforward for conditions inherited in a Mendelian manner. However, risk estimation can be complicated by factors such as reduced penetrance and delayed onset of disease. Risk assessment begins with a process of estimating initial probabilities based on the specific diagnosis and the type of inheritance involved. These initial estimates are then modified on the basis of additional information derived from the family history and results from genetic marker studies. The revised risk is based on the relationship known as Bayes' theorem which provides a conditional or posterior estimate of risk based on the prior risk estimate and additional information derived from tests with specified performance characteristics, i.e., sensitivity and specificity.

For an autosomal dominant condition such as hereditary retinoblastoma, the probability that offspring in the next generation will inherit the mutant gene *RB1* from a parent with the gene is 1 in 2 (50%). However, the probability that that offspring will actually manifest an ocular retinoblastoma is in direct proportion to the penetrance of the disorder, which is the probability of the disease in an individual carrying the mutant gene. Therefore, the presence or absence of the disease can be viewed as a test of whether the mutant gene is present. In this context, the penetrance represents the sensitivity of phenotypic expression of the mutant gene while the specificity or the absence of disease in the absence of the mutant gene is nearly 100%. If the penetrance is 2/3, then the probability of disease in the offspring of an *RBI* carrying parent is (1/2) (2/3) = 1/3. From Bayes' theorem, it is evident that the probability of possessing the gene in the absence of evidence of the disease in this setting is 1/4 (Table 7). Therefore, the probability that an unaffected offspring from a carrier will pass the gene to the next generation is 1/8, while the probability of disease is (1/8) (2/3) = 1/12. It can be shown that this represents approximately the maximal risk of the disorder in this situation under varying assumptions.

For an autosomal recessive condition, the presence of the disease in offspring almost always requires that both parents be heterozygotes in which case 1/4 of the offspring will be homozygous for the disorders, 1/2 will be heterozygous and 1/4 will be normal. Therefore, 2/3 of the unaffected offspring will be carriers of the altered gene. As the probability of a heterozygous genotype is $2pq$, the overall probability of manifesting an autosomal recessive condition in the offspring of an unaffected individual whose parents were heterozygotes is $(2/3)(1/4)\,2pq = 1/3pq$. As discussed previously, gene frequencies can be derived from a knowledge of the prevalence of disease (q^2). For sex-linked disorders which are usually recessive, the disease is almost always transmitted to half the male offspring by a female carrier. Half the female offspring will be carriers. When transmitted by an affected son, none of the grandsons will be affected, but all of the granddaughters will be carriers.

Risk estimation for disorders demonstrating multifactorial patterns of inheritance is generally empirical based on cohort studies of affected families. Such estimates are complicated by the wide variation in incidence rates between different regions related to differences in gene frequencies as well as differences in lifestyle and environmental cofactors. The risk of a cancer with multifactorial features among first-degree relatives presenting at a young age is approximately the square root of the prevalence of the cancer in the general population. For instance, the lifetime risk of breast cancer in the general population is approximately 1/9 while the risk when one or more first-degree relatives are affected at an early age is approximately 1/3. Nevertheless, the empirical use of cohort experience among the families of affected individuals represents the best source of risk estimation in such situations.

With increasing awareness of specific gene mutations among selected individuals with common malignancies, risk assessment can now proceed on a more rational basis. Risk estimation must again start with a knowledge of the prior frequency of the mutant gene in the population. Information is also needed concerning the frequency of specific clinical risk factors in those with and without the mutation. Bayes' theorem may then be applied to estimate the posterior probability of carrying the gene among those possessing the risk factor. For instance, the mutant *BRCA1* allele on chromosome 17 is estimated to occur in approximately 1 in 1000 women in the population. The probability of a woman with a family history of breast cancer possessing the gene can be estimated from this frequency and a knowledge of the probability of a family history among those with and without the mutation. Available estimates indicate that the presence of a family history considerably increases the probability that a woman is a carrier of the mutant gene. The greatest impact is seen when the affected relative developed breast cancer at a young age. In this situation, the risk of being a gene carrier is increased one to two orders of magnitude over the age-matched population without a family history [33]. Such estimates may guide the cost-effective application of genetic testing to select women at high risk for the specific mutations.

Table 7. Sample Bayesian risk analysis for offspring of RB1 *gene carrier*

		RB1 gene		
		Yes	No	
Retinoblastoma	Yes	1/3	0	1/3
	No	1/6	1/2	2/3
		1/2	1/2	1

Assumption: Penetrance (2/3)

Bayes' theorem:

Prob(Disease | Negative Results)

$$= \frac{(1 - \text{Sensitivity})(\text{Prevalence})}{(1 - \text{Sensitivity})(\text{Prevalence}) + (\text{Specificity})(1 - \text{Prevalence})}$$

$$\text{Prob}(M/U) = \frac{(1 - \text{Pen})\text{Prob}(M)}{(1 - \text{Pen})\text{Prob}(M) + 1[1 - \text{Prob}(M)]}$$

$$= \frac{(1/3)(1/2)}{(1/3)(1/2) + 1(1/2)} = \frac{1/6}{1/6 + 1/2} = 1/4$$

Where Pen = Penetrance

Prob (M) = Probability of mutant allele

Prob ($M|U$) = Probability of mutant allele among unaffected offspring

Cancer screening

The identification of high-risk individuals and families can be of great value to the effective application of genetic counselling and cancer screening efforts. Prevention and early detection efforts can be effectively applied to high-risk individuals identified by such programmes. The appropriate prevention and screening effort depends upon the specific pattern of inheritance identified. To be useful, cancer screening tests should be capable of detecting premalignant or early-stage disease with high sensitivity before the onset of symptoms. In addition, there must be treatment available for early-stage disease which improves the outcome of patients with cancer. Ideally, such screening programmes should be inexpensive, non-invasive and associated with a low false-positive rate (high specificity).

Effective screening methods for specific disorders include mammography for breast cancer syndromes and colonoscopy for both familial adenomatous polyposis and the hereditary non-polyposis syndromes. Ultrasound and computer-assisted tomography may be useful in screening for ovarian, renal, endometrial and thyroid malignancies in those at risk. Biochemical studies may be very useful for screening those with various endocrine neoplasm syndromes. Important to the effectiveness of such screening efforts is their initiation at a young age commensurate with the earliest age of manifestation of malignancy in the syndrome. While these expensive and often invasive procedures may not be warranted in screening the general population, they may be applied cost-effectively to individuals at high risk due to the high predictive value associated with the increased incidence of malignancy in these individuals.

DNA testing is now available for several identified cancer susceptibility genes. It is apparent that particular care should be utilised in selecting candidates for such testing. Features of the malignancy which should increase the likelihood of an inherited cancer susceptibility include early age of onset, multiple tumours in different organs, bilateral tumours, as well as several first- or second-degree relatives with related common tumours or the same rare tumour. Such testing not only requires a complete understanding of population and cancer genetics, but raises many social and ethical dilemmas for the physician, patient and family.

1. Bishop JM. The molecular genetics of cancer. *Science* 1987, **235**, 308–311.
2. Knudson AG. Genetics of human cancer. *Ann Rev Genet* 1986, **20**, 231–251.
3. Ponder JD. Inherited predisposition to cancer. *Trends Genet* 1990, **6**, 213–218.
4. Weinberg RA. The genetic basis of cancer. *Arch Surg* 1990, **125**, 257–260.
5. Sorsby A. Gregor Mendel. *Br Med J* 1965, **i**, 333–338.
6. McKusick VA. *Mendelian Inheritance in Man*, 11th edn. Baltimore, Maryland, Johns Hopkins University Press, 1994.
7. Hartl DL. *A Primer of Population Genetics*, 2nd edn. Sunderland, U.K., Sinauer Associates, 1987.
8. Todd JA, Bell JI, McDeveitt HO. HLA-DQβ gene contributes to susceptibility and resistance to insulin-dependent diabetes mellitus. *Nature* 1987, **329**, 599–604.
9. Kurnit DM, Layton WM, Malthyse S. Genetics, chance and morphogenesis. *Am J Hum Genet* 1987, **41**, 979–995.
10. Humphries SE. DNA polymorphisms of the apolipoprotein genes: their use in the investigation of the genetic component of hyperlipidemia and atherosclerosis. *Atherosclerosis* 1988, **72**, 89–108.
11. Vogel F, Motulsky AG. *Human Genetics*, 2nd edn. Berlin, Springer-Verlag, 1987.
12. Friedman MJ, Trager W. The biochemistry of resistance to malaria. *Sci Am* 1981, **244**, 154–164.
13. Allison AC. Protection afforded by sickle-cell trait against subtertian malarial infection. *Br Med J* 1954, **i**, 290–294.
14. Emery AEH. *Methodology in Medical Genetics*, 2nd edn. Edinburgh, Churchill Livingstone, 1986.
15. Hardy GH. Mendelian proportions in a mixed population. *Science* 1908, **28**, 49–50.
16. Eriksson AW, Forsius HR, Hevanlinna HR, *et al. Population Structure and Genetic Disorders*. London, Academic Press, 1980.
17. Haldane JBS. The rate of spontaneous mutation of a human gene. *J Genet* 1935, **31**, 317–326.
18. Carter CO. Multifactorial genetic disease. *Hosp Practice* 1970, **5**, 45–59.
19. Falconer DS. The inheritance of liability to certain diseases estimated from the incidence among relatives. *Ann Hum Genet* 1965, **29**, 51–76.
20. Lynch HJ. 'Cancer families': adenocarcinomas (endometrial and colon carcinoma) and multiple primary malignant neoplasms. *Rec Results Cancer Res* 1967, **12**, 125–142.
21. Li FP, Fraumeni JF. Soft tissue sarcomas, breast cancer, and other neoplasms: a familial syndrome? *Ann Intern Med* 1969, **71**, 747–752.
22. Hodgson SV, Mater ER. *A Practical Guide to Human Cancer Genetics*. Cambridge, Cambridge University Press, 1993.
23. Varmus HE. The molecular genetics of cellular oncogenes. *Annu Rev Genet* 1984, **18**, 553–573.
24. Miller DM, Blume S, Borst M, *et al.* Oncogenes, malignant transformation and modern medicine. *Ann J Med Sci* 1990, **300**, 59–69.
25. Friend SH, Dryja TP, Weinberg RA. Oncogenes and tumor-suppressing genes. *N Engl J Med* 1988, **318**, 618–625.
26. Mitelman F. *Catalogue of Chromosome Abberations in Cancer*, 3rd edn. New York, Alan R. Liss, 1988.
27. Rowley JD. The Philadelphia chromosome translocation: a paradigm for understanding leukemia. *Cancer* 1990, **65**, 2178–2184.
28. Harris H, Miller OJ, Klein G, Worst P, Tachibam T. Suppression of malignancy by cell fusion. *Nature* 1969, **350**, 377–378.
29. Sager R. Tumor suppressor genes: the puzzle and the promise. *Science* 1990, **246**, 1406–1412.
30. Knudson AG. Mutation and cancer: statistical study of retinoblastoma. *Proc Natl Acad Sci USA* 1971, **68**, 820–823.
31. Vogelstein R, Fearon ER, Hamilton SR, *et al.* Genetic alterations during colorectal tumor development. *N Engl J Med* 1988, **319**, 525.
32. Stanbridge EA. Identifying tumor suppressor genes in human colorectal cancer. *Science* 1990, **247**, 12–13.
33. Berry DA, Parmigiani G, Sanchez J, Schildkraut J, Winer E. Probability of carrying a mutation of breast-ovarian cancer gene BRCA1 based on family history. *J Natl Cancer Inst* 1997, **89**, 227–238.