Evidence-Based Medicine — Learn More

The concept of **evidence-based medicine** grew out of the recognition that traditional strategies for clinical decision making had limitations. Clinical medicine has always relied on "expert opinion" — that is, the considered judgment of physicians with longstanding experience in caring for patients with particular disorders. However, even when experienced physicians share their findings, opportunities for systematic observation are limited, and impressions about the effect of different treatment strategies may be subjective.

Pathophysiological reasoning is another approach to clinical decision making. In the first half of the twentieth century, dramatic gains in the understanding of disease biology led to new therapeutic advances, such as insulin treatment for diabetes. This approach led to the use of pathophysiological reasoning, or understanding of the disease process, as the basis for clinical decision making, and to advances such as anti-hypertensive treatment and curative cancer surgeries. This approach has had important successes, but also has limitations. For example, some clinical trials based on pathophysiological reasoning have failed to document anticipated benefits. In one recent study, data from the Women's Health Initiative demonstrated that estrogen replacement therapy has higher risks and fewer benefits than anticipated, and, contrary to recent medical practice, may have few medical indications [Nelson et al 2002, Rossouw et al 2002].

Evidence-based medicine recognizes that expert opinion and pathophysiological reasoning make important contributions to clinical decision making, but that the greatest certainty about medical benefit derives from studies that measure clinical outcomes directly and systematically. In this approach, clinical decision making is guided by an evaluation of the available evidence, both for its quality and for its relevance to the problem at hand. Clinical recommendations have greatest certainty when they are based on outcome data of high quality.

Relevant studies are classified by their design, to determine how much certainty they provide. Within design categories, studies should be examined for internal and external validity and for the concordance of results across similar studies. There should be an explicit link between the quality of evidence and the degree of certainty about management recommendations.

For **studies of medical interventions**, the quality of information available from a well-planned and executed randomized controlled trial is greater than that from a cohort or case-control study, which in turn is greater than that from case series or expert opinion (Table 1).

For **studies of prevalence and risk** (for example, the risk associated with a particular genotype), the quality of information available from a large population-based sample is greater than from a small population-based sample, case-control studies, and public recruitment (Table 2).

For **studies of diagnostic accuracy**, which are of particular importance in genetics, careful attention is needed to the populations that serve as the source for cases and controls, and also to the clinical measures used to define affected and unaffected persons.

Table 1. Evaluating Medical Interventions: Study Designs Ranked by Quality of Evidence

I.	Properly randomized controlled trial (RCT) or systematic reviews of RCTs
II. A.	Well-designed controlled trial without randomization
II. B.	Well-designed cohort or case-control analytic study
II.C.	Time series with or without the intervention or dramatic results in an uncontrolled experiment
III.	Opinion of respected authority, descriptive study or case report, or report of expert committee

Adapted from US Preventive Services Task Force 1996

Table 2. Evaluating Prevalence and Risk: Study Designs Ranked by Quality of Evidence

I.A.	Large population-based sample from the same population as patient in question
I.B.	Large population-based sample from a population geographically similar to that of the patient in question
I.C.	Large population-based sample not associated with patient in question
II. A.	Small population-based sample, or proxy for population sample
II.B.	Clinical case-control series

II. C.	Public recruitment of volunteer subjects by reproducible method
III.	Samples of cases and controls meeting clinical criteria, without defined recruitment method

Adapted from Seymour et al 1997

Prevalence of Genetic Conditions

The prevalence of genetic conditions varies across a wide range. However, even relatively common genetic conditions are infrequent compared to medical conditions commonly seen in primary care.

Table 3a. Estimated Prevalence of Various Genetic Conditions [1]

Huntington disease	1/14,000 - 1/33,000
Familial adenomatous polyposis (FAP)	1/8,000 - 1/40,000
Duchenne muscular dystrophy (at birth)	1/6,000
Cystic fibrosis (at birth)	1/2,500 - 1/3,300
Hereditary non-polyposis colon cancer (HNPCC)	1/500 - 1/6,700
Klinefelter syndrome (males)	1/500
Familial hypercholesterolemia	1/500
Familial combined hyperlipidemia	1/100
Factor V Leiden (heterozygote)	1/20 - 1/100

1. Range indicates estimates from different studies or populations.

Table 3b. Estimated Prevalence of Various Common Conditions [1]

Depression (adults)	1/9 - 1/18
Asthma (all ages)	1/9
Obesity (adults)	1/4 - 1/7
Hypertension (adults)	1/3 - 1/4

1. Range indicates estimates from different studies or populations.

Observations on Genotype-Phenotype Correlation in Medical Genetics

The traditional explanation for genetic disease is that a mutation in a

particular gene leads to the production of an abnormal protein, or to the lack of a protein, which in turn causes disease. Many genetic disorders fit this model (for example, Duchenne muscular dystrophy; hemophilia A). However, progress in genetic research has made the complexity of the relationship between genotype and phenotype apparent.

Concepts of incomplete penetrance and variable expressivity

Geneticists use two terms to describe the complexity in the relationship between genotype and phenotype:

- **Incomplete penetrance**, referring to the circumstance in which some people with a disease-associated genotype do not develop the disease
- **Variable expressivity**, referring to variable clinical manifestations among people with the same genotype

These phenomena have long been thought to be rare, but molecular studies now suggest that they are common, even in classically severe genetic diseases. For example, pulmonary disease severity may vary widely among people with the same cystic fibrosis (CF) genotype. Similar variation in phenotype is seen in sickle cell disease. Many genetic conditions also result in variable clinical manifestations: for example, people with neurofibromatosis type 1 may experience mild or severe skin manifestations, and a few may experience other complications such as bony dysplasia, visual impairment, or malignant schwannoma (See *GeneReview: CFTR*-Related Disorders; Case 12. CF; Case 13. CF; Salvatore et al 2002; CF Genotype-Phenotype Consortium 1993; *GeneReview*: NF1; Case 40. Neurofibromatosis; Beutler 2001; Dipple & McCabe 2000.)

Genotypes as risk factors

Increasingly, molecular genetic research is also identifying genotypes that are properly considered risk factors rather than indicators of genetic disease. These genotypes contribute to the occurrence of disease as one of several risk factors. An example is factor V Leiden, a gene variant associated with increased risk of venous thromboembolism. The estimated lifetime risk of venous thrombosis for a person with factor V Leiden is 10% to 12%. (See *GeneReview:* Factor V Leiden Thrombophilia; Case 39. Thrombophilia; Juul et al 2004.)

Another example is genotype testing for hereditary hemochromatosis (HHC), a genetic disorder associated with excess iron accumulation. HHC is caused by mutations in the *HFE* gene. Most individuals with clinical symptoms related to HHC are homozygotes for the *HFE* mutation designated C282Y. That is, they have two copies of the C282Y mutation, so that their *HFE* genotype is C282Y/C282Y. The earliest clinical findings in HHC are nonspecific symptoms such as fatigue, joint pain (arthralgia), and abdominal pain; later complications include cirrhosis, diabetes mellitus, and cardiomyopathy.

Carriers — people with only one copy of the C282Y mutation — are typically unaffected. About 10% of people of northern European origin carry the C282Y mutation. If both parents are carriers of C282Y, their offspring have a 25% of having a C282Y/C282Y genotype; thus, risk is inherited in an autosomal recessive manner. However, current studies suggest that only a small percentage of people with the C282Y/C282Y genotype develop clinical symptoms of HHC.

In this case, the C282Y/C282Y genotype identifies a risk state rather than a disease condition. That is, an individual with the C282Y/C282Y genotype has an increased risk of developing future, serious complications of iron overload. Although early treatment with phlebotomy could potentially prevent complications such as cirrhosis, cardiomyopathy, and diabetes mellitus, most people with this genotype are likely to remain well without treatment. Thus, the genotype is an indicator for surveillance of iron status (for example, with serum iron measures such transferrin saturation and serum ferritin) rather than for treatment. (See Asberg et al 2001; Beutler et al 2002; McCune et al 2002; Gleeson et al 2004; *GeneReview:* HFE-Associated Hereditary Hemochromatosis; Case 25. HHC; Case 26. HHC.)

Genetic Test Characteristics

Many genetic tests use DNA-based technology, but any laboratory test used primarily to identify an inherited condition is considered a genetic test. Geneticists have created terms that help define their performance characteristics. Three basic questions must be answered:

- Does the test accurately identify the genetic variant of interest? (analytic validity)
- Does identifying that variant accurately predict the presence or risk of having the related clinical condition? (clinical validity)
- What are the outcomes associated with identifying the clinical

condition? (clinical utility)

Analytic validity. The term analytic validity refers to the accuracy with which a particular genetic characteristic (for example, a DNA sequence variant) can be identified by a given laboratory test.

Example: The accuracy with which a particular lab technique identifies whether a person has the C282Y/C282Y genotype. Analytic validity is usually very high for molecular genetic tests to identify specific gene variants.

Clinical validity. Clinical validity refers to the accuracy with which a genetic test identifies a particular clinical condition. It is described in terms of sensitivity, specificity, positive predictive value, and negative predictive value.

Examples:

Diagnostic testing. When a test is used diagnostically, clinical validity measures the accuracy with which the test identifies a person with the clinical condition in question. For example, a test for mutations in the RET gene detects a disease-causing mutation in 95% of persons with medullary endocrine neoplasia type 2 (MEN 2). Specificity is assumed to be at or close to 100%, based on the high penetrance observed in MEN 2 families, that is, the high likelihood — approaching 100% — that disease will occur in an individual with the disease-related genotype. (See Case 29. MEN 2; Case 30. Medullary Thyroid Cancer.)

Carrier testing. Genetic tests are sometimes used to detect carriers of autosomal or X-linked recessive diseases. Sensitivity is the key parameter for this type of testing; it is evaluated by assessing test performance in obligate carriers — that is, people who are known to be carriers because they have affected children. Counseling based on carrier testing needs to take into account any limitations in the sensitivity of the test; for example, CF carrier testing has different sensitivity in different racial/ethnic groups. (See Case 12. CF.)

Predictive testing. Genetic testing can be done in asymptomatic individuals to identify genetic susceptibility to

future disease. In this use of genetic testing, clinical validity measures the accuracy with which the test predicts a future clinical outcome. This measure depends on the penetrance of the genetic trait being measured and the prevalence of the clinical condition. For example, current estimates, based on a meta-analysis of population-based data, suggest that for a woman with a BRCA1 mutation, the average risk of breast cancer by age 70 years is about 65% and the average risk of ovarian cancer is about 40% [Antoniou et al 2003]. Given that risk is under 100%, and that there are wide confidence intervals in these risk estimates, these results suggest that other factors, either genetic or environmental, modify the effect of these mutations. (See Case 2. BRCA; Case 3. BRCA.)

Clinical utility. Clinical utility refers to the risks and benefits resulting from genetic test use. The most important considerations in determining clinical utility are: 1) whether the test and any resulting interventions lead to an improved health outcome among people with a positive test result; and 2) what risks occur as a result of testing. Complete measurement of clinical utility requires evaluation of the medical and social outcomes associated with testing (and any subsequent interventions) for people with both positive and negative test results. When treatment is unavailable, a genetic test with high clinical validity may be useful to establish a diagnosis or provide prognosis; in this situation, the value of testing is determined by clinical validity.

Evaluating Clinical Outcomes in Genetic Conditions

Difficulties in studying genetic conditions

Because genetic conditions are typically rare (Table 3a), the number of affected individuals is small and studies of clinical outcomes can be difficult. Even "common" genetic conditions occur in a small percentage of the population. For example, factor V Leiden is present in 1-5% of the population, and familial hypercholesterolemia is estimated to have a prevalence of 0.2%. Definitive understanding of the risk associated with a particular genotype, and of the benefit of interventions, may be dependent on case-control studies or well done pooled analyses of different studies. An understanding of the hierarchy of epidemiological studies helps providers evaluate the certainty attributable to different studies (Table 2).

An additional problem in the study of genetic conditions is that initial

studies may involve selection biases. For example, the estimate of cancer risk associated with *BRCA1* and *BRCA2* mutations was higher in the initial studies, based on high-risk families seen in referral centers, than in later population-based studies [Antoniou et al 2003].

Evaluation of treatment

For many genetic conditions, treatment is based on knowledge of disease biology, with benefits assessed by historic controls. An example is the use of prophylactic thyroidectomy in children with MEN 2 to prevent medullary thyroid cancer. Evaluation of small cohorts receiving this therapy indicate a definite benefit, with few cases of medullary carcinoma occurring over several years of follow-up among patients who would historically have been at high risk [Brandi et al 2001, Case 29, Case 30]. This example illustrates that observational data may be convincing despite small study samples and the lack of randomization, blinding, or other controls used to improve the quality of data in clinical trials.

Small numbers are not the only barrier to RCT evaluation of treatment. The expectation of treatment benefit may be high enough with some genetic conditions as to make randomization ethically unacceptable, as in the case of early screening colonoscopy in hereditary non-polyposis colon cancer [CRC Summary, Case 8]. In other cases an RCT may not be possible because many patients find the intervention (for example, prophylactic mastectomy) unacceptable.

Even drug treatment for genetic conditions may be based on limited data. Clinical trials are required for the approval of new drug treatments, but a randomized study design may not be required for rare conditions, and outcomes may be limited to intermediate biological measures. For example, replacement therapy for alpha(1)-antitrypsin (AAT) deficiency was approved on the basis of clinical studies demonstrating that replacement therapy could maintain target serum levels in people with severe deficiency, rather than on the basis of improved clinical outcomes [World Health Organization 1997].

Given these considerations, the appropriate management of genetic diseases, and therefore the clinical utility of the associated genetic tests, is often best determined by the collection of high quality observational data, such as well-designed cohort studies or case series [Wilcken 2001]. With this approach, however, uncertainties about optimal practice may remain long after the treatment is introduced, particularly regarding the timing of

treatment and the selection of patients most likely to benefit.

Evidence standard for genetic risk assessment in common complex diseases

Although observational data may provide an acceptable basis for determining the clinical utility of tests for rare, high-risk genetic conditions, this standard is not likely to be acceptable for gene variants associated with common, complex disorders. For example, many genetic tests identify individuals with an increased risk for cardiovascular disease. However, they are typically less informative than intermediate measures of risk, such as lipid profiles, which capture the effects of both genotype and diet and other lifestyle factors [Humphries et al 2004, Case 4, Case 5, Case 6]. Even when a genetic test is an established independent risk factor, its clinical utility may be low. For example, factor V Leiden (FVL) confers an increased risk for venous thromboembolism, but randomized clinical trials indicate that FVL testing does not provide useful guidance for anticoagulant therapy, limiting the clinical utility of the test [Ridker et al 2003, Case 39]. These observations argue for the importance of controlled studies of outcomes in determining the clinical utility of genetic tests for low penetrance gene variants.

References

Antoniou A, Pharoah PD, Narod S, Risch HA, Eyfjord JE, Hopper JL, Loman N, Olsson H, Johannsson O, Borg A, Pasini B, Radice P, Manoukian S, Eccles DM, Tang N, Olah E, Anton-Culver H, Warner E, Lubinski J, Gronwald J, Gorski B, Tulinius H, Thorlacius S, Eerola H, Nevanlinna H, Syrjakoski K, Kallioniemi OP, Thompson D, Evans C, Peto J, Lalloo F, Evans DG, Easton DF (2003) Average risks of breast and ovarian cancer associated with BRCA1 or BRCA2 mutations detected in case series unselected for family history: a combined analysis of 22 studies. *Am J Hum Genet* 72:1117-30 [Medline]

Asberg A, Hveem K, Thorstensen K, Ellekjter E, Kannelonning K, Fjosne U, Halvorsen TB, Smethurst HB, Sagen E, Bjerve KS (2001) Screening for hemochromatosis: high prevalence and low morbidity in an unselected population of 65,238 persons. *Scand J Gastroenterol* 36:1108-15 [Medline]

Beutler E (2001) Discrepancies between genotype and phenotype in hematology: an important frontier. *Blood* 98:2597-602 [Medline]

Beutler E, Felitti VJ, Koziol JA, Ho NJ, Gelbart T (2002) Penetrance of 845G>A (C282Y) *HFE* hereditary haemochromatosis mutation in the USA. *Lancet* 359:211-8 [Medline]

Brandi ML, Gagel RF, Angeli A, Bilezikian JP, Beck-Peccoz P, Bordi C, Conte-Devolx B, Falchetti A, Gheri RG, Libroia A, Lips CJ, Lombardi G, Mannelli M, Pacini F, Ponder BA, Raue F, Skogseid B, Tamburrano G, Thakker RV, Thompson NW, Tomassetti P, Tonelli F, Wells SA Jr, Marx SJ (2001) Guidelines for diagnosis and therapy of MEN type 1 and type 2. *J Clin Endocrinol Metab* 86:5658-71 [Medline]

Cystic Fibrosis Genotype-Phenotype Consortium (1993) Correlation between genotype and phenotype in patients with cystic fibrosis. *N Engl J Med* 329:1308-13 [Medline]

De Braekeleer M, Allard C, Leblanc JP, Simard F, Aubin G (1997) Genotype-phenotype correlation in cystic fibrosis patients compound heterozygous for the A455E mutation. *Hum Genet* 101:208-11 [Medline]

Dipple KM, McCabe ER (2000) Modifier genes convert "simple" Mendelian disorders to complex traits. *Mol Genet Metab* 71:43-50 [Medline]

Gleeson F, Ryan E, Barrett S, Crowe J (2004) Clinical expression of haemochromatosis in Irish C282Y homozygotes identified through family screening. *Eur J Gastroenterol Hepatol* 16:859-63 [Medline]

Humphries SE, Ridker PM, Talmud PJ (2004) Genetic testing for cardiovascular disease susceptibility: a useful clinical management tool or possible misinformation? *Arterioscler Thromb Vasc Biol* 24:628-36 [Medline]

Imperatore G, Pinsky LE, Motulsky A, Reyes M, Bradley LA, Burke W (2003) Hereditary hemochromatosis: perspectives of public health, medical genetics, and primary care. *Genet Med* 5:1-8 [Medline]

Juul K, Tybjaerg-Hansen A, Schnohr P, Nordestgaard BG (2004) Factor V Leiden and the risk for venous thromboembolism in the adult Danish population. *Ann Intern Med* 140:330-337 [Medline]

McCune CA, Al-Jader LN, May A, Hayes SL, Jackson HA, Worwood M (2002) Hereditary haemochromatosis: only 1% of adult HFEC282Y homozygotes in South Wales have a clinical diagnosis of iron overload. *Hum Genet* 111:538-43 [Medline]

Middeldorp S, Meinardi JR, Koopman MM, van Pampus EC, Hamulyak K, van Der Meer J, Prins MH, Buller HR (2001) A prospective study of asymptomaic carriers of the factor V Leiden mutation to determine incidence of venous thromboembolism. *Ann Intern Med* 135:322-7 [Medline]

Mullins CD, Blatt L, Wang J (2002) Societal implications of the pharmacoeconomics of alpha1-antitrypsin deficiency. *Expert Rev Pharmacoeconomics Outcomes Res* 2:243-9

Nelson HD, Humphrey LL, Nygren P, Teutsch SM, Allan JD (2002) Postmenopausal

hormone replacement therapy: scientific review. *JAMA* 288:872-81 [Medline]

Ridker PM, Goldhaber SZ, Danielson E, Rosenberg Y, Eby CS, Deitcher SR, Cushman M, Moll S, Kessler CM, Elliott CG, Paulson R, Wong T, Bauer KA, Schwartz BA, Miletich JP, Bounameaux H, Glynn RJ; PREVENT Investigators (2003) Long-term, low-intensity warfarin therapy for the prevention of recurrent venous thromboembolism. *N Engl J Med* 348:1425-34 [Medline]

Rossouw JE, Anderson GL, Prentice RL, LaCroix AZ, Kooperberg C, Stefanick ML, Jackson RD, Beresford SA, Howard BV, Johnson KC, Kotchen JM, Ockene J; Writing Group for the Women's Health Initiative Investigators (2002) Risks and benefits of estrogen plus progestin in healthy postmenopausal women: principal results From the Women's Health Initiative randomized controlled trial. *JAMA* 288:321-33 [Medline]

Salvatore F, Scudiero O, Castaldo G (2002) Genotype-phenotype correlation in cystic fibrosis: the role of modifier genes. *Am J Med Genet* 111:88-95 [Medline]

Seymour CA, Thomason MJ, Chalmers RA, Addison GM, Bain MD, Cockburn F, Littlejohns P, Lord J, Wilcox AH (1997) Newborn screening for inborn errors of metabolism: a systematic review. *Health Technol Assess* 1:i-iv, 1-95 [Medline]

US Preventive Services Task Force (1996) Guide to Clinical Preventive Services, 2 ed. Williams & Wilkins, Baltimore

Wilcken B (2001) Rare disease and the assessment of intervention: What sorts of clinical trials can we use? *J Inherit Metab Dis* 24:291-8 [Medline]

World Health Organization (1997) Alpha 1-antitrypsin deficiency: memorandum from a WHO meeting. *Bull World Health Organ* 75:397-415 [Medline]