Commentary

Epidemiologic Evaluation of Screening for Risk Factors: Application to Genetic Screening

MUIN J. KHOURY, MD, CAROL A. NEWILL, PHD, AND GARY A. CHASE, PHD

Abstract: To assess the usefulness of screening for risk factors, we derived arithmetic relationships between screening parameters (sensitivity, specificity, and positive predictive value PPV) and risk factor frequency, disease frequency and relative risk. We evaluated these relationships in the special case of genetic markers and disease susceptibility. It can be shown that even in the face of very large relative risks, sensitivity and positive predictive value are affected by the relative magnitude of disease and genetic marker frequencies. When the genetic marker is less frequent than the disease, PPV increases with increasing relative risk but sensitivity remains low. When the genetic marker is more frequent than the disease, sensitivity increases with increasing relative risk but PPV remains low. When marker and disease frequencies are equal, both PPV and sensitivity increase with increasing relative risks, but very high relative risks (> 100) have to be obtained for rare diseases. Depending on the goals of the screening program, these relationships can be used to predict the relative magnitudes of false positives (low PPV) and false negatives (low sensitivity). This approach can be generalized to evaluate nongenetic risk factors in screening programs as well. (*Am J Public Health* 1985;75:1204–1208.)

Introduction

Over the past few years, there has been an increasing interest in screening for risk factors to identify healthy people at risk to develop disease.[1] In the workplace, some railroad companies use low back X-rays to predict which individuals are at increased risk of back injuries,[2] and asbestos workers may be screened for cigarette smoking because of the synergistic effects in lung cancer risk.[2]

Genetic screening has recently received attention,[3–5] especially in the workplace,[6–12] based on evidence that genetic differences (e.g., isoenzymes, serum proteins, blood groups, and histocompatibility antigens) may be associated with different diseases or predispose individuals to the effects of physical, chemical, and biologic agents.[13–15] Genetic screening identifies persons with specific genotypes who are at risk of developing future disease but are usually healthy at the time of screening. For most genetic tests, laboratory methods are usually accurate enough to have high sensitivity and specificity with respect to correct classification of the person's genotype. What we are concerned with here are the sensitivity and specificity of the genetic marker with respect to the development of disease subsequent to screening.

At the present time, there are no clear criteria to evaluate screening for risk factors. Rothstein points to the limitations of using low back X-ray in screening workers because of its poor predictive value.[2] Omenn has recently proposed several criteria for research development in the area of genetic screening.[9] These include among others: a high prevalence of the genetic trait (at least 5 per cent), and high relative risk (at least 3, preferably 10) associating the marker with the disease. While these guidelines appear intuitively sound, disease frequency must also be considered because of its large impact on positive predictive value.[1]

From the Johns Hopkins University School of Hygiene and Public Health, Departments of Epidemiology and Mental Health. Address reprint requests to Dr. Muin J. Khoury, Department of Epidemiology, JHU School of Hygiene and Public Health, 615 N. Wolfe Street, Baltimore, MD 21205. This paper, submitted to the Journal November 30, 1984, was revised and accepted for publication April 30, 1985.

We have derived simple arithmetic relationships between screening parameters (sensitivity, specificity and positive predictive value) and genetic marker (or risk factor) frequency (m), disease frequency (p), and the relative risk (R). Although the formulas are applied here to genetic screening, they may be used in evaluating screening for nongenetic risk factors as well. Other issues important in coming to decisions about screening—such as cost and benefit, acceptability, and intervention measures—are not considered here.

Methods/Formulation

Consider Table 1 in which screened individuals in the population are stratified according to the presence (or absence) of the genetic marker, and to whether they will (or will not) develop a particular disease subsequent to screening. To compute values of sensitivity and specificity, we assume that the marginal probabilities are known:

Sensitivity (b) of the marker with respect to the disease is the conditional probability of carrying the marker given the presence of disease, Prob(M+|D+). Specificity (a) of the genetic marker is the conditional probability of not carrying the marker given the absence of disease development, Prob(M−|D−).

If the relative risk R relating the marker to the disease is known, values of b and a in terms of m, p, and R can be calculated. Essentially, R is the ratio of the probability of

TABLE 1—Relationship of Genetic Marker Sensitivity and Specificity to Marker and Disease Frequencies

	Disease		
Marker	Will Develop Disease (D+)	Will Not (D−)	Total
Present (M+)	bp	(1 − a)(1 − p)	m
Absent (M−)	(1 − b)p	a (1 − p)	1 − m
Total	p	1 − p	1

b = sensitivity; a = specificity.
m = marker frequency; p = disease frequency.

developing disease given the presence of the marker to the probability of developing disease given the absence of the marker, or R = Prob(D+|M+) / Prob(D+|M−), and can be shown (Table 1) to be equal to (bp/m) / (((1-b)p)/(1-m)). By rearranging the above equation, the value of b can be obtained:

$$(1)\ \text{Sensitivity} = b = \frac{Rm}{1+m\,(R-1)}$$

As the probability of carrying the marker is the sum of joint probabilities in cells 1 and 3 (Table 1), the value of a can be obtained from the following equation:

$$m = bp + (1-a)(1-p)$$

$$(2)\ \text{Specificity} = a = 1 - \frac{m\left(1 - \frac{Rp}{1+m\,(R-1)}\right)}{1-p}$$

The positive predictive value of the genetic marker (PPV) is the conditional probability of developing disease given the presence of the marker. From Table 1, PPV = bp/m, or by replacing the value of b in the equation:

$$(3)\ \text{Positive predictive value} = \text{PPV} = \frac{Rp}{1+m\,(R-1)}$$

Lastly, dividing equations (1) and (3), it can be shown that:

$$(4)\ \text{Sensitivity/PPV} = m/p$$

Results/Outcomes

Features of Genetic Marker Sensitivity

By examining equation 1, several features of sensitivity can be noted:

- Sensitivity increases with relative risk.
- Sensitivity increases with increasing marker frequency.
- Sensitivity is generally low for RR < 10 and m < 10 per cent. Large values of m and R are needed to give sensitivity values that are greater than 90 per cent. For example, if R is as high as 100 and the marker is present in 1 per cent of the population, the sensitivity of the marker is only around 50 per cent.
- Although sensitivity is not directly related to disease frequency, it can be shown from equation 4 that it is always less than or equal to m/p (even if PPV = 1). For example, if p = .01 and m = .001, sensitivity cannot exceed 10 per cent. Thus, when the marker is rarer than the disease, a limited proportion of cases will be associated with the marker.

Features of Genetic Marker Specificity

Table 2 depicts values of marker specificity for different m and R for a disease frequency of 1 per cent and shows some features of specificity:

- Specificity is essentially unaffected by changes in R but decreases with increasing m.
- Specificity of genetic markers remains high for a wide range of m, p and R.

Features of Genetic Marker Positive Predictive Value

From equation 3, several features of PPV can be noted:

- PPV increases with increasing relative risk.

TABLE 2—Relationship of Marker Specificity (a) to Marker Frequency (m) and Relative Risk (R) (for disease frequency p = .01)

	Marker Frequency			
	m < p	m = p	m > p	
Relative Risk	.001	.01	.10	.50
1*	.999	.990	.900	.500
2	.999	.990	.901	.502
5	.999	.990	.903	.503
10	.999	.991	.904	.504
100	.9999	.995	.908	.505
R_{max}	1.0† (R_{max} = 111)	1.0†† (R_{max} = ∞)	.90909††† (∞)	.5556††† (∞)

*when R = 1, a = 1 − m.

†when m < p, $R_{max} = \frac{1-m}{p-m}$ and a_{max} = 1.0.

††when m = p, R_{max} = ∞ and a_{max} = 1.0.

†††when m > p, R_{max} = ∞ and $a_{max} = \frac{1-m}{1-p}$.

- In contrast to sensitivity, PPV decreases with increasing marker frequency. For example, when a disease frequency in the population is 1 per cent and the relative risk for the marker-disease association is 100, 91 per cent of individuals with the marker are expected to develop the disease if marker frequency is .001 but only 2 per cent of individuals with the marker will develop disease if the marker frequency is 50 per cent.
- For relative risks that are below 10, PPV is generally low (< 10 per cent) if disease frequency is .001 or less.
- PPV increases markedly with increasing disease frequency but can never exceed p/m even for large relative risks (shown in equation 4). Thus, for very common markers in the population, PPV remains generally low if these are applied to screening for rare diseases even if relative risks are very high.

Interplay of Marker and Disease Frequencies

The relative magnitude of marker and disease frequencies may have a tremendous impact on values of sensitivity and positive predictive value. Figure 1 shows variations in values of screening parameters by relative risk, for a disease frequency of .01 when the marker is more frequent than the disease (m = .10). While the sensitivity of the marker increases with increasing R, PPV is low and remains low even for large R (< .10). Figure 2 shows changes in values of screening parameters by relative risk for a disease frequency of .01 but where the marker is less frequent than the disease (m = .001). While PPV increases tremendously with increasing R, the sensitivity of the marker remains low despite changes in R and is always < 10 per cent. Figure 3 shows the situation where m and p are equal (both = .01). In this case, sensitivity and PPV are equal and both increase with increasing R. However, even when the marker and the disease are approximately equally frequent in the population, it can be shown (Table 3) that to achieve certain levels of PPV (or sensitivity) for very rare diseases, the relative risk is required to be astronomically high.

Discussion

We have illustrated how the interplay between marker and disease frequencies affects sensitivity and PPV. This interplay should be considered in evaluating the usefulness of a marker for particular screening situations. Even with a very

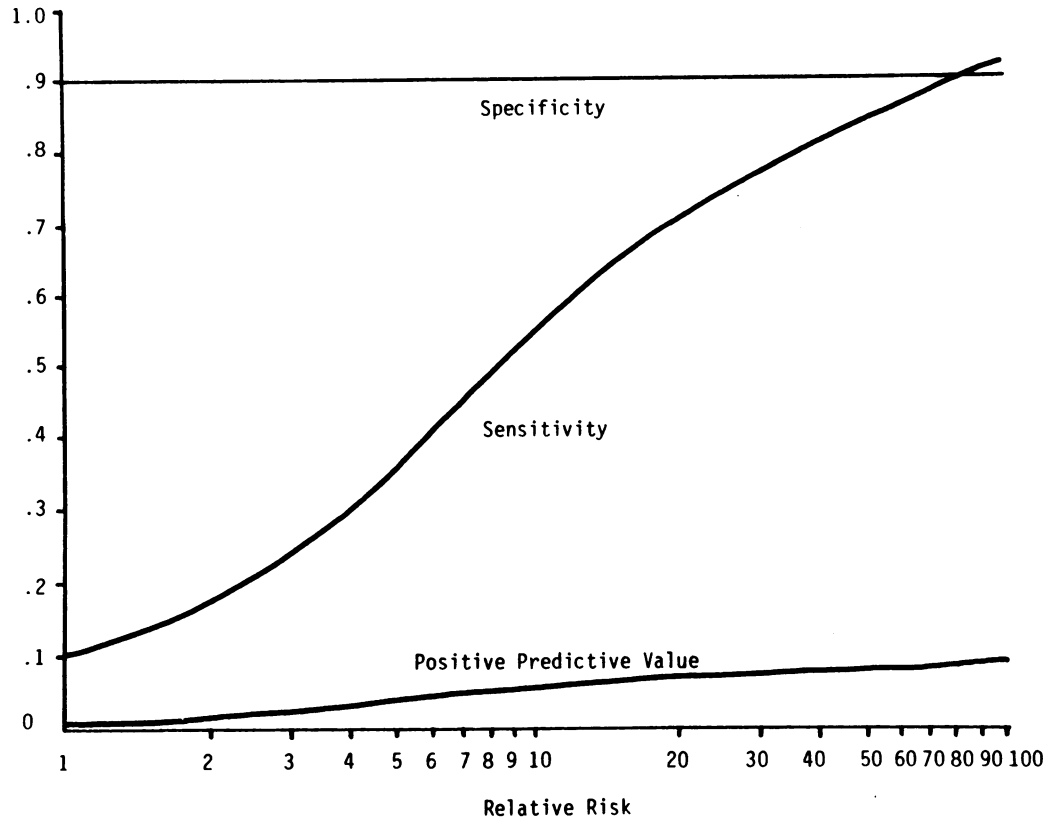

FIGURE 1—The Effects of Relative Risk on Sensitivity, Specificity, and Positive Predictive Value when m>p (m=.10, p=.01)

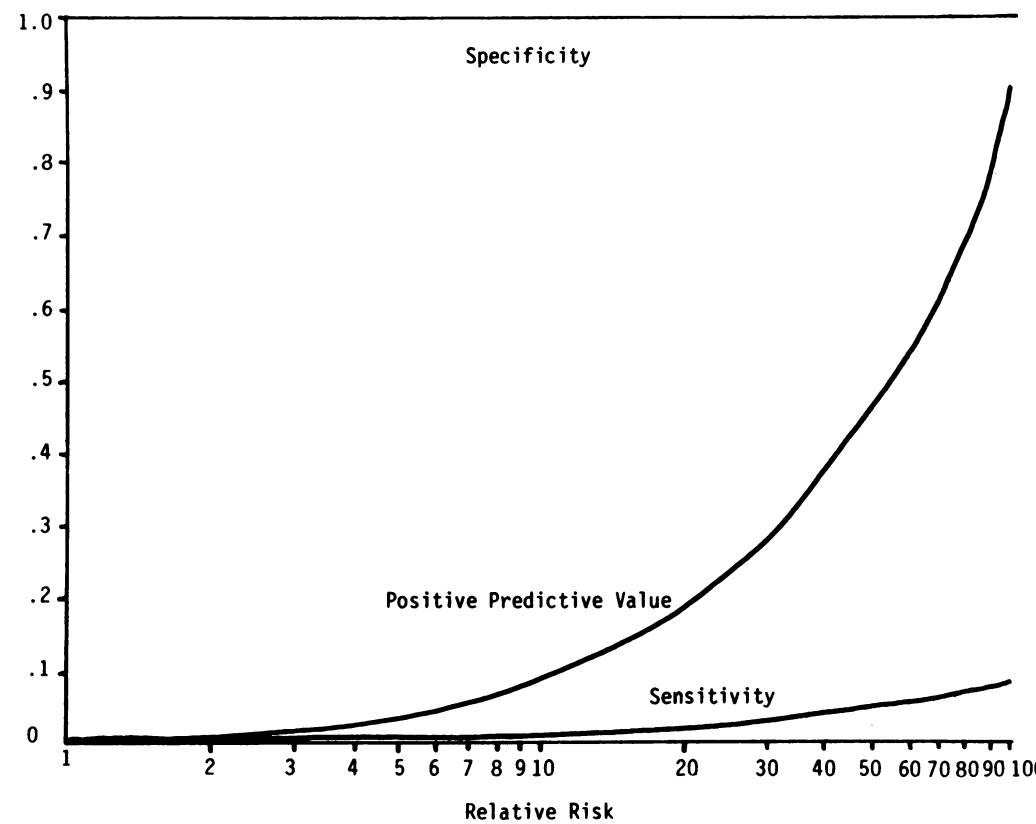

FIGURE 2—The Effects of Relative Risk on Sensitivity, Specificity and Positive Predictive Value when m<p (m=.001, p=.01)

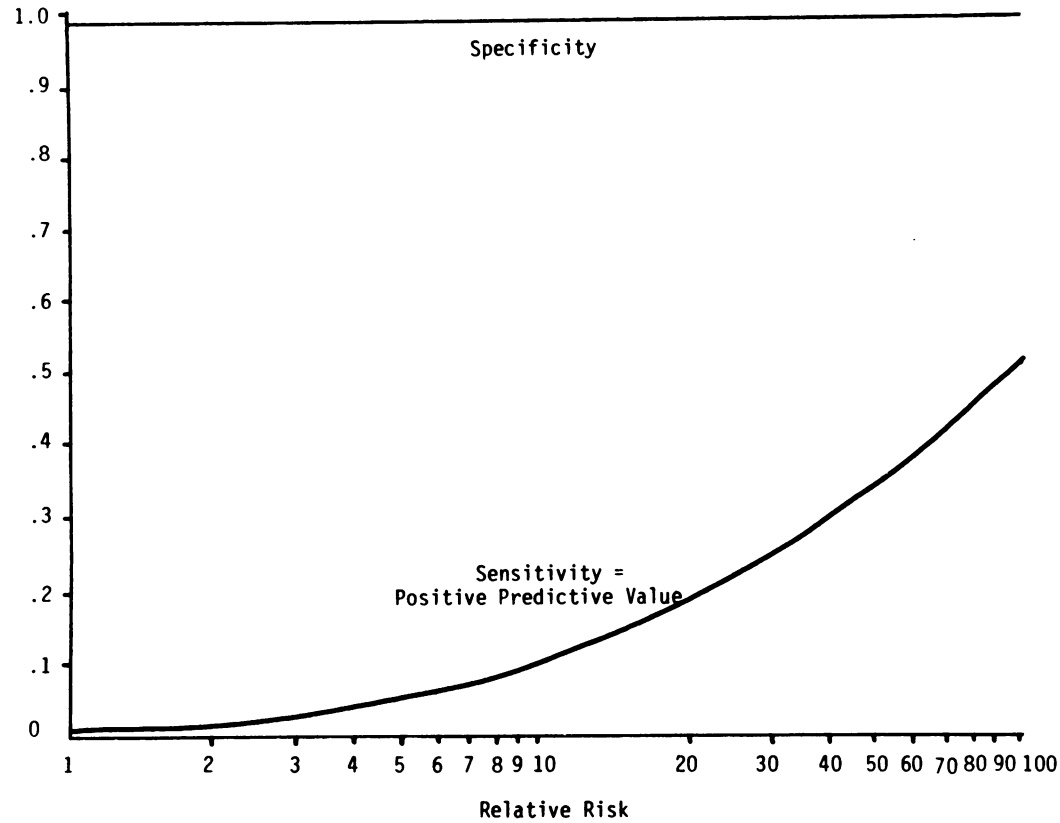

FIGURE 3—The Effects of Relative Risk on Sensitivity, Specificity and Positive Predicitive Value when m=p (m=.01, p=.01)

TABLE 3—Relative Risks Needed to Obtain a Given Level of Positive Predictive Value (= sensitivity) by Marker Frequency (when m = p)

PPV (= sensitivity)	Disease Frequency (= marker frequency)				
	.0001	.001	.01	.10	.50
.01	101	10.1	1.0	.09*	.01*
.10	1111	111	11	1.0	.11*
.50	9999	999	99	9.0	1.0
.75	29997	2997	297	27.0	3.0
.90	89991	8991	891	81	9.0

*When PPV < p, R < 1, marker is protective against disease.

high relative risk, the combination of a rare genetic marker and a common disease yields high predictive value but very low sensitivity. As an example, consider the PiZ phenotype (alpha-1 antitrypsin deficiency) and its relation to chronic obstructive pulmonary disease (COPD). Only 1 in 2000 to 4000 individuals in the population have the PiZ phenotype.[16,17] However, 5–10 per cent of the population may have COPD during their lifetime. If PiZ is to be used for screening purposes for COPD, although the PPV for such testing will be very elevated (> 95 per cent), its sensitivity can never exceed 1 per cent even in the face of the very large R (about 30).[17] On the other hand, the combination of a common marker and a rare disease yields low predictive value and high sensitivity. For example, HLA-B27 occurs in approximately 7 per cent of Whites,[15,18] but the disease frequency is around 2 per 1000. Even in the face of the very large relative risk reported (about 100),[19] the sensitivity of such testing is about 90 per cent, but its PPV is only about 2.5 per cent. The only situation where PPV and sensitivity are equal is when the marker and the disease are approximately equally prevalent in the population. However, as noted above, to achieve 50 per cent or more in sensitivity or PPV, the disease and the marker should be quite prevalent.

The suitability of a particular marker to be used for screening purposes depends on the objectives of screening. If the emphasis is placed on identification of individuals at increased risk—for example, in screening in the workplace—then PPV may be considered more important than sensitivity. In that situation, the choice of common markers as advocated by Omenn generally will result in a low PPV (high proportion of false positives). On the other hand, the use of a rare marker yields high PPV but has the obvious limitation of identifying very few susceptibles in the population (high proportion of false negatives).

This approach can also be applied to screening for nongenetic risk factors such as smoking, alcohol, diet, medical and family history. To evaluate any risk factor, estimates of the risk factor frequency, disease frequency and the relative risk must be available on the screened population. Such estimates can be obtained from epidemiologic studies of risk factors and disease associations and ought to be carefully used in relation to the goals of the screening program, cost/benefit considerations, acceptability of the screening and its application, and availability of efficacious and ethically palatable intervention measures for "positive" individuals.

ACKNOWLEDGMENTS

The authors thank Dr. Bernice H. Cohen for her comments, and Alice Rodman and Michele Neveu for technical help. Dr. Khoury was supported by the Lebanese National Council for Scientific Research.

REFERENCES

1. Morrison AS: Screening for Chronic Disease. New York: Oxford University Press, 1985.
2. Rothstein MA: Medical Screening of Workers. Washington, DC: Bureau of National Affairs, 1984.
3. Scriver CR: Genetic screening: implications for preventive medicine. Am J Public Health 1983;73:243–245.
4. Childs B: Genetic screening. Ann Rev Genet 1975;9:67–87.
5. Kazazian HH: Genetic screening: prerequisites and further considerations. *In:* Cohen BH, Lilienfeld AM, Huang PC (eds): Genetic Issues in Public Health and Medicine. Springfield: Charles Thomas; 172–192. 1978.
6. Calabrese EJ: Pollutants and High Risk Groups. New York: John Wiley, 1978.
7. Calabrese EJ: Genetic screening for hypersusceptibles in industry. Med Hypothesis 1981;7:393–399.
8. Holden C: Looking at genes in the workplace. Science 1982;217:336–337.
9. Omenn GS: Predictive identification of hypersusceptible individuals. JOM 1982;24:369–374.
10. Stokinger HE, Scheel LD: Hypersusceptibility and genetic problems in occupational medicine—a consensus report. JOM 1973;15:564–573.
11. Office of Technology Assessment: The Role of Genetic Testing in the Prevention of Occupational Disease. Washington, DC: OTA, 1983.
12. Murray TH: Warning: Screening Workers for Genetic Risk. Hastings Center Report, February 1983; 13:5–8.
13. Mourant AE, Kopec AC, Domaniewski-Sobczak K: Blood Groups and Disease. London, New York: Oxford University Press, 1978.
14. Bias WB: Genetic polymorphisms and human disease. *In:* Biocultural Aspects of Disease. New York: Academic Press, 1981;95–131.
15. Omenn GS, Motulsky AG: Ecogenetics: genetic variation in susceptibility to environmental agents. *In:* Cohen BH, *et al,* (eds): Genetic Issues in Public Health and Medicine. Springfield: Charles Thomas, 1978;83–111.
16. Kueppers F: Inherited differences in alpha-1 antitrypsin. *In:* Litwin SD (ed): Genetic Determinants of Pulmonary Disease. New York and Basel: Marcel Dekker, 1978;23–74.
17. Fagerhol MK, Cox DW: The Pi polymorphism: genetic, biochemical, and clinical aspects of human alpha-1 antitrypsin. Adv Hum Genet 1981;11:1–62.
18. Gilliland B, Mannik M: Ankylosing Spondylitis. *In:* Harrison's Principles of Internal Medicine, 10th Ed, Vol 2. New York: McGraw-Hill, 1983;1986–1988.
19. Sutton HE: An Introduction to Human Genetics, 3d Ed. Philadelphia: Saunders College, 1980;318.

Errata

In: Phillips K, Holm H, Wu AC: Contemporary table salt practices and blood pressure. Am J Public Health 1985; 75:405–406. On page 405, column 2, line 25, sentence should read: "Table 3 shows that in spite of the fact that high salt users were younger than low salt users, systolic blood pressure levels were higher in the *latter* group." (not *former* group)

In: Budnick LD, Ross DA: Bathtub-related drownings in the United States, 1979–81. Am J Public Health 1985; 75:630–633. In final paragraph of text on page 633, the sentence should read: "Because bathtub-related drownings occur suddenly and in a presumably protected environment, they may be more psychologically traumatic for families than are deaths from some other causes.[35,36]" The correction is the reference numbers 35 and 36, not 27 and 28 as printed.

In: King H, Li J-Y, Locke FB, Pollack ES, Tu J-T: Patterns of site-specific displacement in cancer mortality among migrants: the Chinese in the United States. Am J Public Health 1985; 75:237–242. On page 237, column 1, last line, "Fujianese" is the correct spelling (not Fujienese); page 237, column 2, line 15, "Guandong" is the correct spelling of the province (not Guangzhou). Also, in the column headings of Table 2, the righthand column should be "95% Confidence Limits**" to refer to the footnote.